DISTORTION OF SUBJECTIVE PROBABILITIES AS A REACTION TO UNCERTAINTY

By William Fellner

I. Observation of slanted probabilities and computation of the correction factor, 670. — II. The question of rationality, 677. — III. Two small experiments, 686.

I. Observation of Slanted Probabilities and Computation of the Correction Factor

Daniel Ellsberg in Santa Monica and, by coincidence, the present writer in New York,[1] have developed a piece of analysis which suggests a link between the nonadditivity of directly observable subjective probabilities and the problem of uncertainty. I shall try to use little space for telling the gist of the story in my particular way. Subsequently, I would like to say something about the question of whether Ellsberg's and my hypothesis is or is not compatible with rational behavior, and I will add a few words about two small experiments with Yale students.

Let us assume we can get a person to disclose that he assigns equal subjective probabilities to an event's happening and to its nonoccurrence. For example, we may ask him to choose whether he prefers to receive a prize of a specific size if an event will happen or to receive the prize if the event will not happen. He fills out an answer, and before looking at it we ask him whether for a small additional amount he would change the answer around. If he says yes, we conclude that the event and its "complement" have roughly equal probabilities to him. Denote the process which produces this event by X, the event itself by E_x, the complement (nonoccurrence) of the event by $\sim E_x$, and denote the subjective probabilities of specific events to our individual by the symbol P. Then, we infer from the test: $P(E_x) = P(\sim E_x)$.

One would at any rate be interested in knowing whether this judgment of equiprobability is strictly comparable with the judg-

1. The writer would like to express his gratitude for the opportunity which Columbia University gave him to work on this project, by inviting him to a Ford Visiting Research Professorship. He is indebted to many members of the faculty of Columbia University for very helpful discussions.

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

ment, "I don't mind whether I stake a prize on drawing red or drawing black from an unbiased deck of cards." The event E_x, which we considered before, may, for example, have been the rise of a stock price, or the victory of a political candidate, or the subsequent confirmation (or refutation) of an answer which the subject was required to fill out in response to a true-or-false question. We shall assume that well-established, "objective" frequency-probabilities do not apply to the process producing the event E_x, i.e., to the process X. On the other hand, we may select a process to which such probabilities do apply — say, the drawing of cards from a deck of guaranteed composition — and we shall call this process our individual's *standard process* (S). We shall assume that he has become conditioned to go along subjectively with all the objective probability characteristics of his standard process,[2] so that for any E_s the value of $P(E_s)$ is a subjective probability and a classical frequency probability at the same time. As was said a moment ago, one would at any rate be interested in knowing whether the statement $P(E_x) = P(\sim E_x)$ is strictly comparable with, say, the statement "red and black in the standard process are subjectively equiprobable," and quite generally with some statement of the form $P(E_s) = P(\sim E_s)$.

Even in cases of equiprobability this is an important question. Furthermore, the general concept of numerical subjective probabilities — the entire conceptual framework which has gradually developed along the Ramsey-de Finetti-Savage line[3] — needs doctoring if $P(E_x) = P(\sim E_x)$ is *not* a strictly comparable statement with $P(E_s) = P(\sim E_s)$. In this case $P(E_x) = 3P(\sim E_x)$ will surely not be strictly comparable with $P(E_s) = 3P(\sim E_s)$. Yet the methods we can use to establish the proposition that to our individual the rise of a stock price seems three times as probable as the failure of that stock price to rise *involve comparing his $P(E_x)$ for the rise of the stock with the known probability of some event in his standard process.* His betting odds on the stock itself will not disclose his numerical probability judgment for E_x because these odds are influenced by the shape of his utility function (i.e., by decreasing or increasing mar-

2. It will be shown on pp. 683–84 that this is a testable characteristic of an individual.

3. See Frank P. Ramsey's essay, "Truth and Probability," in *The Foundations of Mathematics* (posthumously published essays; London: Harcourt Brace, 1931); Bruno de Finetti, "La prévision: ses lois logiques, ses sources subjectives," *Annales de l'Institut Henri Poincaré*, Vol. VII (1937); and Leonard J. Savage, *The Foundations of Statistics* (New York: John Wiley, and London: Chapman and Hall, 1954).

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

ginal utility). But if we could disregard the distortion or noncomparability to which I shall turn in a moment, we could postulate that a rise of the stock price has probability 0.75 to him, and its failure to rise[4] has probability 0.25, in the event that he satisfies the following condition: he is indifferent as between (1) receiving a prize which is made contingent upon a rise of the stock price, (2) receiving the identical prize contingent upon his drawing a card belonging in any one of *three* of the existing four suits. This statement is a straightforward application of Savage's theory of testable numerical subjective probability. It is an application to a reasonable person, whose standard process for determining his numerical subjective probabilities can be chosen in such a way that his subjective probabilities in that process should equal the objective ones. The theory doubtless postulates that probability judgments for various processes are strictly comparable with each other. One cuts across processes, so to speak, without fear of distortion.[5]

I will suggest, as does Ellsberg, that subjective probability judgments relating to various processes are not strictly comparable. The same general idea was expressed on an earlier occasion by N. Georgescu-Roegen.[6] The use of a correction factor seems to me suitable for *restoring* the comparability which is essential to the Savage type concept of subjective probability. The correction factor, which will be described presently, measures the reaction to "uncertainty" in one of the essential senses of this word. In general, the factor is different for different nonstandard processes, and it depends also on the magnitude of the prize used for subjective probability tests.[7]

Let us examine the difficulty. I will illustrate with the simple case of equiprobability, although the nature of the problem is the same for probability ratios other than $0.5 \div 0.5$.

4. The phrase "fails to rise" stands here, of course, for "either stays unchanged or declines."

5. It is not literally true that *all* methods ever proposed for detecting numerical subjective probabilities raise the problem in this particular form. But what is true is that the only other method of which I am aware, the method suggested by the great pioneer, Frank Ramsey, causes a difficulty which is prohibitive from our point of view. In the event of incomparability, the method based on Savage's theory registers a distortion clearly, and enables us to face the problems created by this distortion. The Ramsey method, on the other hand, covers up the difficulty. This is because the method establishes numerical subjective probabilities for each process within the process itself, thus merely *implying* strict comparability for the probabilities applying to all processes.

6. See *Expectations, Uncertainty and Business Behavior,* ed. Mary Jean Bowman, Social Science Research Council (New York, 1958).

7. See n. 1 on p. 677.

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

When an individual says that red and black from an unbiased deck are equiprobable, he knows he makes this judgment in possession of *all the available information* concerning the matter. This kind of situation is at one end of a range. Most probabilistic problems of real life are rather far removed from this limit. When our individual proves indifferent as between making a prize contingent upon E_x or $\sim E_x$, he usually knows that *he is ignorant of much potentially available information.* Consequently, his state of mind could be described by a statement such as this: "the probability of E_x might be anywhere between 0.7 and 0.3, the probability of $\sim E_x$ anywhere between 0.3 and 0.7; I would need further objective guidelines of the right kind to narrow this range." Now, such a person may be induced to do things from which we wish to *infer* the probabilities $0.5 \div 0.5$ for him. For example, the equiprobability test which was described in the beginning of this paper can accomplish this. But this merely means that if he assigns a probability of somewhere between 0.7 and 0.3 to E_x and to $\sim E_x$ alike, then in the subjective probability test he will prove indifferent as between E and $\sim E$. *This need not imply that he is equally willing to say "the range 0.7 to 0.3 can be represented by the figure 0.5," as he is to say "red or black from an unbiased deck is 0.5."*

It follows that our individual may prove indifferent as between staking a prize on E_x and staking the identical prize on *a standard-process event which has the probability 0.3*; and that his indifference point may be the same for $\sim E_x$ and an event in his standard process (provided of course that we feel out his indifference values for E_x and $\sim E_x$, respectively, without his knowing on which of these two expressions of willingness he will have to make good). That is to say, he knows for sure that either E_x or $\sim E_x$ will obtain, he is reluctant to judge whether the probability ratio is $0.7 \div 0.3$ or $0.3 \div 0.7$ or anything in between, and he is of such a disposition that in these circumstances *he assigns the weight 0.3 to that occurrence on which he stakes his fortunes* (E *or* $\sim E$). Alternatively, he could be the kind of person who in such circumstances assigns the weight 0.7 to the event in which he acquires an interest. In general, this figure could be anywhere between 0.3 and 0.7. While I shall go on illustrating the problem with the number 0.3, all that matters here is that the number need not be 0.5.

Let me repeat that we, as observers, will wish to infer that this individual's "true" subjective probabilities for E_x and $\sim E_x$ are $0.5-0.5$. This involves introducing the concept of *uncorrected* and of *corrected* subjective probabilities. The uncorrected probability

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

of E_x is 0.3 to our individual, since he is indifferent as between making a prize contingent upon E_x on the one hand, and an event in his standard process, on the other, whose probability is 0.3. In this particular case we have assumed that the uncorrected probability is 0.3 also for $\sim E_x$, although we will, of course, very frequently get a different figure for $\sim E_x$. In general, these uncorrected probabilities need not add up to unity. In our particular case they add up to 0.6. We now inflate or deflate, as the case may be, the uncorrected probabilities of E_x and of $\sim E_x$ in such a way that their sum should equal unity, and that the ratio of the corrected probabilities should be the same as was that of the uncorrected ones. In our case we multiply the uncorrected probabilities of E_x and of $\sim E_x$ by the ratio 5/3. This gives us the corrected probabilities.

It is admittedly somewhat arbitrary to preserve in our correction procedure the *ratio* of the uncorrected probabilities defined as *hope-of-gain indifference points relative to an event in the standard process*, rather than as *fear-of-no-gain* indifference points relative to an event in the standard process. To be more specific: aside from the special case of equiprobability, the ratio of the uncorrected probabilities is one thing if we place the emphasis (as we have done) on the fact that our individual slants down his genuine $P(E_x)$ when a prize is staked on E_x and slants down his genuine $P(\sim E_x)$ when a prize is staked on $\sim E_x$; the ratio would be another thing if we had wanted to place the emphasis on the fact that the same individual slants up his genuine $P(E_x)$ when the prize is staked on $\sim E_x$, and slants up his genuine $P(\sim E_x)$ when the prize is staked on E_x. These are two different concepts of slanted or uncorrected probability, where one concept relates to a hope of gain and the other to the fear of not making a gain.[8] But it will be shown that the slanting tend-

8. In our simple case of equiprobability the "hope-of-gain" type uncorrected probability ratio for the event and its complement was 0.3 to 0.3; the "fear-of-no-gain" ratio was 0.7 to 0.7; and it therefore made no difference which ratio we carried over into the corrected (additive) probabilities. If, however, the hope-of-gain ratio had been, say, 0.6 to 0.2 (instead of 0.3 to 0.3), then the fear-of-no-gain ratio would have been 0.8 to 0.4, and it would have made a difference whether we decided to preserve the hope-of-gain ratio (thus obtaining the corrected probabilities 0.75 and 0.25 for the event and its complement), or decided to preserve the fear-of-no-gain ratio (thus obtaining the corrected probabilities 0.66 and 0.33). My suggestion that in such cases it is more convincing to preserve the hope-of-gain ratio is based on the following. A person with no utility or disutility of gambling would interpret the value of the conditional offer of a prize as being the *precise equivalent* of his genuinely probabilistic expectations, say of *0.75 times the value of the prize itself if this is staked on the event, and of 0.25 times the*

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

encies result from the "utility or disutility of gambling"; and it seems to me more convincing to relate this phenomenon to an expected change in the individual's wealth (i.e., to his hope of gain) than to the possibility of his staying in his initial position (i.e., to the fear of no gain). However, more views than one could be maintained with respect to the problem of the correction factor for restoring additivity. What ultimately matters is merely the fact that nonadditivity measures a distortion of the genuine probabilities.

If, and only if, by probabilities in nonstandard processes we meant *uncorrected* probabilities, would the operational utility function of our individual — the kind of Bernoulli function or von Neumann-Morgenstern function which Mosteller and Nogee could have found for him — be identical when established in processes such as X as when established in the standard process S. The *uncorrected* probabilities do measure indifference points as between the various processes. The use of "uncorrected probabilities" for the establishment of operational utility functions in nonstandard processes would make it immediately obvious that we are not postulating the maximization of "mathematical expectations," since computing mathematical expectations from "uncorrected probabilities" would be faulty procedure. The uncorrected probabilities violate essential postulates of the probability calculus. These apparent probabilities — psychological weights used in decision-making — result from the slanting in one direction of the probability of the event on which the individual happens to stake his fortunes, and from the slanting of the complement's probability in the other direction. If, on the other hand, we take our *corrected* probabilities, then our particular individual will show a *higher* utility function for the standard process than for the process X. This is because if the individual "barely accepts" a bet (if he proves to be at the boundary of his bet-acceptance region) when promised the gain g with probability p against a possible loss of one unit of utility with probability $1-p$, *then the utility of g to him is* $\frac{1}{p} - 1$, with the no-bet position as the zero point of his utility

value of the prize if this is staked on the complement. It seems convincing to me to use these hope-of-gain probabilities as points of departure when allowances must be made for the utility or disutility of gambling. It does *not* seem convincing to me to use the fear-of-no-gain probabilities as points of departure, because in the absence of any utility or disutility of gambling the individual would act as if there were no "fear of no gain." This is the conception which underlies my correction procedure in the text, but the analytical results presented in the text do not depend on this procedure in any essential way.

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

scale.[9] We have seen that our particular individual is willing to take a bet at lower probabilities of gain in his standard process than are the corrected probabilities at which he takes a bet in the process X, and we now have shown that to lower values of p there correspond higher utilities. When the corrected subjective probabilities are used for various nonstandard processes, the differences between the correction factors applicable to these will lead to the establishment of different operational utility functions for all these processes. The discrepancy between the operational utility functions found in different processes expresses the same reaction to uncertainty which is measured by the distortion (nonadditivity) observable on the uncorrected subjective probabilities.

Instead of postulating that individuals maximize the mathematical expectations of the utility of wealth, we should therefore postulate that they maximize a term which is analogous to the mathematical expectations of utility but is, at least for some processes, based on psychological weights that are in the nature of *distorted* probabilities. An alternative statement which is also acceptable for some purposes is that the *true mathematical expectations of an impure kind of utility are being maximized*, where the maximized concept is rendered impure by the reaction to uncertainty, i.e., by the utility or disutility of gambling in each process. But this alternative statement which implies the use of undistorted (or corrected) probabilities merely means that for pragmatic reasons we may sometimes wish to channel the impurity into the utility concept itself rather than catch it at the level of the weighting system. In this case the distortion of

9. This in turn is because, $U(g)$ standing for the utility of g, we postulate the relationship

$$pU(g) = 1-p.$$

Hence $$U(g) = \frac{1-p}{p}.$$

For marginal utility it follows that

$$U' = -\frac{1}{p^2}\cdot\frac{dp}{dg}.$$

The function $p = p(g)$ of which $\frac{dp}{dg}$ is the derivative is the function which gives us for a promised gain in the amount of g the minimum probability of winning required for *barely inducing* the subject to put up an amount whose utility is defined as one utile. ("Barely inducing" is not absolutely precise because $p = p(g)$ really marks the *boundary* of the region of bet-acceptance, and thus it is not literally p but any probability *ever so little in excess of* p which dependably induces the subject to accept the bet.)

We have here "measurement of utility up to a linear transformation." The unit of measurement and the zero point must be chosen arbitrarily.

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

the probabilities gives the appearance of a distortion of the utility concept rather than of the probabilities.

II. The Question of Rationality

Is the phenomenon of nonadditivity — the phenomenon which calls for our correction factor — compatible with rationality?

This is no doubt a subtle question. It is also an important question because some of the freaks that can be found in the observed behavior of almost all people call more for general qualifications than for efforts to revise received theory.

I will distinguish three aspects of the rationality problem, and I will be brief on the first two. The third, I think, is the most essential for the appraisal of the present problem, although all three are interrelated.

In the first place, there is the question whether at any given time our individual, with his nonadditive uncorrected probabilities, does or does not satisfy certain consistency requirements within this type of behavior. For example, if an individual is indifferent as between staking a prize on E_x and staking it on $\sim E_x$, and if the uncorrected subjective probability of E_x is say 0.3 to him, then the uncorrected subjective probability of $\sim E_x$ should also be 0.3 to him. I shall make some suggestions concerning such consistency requirements in the footnote below. If he satisfies requirements such as these, he passes one of the essential tests of rationality.[1]

1. For any nonstandard process, denote by P_u the uncorrected subjective probabilities found by indifference tests as between staking a prize of given size on an event in the nonstandard process in question and staking it on some event of known probability in the standard process. For the *standard process* S we do not distinguish between uncorrected probabilities (P_u) and corrected probabilities (P_c), but we postulate that for S the "classical" or "objective" frequency-probabilities (P) coincide with the subjective ones. For the *nonstandard processes* X, Y, *etc.*, the P_u terms for an event's happening and its not happening do not generally add up to unity.

Denote by P_c the corrected subjective probabilities;

denote by E_s, E_x, E_y . . . events happening in the various processes expressed by the subscripts; by $\sim E_s$, $\sim E_x$, $\sim E_y$ . . . the complements of these events [so that $P(E_s) + P(\sim E_s) = 1$ but $P_u(E_x) + P_u(\sim E_x) \gtrless 1$, even though it should be remembered that $P_u(E_x U \sim E_x) = 1$];

denote by σx, σy . . . the sum of the uncorrected probability of an event and of its complement in processes X, Y, . . . [so that $\frac{P_u(E_x)}{\sigma x} + \frac{P_u(\sim E_x)}{\sigma x} = 1$ and quite generally P_c is defined by $\frac{P_u}{\sigma} \equiv P_c$]; and

denote by $P_w(E_x) = P_w(\sim E_x)$ an equiprobability found directly *within the process* X, i.e., without reference to the standard process, by means of estab-

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

Secondly, there is the question whether, if we observe in him the trait of nonadditivity, he is or is not likely gradually to lose this

lishing indifference as between staking a prize on E_x and on $\sim E_x$; further, by analogy, use the symbol $P_y(E_x) = P_x(E_y)$ for equiprobability established directly across nonstandard processes, i.e., by indifference as between staking a prize on an event produced by the nonstandard process X and an event produced by the nonstandard process Y.

Then a consistent probability structure requires

(1) if $P_w(E_x) = P_w(\sim E_x)$, then and only then: $P_u(E_x) = P_u(\sim E_x)$

(2) if $P_y(E_x) = P_x(E_y)$, then and only then: $P_u(E_x) = P_u(E_y)$.

Both conditions are testable. If these conditions are not satisfied, the concept $P_c = \frac{P_u}{\sigma}$ lacks the properties which would make it a useful concept of probability.

The foregoing does *not* make it a consistency requirement that P_u should be *invariant to the magnitude of the prize* used for establishing the uncorrected (or preliminary) probabilities! However, it obviously is a reasonable consistency requirement that all $P_w(E) = P_w(\sim E)$ relationships should be invariant to the magnitude of the prize used for testing, and this implies that $\frac{P_u(E)}{P_u(\sim E)}$ *should be invariant to the magnitude of the prize for any E*, and, of course, $P_u(EU\sim E) = 1$ also should stay valid for any prize. In other words, what need *not* be invariant to the magnitude of the prize is $\sigma_x, \sigma_y. \ldots$

Note also that we did not make it a consistency requirement that the σ terms should be invariant to how far we carry the partitioning of the entire possibility set. That is to say, not only do we recognize that $P_u(E_x) + P_u(\sim E_x) \gtrless 1$, while naturally $P_u(E_xU\sim E_x) = 1$, but we recognize *also* that for any mutually exclusive events $E^1{}_x$, $E^2{}_x$ such that $E^1{}_xUE^2{}_x = E_x$, $P_u(E^1{}_x) + P_u(E^2{}_x) \gtrless P_u(E_x)$ (and the same statement can be made for events produced by all processes other than S). The σ terms, however, are defined in such a way as to assure additivity to unity for the *entire* possibility set, regardless of how far the partitioning is carried in any given instance.

This amounts to the following. Preliminary subjective probability tests involving events produced by different processes are influenced ("distorted") by the "utility of gambling" which depends in part on the magnitude of the amount involved in the bet. Tests cutting across various processes are influenced also by the dependence of the utility of gambling on how far the partitioning of the possibility set into various subsets (individual possibilities) is carried. Fortunately, it is possible to anchor such a distorted system to a point which stays fixed, and then there exists a reasonable method for undoing the distortion. The point that stays fixed (undistorted) is the ratio of the probability of an event's happening to its not happening. For this relationship the utility of gambling cancels out.

This last sentence implies that if a person consistently prefers to bet on peace rather than on war, then this is indistinguishable from his assigning greater probability to peace than to war. One might wish to argue that this is a weakness (but then an *inevitable* weakness) of the operational concept of subjective probability. It is this "weakness" which prompted Ramsey to limit the concept to "ethically neutral events." But, what is an ethically neutral event?

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

trait *as he gets used to the uncertainty with which he is faced.* To this I cannot suggest a wholly satisfactory answer. The unsatisfactory part of the answer is that the rather primitive methods of which I am aware are not well suited to testing the same subject on several successive occasions. If tests are repeated, he is practically certain to "find out" the experimenter in various respects, and the consequences of this are not likely to be very similar to those of a subject's getting used to the more involved processes of nature or of social life. Already the equiprobability test described in the opening passages of this paper involves using a trick which the subject must not anticipate. In general, his subjective probabilities for the problems with which we face him are certain to change as he begins to read a pattern into the problems the experimenter presents to him. Finding situations to which the individual develops the reaction of at least rough subjective equiprobability plays a large role in such tests. The task is difficult at any event, and it might prove practically unmanageable for a subject who is developing increasingly strong hunches, right or wrong ones, concerning the procedures of the experimenter. Even where this can be made vaguely analogous to the acquisition of knowledge about the processes of real life, the fact still remains that observing the individual's reaction to some degree of uncertainty to which he has gotten used is a different task from observing his reaction to rapidly changing degrees of uncertainty.

On the other hand — and this perhaps is the more satisfactory part of the answer — the subject is quite likely to realize that the kind of situation with which he becomes faced in an experiment has the same *general* characteristics as many situations with which life has already confronted him. This should help, even though the special circumstances are inevitably different from those to which the subject has gotten used. To take an illustration, in a half-conscious way many students have probably asked themselves on some occasions whether a true-or-false question about which a candidate has next to no information does or does not create the same problem as a pure-chance situation with known characteristics. If in a test we let him choose between answering such true-or-false questions and staking his fortunes on "red or black from an unbiased deck," and if we give him time to make his choice carefully, then in one sense we *do* and in another we *do not* test him "for the first time." Among colleagues who have commented to me on the problem of this paper, Gerard Debreu has placed emphasis on the inconclusiveness of testing a subject just once (even if we test him with several problems).

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

These doubts can perhaps be answered in a semisatisfactory fashion, but not in a fully satisfactory one. I will let it go at that, since the third question, to which I now turn, raises a very similar problem from a somewhat different angle.

Say that our individual meets the consistency requirements within his type of behavior and that over a reasonable period of time he persists in this behavior. In other words, we assume now that he consistently prefers "red or black from an unbiased deck" to answering a question involving available information none of which he happens to possess. He could also have the contrary preference, but a preference for the deck of known composition is what we shall now assume. To our individual an objective (classical type) probability is psychologically equivalent to a *more* favorable probability of the typically subjective kind; or, differently expressed, if the probabilities are the same for the two kinds of processes, he will take the objective probabilities even for a lesser reward. The subjective probabilities about which we are speaking here are, of course, our "corrected" ones, since by definition the individual is indifferent as between objective and "uncorrected" subjective probabilities.

So our individual is consistent and, as long as we leave him alone, he has stable preferences. But the question still remains whether leaving him alone is not like leaving an otherwise rational person alone who consistently prefers three dollars to *quatre dollars*. This latter person needs to be supplied with a dictionary rather than to be assured of our respect for his preference scales. He is making a mistake. Is not our individual, too, making a simple mistake? What if we explained to him that when for E_x as well as for $\sim E_x$ he shows an uncorrected probability of 0.3, then he is *really* saying that he assigns to both these possibilities the probability 0.5, and that when for E_x he shows the uncorrected probability 0.6 and for $\sim E_x$ the uncorrected probability 0.2, he is *really* assigning 0.75 and 0.25. Why does he not correct his probabilities for himself in actual practice, instead of forcing us to distinguish between directly observable "uncorrected" probabilities and theoretical "corrected" ones? Given his probabilistic appraisals, he would maximize his mathematical expectations if he were willing to make the correction for himself.

There is room for more opinions than one on this aspect of the problem, and I have hopes that in the present symposium more than one opinion will be expressed. The most interesting comments on this point I received from Thomas Schelling. Meanwhile, I will say

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

why I feel inclined to the view that our individual, with his observable "uncorrected" probabilities, is not simply "making a mistake."

My proposition here is this: if, when tested in a game with purely "objective" (classical type) probabilities, our individual is entitled to some amount of utility or disutility of gambling, without this behavior being called a "mistake," then he is also entitled to quite a bit of the effect which Ellsberg and I are attributing to him. Let me try to justify this position.

Say that some particularly sharp insight enables us to make the intuitive (nonoperational) assertion that in a specific range of potential gains and losses our individual's "pure" (or "neoclassical") utility-of-wealth function is roughly linear. I think the reader will agree that for many individuals this assertion may have strong intuitive plausibility, within specific ranges of betting. Equally great plausibility attaches to the statement that such a person may nevertheless insist on appreciably more favorable probabilities than 0.5–0.5 to put up five dollars for a potential gain of five dollars, even when he is drawing cards from a deck of guaranteed composition. Say he insists on a probability of 0.6 in his favor when trying for red and on the same probability in his favor when trying for black. One might explain this by saying that in the event of a loss he does not expect to be fully compensated *in the psychological sense* by knowledge of the fact that he had an *equal* chance of winning, but *does* expect to be compensated by knowledge of the fact that he had a 0.6 chance of winning. Few would feel like explaining to him that he is making a "mistake."

His mathematical expectations of the utility of wealth would, of course, be greater if, when faced with bets carrying probabilities of between 0.5 and 0.6 in his favor, he did not reject them; *but he has valid reasons for not being guided by mathematical expectations alone.* No matter how many rounds of various "games" he is going to play in his lifetime, he will end up not at mathematical expectations but at a higher or at a lower value. This gives rise to a problem of "uncertainty." Each "present round" in which he participates has traits of a marginal item, in that the final outcome of a long sequence of rounds will be higher or lower by the amount corresponding to the outcome of the present round. Our individual surely need not be irrational or uninformed to ask himself each time the question whether, in the event of the loss of the money he puts up, he would or would not be fully compensated by the knowledge that he could have made

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

a gain. It is not a requirement of rationality that a person whose "pure" or "neoclassical" utility-of-wealth function is linear should be indifferent to putting up a specific number of dollars for a potential gain of the same number of dollars at a 0.5 probability of winning. Needless to say, this does not mean that a rational individual may leave mathematical expectations out of account. In setting his terms for the acceptance of uncertainty he must be aware of the fact that if he is considering participation in *many* additional rounds at favorable odds, then the uncertainty which he faces relates to *deviations from higher mathematical expectations* than if he is considering merely one additional round.

The phenomenon responsible for the fact that individuals do not *maximize* the mathematical expectations of the "pure" or "neoclassical" utility of wealth is sometimes called the "utility of gambling" (positive or negative). This expresses itself in a reaction to uncertainty. A person may attach a psychic cost to moving himself back and forth along his utility-of-wealth function by specific risky decisions; or he may place a positive value on such decisions. We have described a situation in which the operational utility function of a person is concave from below (appears to indicate declining marginal utility), not because his "pure" or "neoclassical" utility-of-wealth function has this property but because the operational utility function shows here an impurity which enters through the *disutility* of gambling. In the more general case, the operational utility function shows some combination of the properties of the "pure" utility-of-wealth function with impurities caused by the utility or disutility of gambling. The impurity is likely to be present in many cases, and sometimes it is quite likely to predominate, although there exists no dependable operational method for separating the pure utility-of-wealth function from the impurity.

Now note that so far as concerns the "impurity," the individual's behavior may be expressed by the statement that even in his standard process he is *slanting* his probabilities in one direction or the other. Our individual, for whom we have assumed disutility of gambling, is slanting them downward *for the event on which he puts up money* and upward *for the event against which he bets.* This is why it takes 0.6 probability of winning to compensate him for 0.4 probability of losing in a bet in which the same quantity of utility-of-wealth can be gained as lost. In his decision-making process, he attaches the weight 0.5 to a favorable event whose probability is 0.6, and he

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

attaches the weight 0.5 to an unfavorable event whose probability is 0.4. Awareness of the fact that his probability judgments were more optimistic than his decision-making weights would make a loss that might occur more acceptable.

To be sure, *for the standard process* we need not go through the double operation of first detecting his slanted probabilities — i.e., probabilities that are distorted by the utility or disutility of gambling — and of subsequently correcting these probabilities for the slant. The fact that he goes along subjectively with the objective probabilities of the standard process (if this *is* a fact for him) can be detected by tests from which the slant cancels out. These are tests by which he discloses indifference as between accepting a prize contingent upon an event which possesses some objective probability in the standard process and any other event which possesses the same objective probability in the identical process (regardless of what the probability of the first event or sequence of events is). Since *within any one process* he slants all probabilities of given sizes in the same proportion whenever he stakes his fortunes on them, these equiprobability tests are not influenced by the slant. Once we know that he goes along with the objective probabilities of the standard process, we need not go through the operation of first finding slanted (or "uncorrected") probabilities and of then correcting these. Indeed, for the standard process itself it takes an "intuitively acceptable assertion" concerning the properties of the individual's "pure" utility-of-wealth function *to enable us at all* to make a statement about the precise extent of the impurity, and thus of the slanting. In the general case, the standard process confronts us with an unidentifiable degree of slanting of the probabilities applied to the pure utility of wealth, and it is pragmatically preferable to express this for the standard process not as a slanting of the probabilities but as an unidentifiable degree of impurity in the operational utility concept itself to which the probabilities are applied. In the general case, only the *differences* between observations relating to nonstandard processes and observations relating to standard processes can qualify as precise quantitative indicators of slanting, i.e., of *additional* slanting, since these differences are *wholly* attributable to an additional element of impurity in the nonstandard processes. But the fact remains that if to our individual gambling has disutility or positive utility, then even in his standard process he does something that is *equivalent* to slanting in one direction the probabilities of the events on which he

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

puts up money, and in the other direction the probabilities of the events against which he bets. He distorts the probabilities applicable to the "pure" utility of wealth.

What if the individual slants *more* in processes for which he has merely very tentative probability judgments than in his standard process? This, it seems to me, he is very likely to do. Say he has rather uncertain views about whether the price of a stock should be expected to rise or not to rise, and that he believes that more information could lead him to probability judgments lying anywhere between 0.8 and 0.4 for a rise and between 0.2 and 0.6 for the other possibility. It is quite reasonable to attribute to him the probability judgment 0.6 for a rise and 0.4 for the contrary, but in this case it will almost certainly require our correction factor to discover these probabilities. A person with a downward-slanting tendency will usually find it very tempting indeed to slant downward *from the median of the range 0.8 – 0.4, that is, to slant below 0.6* if he is betting on a judgment which has a spread of such size from the outset. He will usually find this appreciably more tempting than to slant the precise probabilities of his standard process; and it is hardly open to doubt that slanting takes place in standard processes, too. In the nonstandard processes he distorts the probabilities applicable to that operational utility concept which can be detected in the standard process.

Moreover, while there presumably exists a personality type which can be gradually conditioned not to slant probabilities in standard processes, many people belonging in this type may not lose their slanting tendency in a number of other processes. After a while, the specific instances in which the unfortunate decision *not to bet* in heads-or-tails has caused a *gain foregone* may start bothering this type of person as much as instances in which the unfortunate decision *to bet* has caused an *actual loss.* Each round is recognized as having basically identical characteristics, and the individual may acquire a more positive attitude to the rounds of the future. But this statement has no close analogy for a series of events each of which has different characteristics from the other. The kind of potentially available information which a person did not actually have the last time is frequently quite different from the kind of relevant information which he feels he lacks this time; and even acute regret over a gain foregone last time means little if the present situation is regarded as sufficiently dissimilar.

However, the reasoning here shows also that all this is a matter of gradations. Some nonstandard processes tend to acquire increasingly the properties of standard processes, as the outcome of succes-

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

sive rounds provides more information about the objective characteristics of the process. There exist nonstandard processes of which this is true, and others of which this is much less true (or, for all practical purposes, is not true). The question here is whether the "basic characteristics" of the process are simple enough, and sufficiently understandable, to permit us to derive significant clues from what is happening in individual instances. Where this is definitely not the case, it is perhaps more usual to speak of a series of events all of which are "unique," i.e., do not in any sense belong in one and the same "process"; but the point here is that the concept of such sets of unique events *shades over* into the concept of nonstandard processes, and that some of the latter do, while others do not, gradually acquire characteristics which are similar to those of the standard processes. Experience reduces uncertainty in some respects, but experience does not eliminate uncertainty. I feel we should not regard it as irrational if a person develops a reaction to uncertainty rather than is guided exclusively by mathematical expectations; and reactions to uncertainty, expressing themselves in slanting tendencies, are likely to be stronger for a good many nonstandard processes than for the standard processes.

The *difference* between the degree of slanting for "unique events" as well as for nonstandard processes on the one hand, and for standard processes on the other, cannot help but make an appearance in the subjective probability tests. It is precisely this *difference* which causes the distortion (nonadditivity) observable on the uncorrected probabilities. The reason is that the probability of 0.6 for the rise of a stock and the probability of 0.4 for the complement of this event is not detected directly *within* the stock process.[2] These probabilities are detected by discovering that the individual is indifferent as between making a prize contingent on an event relating to the stock price, on the one hand, and on an event of known probability in his standard process, on the other. *But if this is how we detect his subjective probabilities, we shall first find the uncorrected (nonadditive) probabilities for the stock because the result of this procedure will be influenced by the fact that the individual slants more in the stock-exchange process than in the standard process.* The correction factor needed for restoring additivity to unity will measure the direction and the size of the difference in the degrees of slanting.

The difference between the degree of *downward slanting* in nonstandard processes on the one hand and standard processes on the other is likely to be increased by knowledge of the fact that in most

2. See pp. 671–72 and the qualification added in n. 5 on p. 672.

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

cases the individual is representing other people as well (some institution, his family, etc.). Upward-slanting inclinations, too, are likely to be greater for "uncertain" judgments than for standard processes. But it is possible that this difference is reduced rather than increased by the existence of others in the background *for whom* the risky decisions are in part made.

A final word about the bearing of our problem on the concept of "uncertainty" versus that of "risk." Should we conclude that an individual views a situation as one entailing uncertainty rather than risk *only* to the extent that "gambling" has positive or negative utility to him? One might argue that aside from the utility of gambling — aside from the phenomenon which creates the slanting tendency — the problem is one of risk rather than uncertainty, since aside from this phenomenon, the individual makes his decisions probabilistically, on grounds of mathematical expectations based on his subjective probability judgments. This is one way of looking at the matter, but I think it would not be methodologically fruitful to *limit* the concept of "uncertainty" to this particular aspect of a more general problem. It is preferable to say that subjective probability judgments belong themselves in the theory of uncertainty, unless they happen to coincide with the public's "average" judgment which determines the current market values of goods and securities (including those of insurance policies). It is in the theory of profit that the Knightian distinction between risk and uncertainty has its most important application. Profit (positive or negative) results after deducting contractual and imputed costs on the basis of market valuations, not after deducting them on the basis of the subjective probability judgments of each individual. The quality of the probability judgments themselves as well as the nature of the slanting tendencies belong in the theory of decision-making under uncertainty.

III. Two Small Experiments[3]

In a first experiment we tested twelve Yale undergraduates who were assigned by the Yale Employment Service. From my point of view, the most obvious characteristic of the members of this group was that in the region in which they were tested they rather generally showed a "risk-loving" behavior. Mosteller and Nogee would have

3. The writer would like to express his appreciation for Mr. Thomas Synnott's valuable assistance at the experiments. It would have been very much more difficult to prepare and perform the tests without Mrs. Rosemarie Arena's constructive advice and help.

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

found that their operational utility functions were typically either linear or somewhat convex from *below*. The amounts involved in the risky decisions were such that it seems more reasonable to attribute the observed behavior to some amount of positive "utility of gambling" than to the properties of "pure" utility-of-wealth functions. For such a group one would not expect to find a downward slanting tendency for "uncertain" probabilities as compared with probabilities in standard processes. Perhaps one would expect a small degree of upward slanting.

The attitude to "uncertain" probability judgments was first tested by examining the willingness of the group to make a prize contingent on drawing the color of their choice from a deck whose composition was wholly unknown to them, as compared with their willingness to draw the color of their choice from a deck of guaranteed composition. Eight proved indifferent (except that the tests were, of course, unsuited to detect very small preferences). None had a preference for the "uncertain deck" when the odds were made slightly worse on this deck than on the guaranteed one. Four subjects had a preference for the guaranteed deck even at worse odds. One of these subjects had a rather small preference, and another a preference-ranking which was inconsistent at one point. The group was tested also for color bias. Eleven of the twelve did not seem to have any noteworthy bias of this sort. (This implies subjective equiprobability of red and black in the guaranteed deck; yet they may still have had hunches about the composition of the other deck.)

I also administered to this group a test by which it was possible to compare their willingness to answer a factual question with their willingness to draw the color of their choice from an unbiased deck. It was my intention to present to them a question the two possible answers to which seemed roughly equiprobable to most. In the first experiment I failed in this endeavor. But in spite of the fact that in the first experiment the majority seems to have had strong hunches concerning the right answer, I believe that the number of subjects who had no strong hunches may have been about five. Only one preferred making a prize contingent on a "standard process" (drawing numbers from a batch of guaranteed composition) to answering the factual question, when the odds in the standard process were slightly worse than fifty-fifty and the prize was identical. In the first experiment the subjects were not required to put up any money of their own.

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

In a second experiment the following devices were used to get the subjects into a range where they would show risk-aversion and would feel some amount of responsibility to others. A group of sixteen Yale students (mostly undergraduates), assigned again by the Yale Employment Service, were divided in two groups of eight subjects, in alphabetical order.[4] The first group was making the risky decisions, this time each in a separate room. Each was representing a member of the other group (they did not know in advance which member of the other group); 80 per cent of the gains and losses went to the other subject, 20 per cent had to be accepted by the subject who was making the decisions. The subjects who were represented by the decision-makers were informed about the nature of the problems which were being decided largely for them. Later, the members of the decision-making group had to give the others an account of their decisions. It was necessary for the decision-maker to put up amounts the loss of which could deprive the other subject of *almost all* of his wages for a rather lengthy Saturday session. On the other hand, the potential net gain over and above wages exceeded seven dollars for each subject.

Five of the eight decision-makers preferred "objective" probabilities (in a deck of guaranteed composition) to the "uncertain" probability of drawing even or odd from a batch of numbers whose distribution by even and odd was *not* guaranteed. The five in question had this preference even though their potential winnings were somewhat smaller if they chose the guaranteed deck, and the stake was the same. Those choosing the guaranteed deck were tested for color bias, and they did not seem to have any noteworthy bias for red or black.

The eight decision-makers were also given a test which in effect compelled them to rank their preferences for (1) the "objective" probability 0.5 of winning \$5.50, (2) the chance to win \$5.75 by answering a true-or-false question which was worded in such a way as to suggest that many students might know the answer but of which I hoped all or most to be ignorant, and (3) a chance to win \$5.75 by answering another true-or-false question of the same kind. Five of

4. There is no reason to assume that dividing the group in two halves alphabetically introduced a bias, except in one respect. I thought I had a better chance of evoking a reaction of subjective equiprobability to the true-or-false questions by confronting the "first half" rather than the "second half" with these questions. This is because I knew that two subjects in the "second half" had been in a curriculum which includes courses that *may* deal with the areas in which these questions fell.

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015

the eight subjects preferred Alternative (1) to at least one of the two true-or-false questions. Four of these five preferred Alternative (1) to both (2) and (3). Of the remaining three, who preferred both of the true-or-false questions to the "objective" fifty-fifty chance, two did *not* pass the equiprobability test for the question which they actually answered. In the test which was described on page 670 *supra*, they did not "switch" for a moderate inducement (even though they knew that I had no opportunity to discover what their initial answer was before offering them an inducement to switch). Consequently, two of the three who preferred both true-or-false questions to the "objective fifty-fifty" were far from interpreting their answers as having merely a fifty-fifty chance. Both gave the correct answers. The subjects who decided to bet on the "objective" probabilities in the guaranteed deck were not tested for subjective equiprobability on the true-or-false questions. Several of them may well have had weak or even strong hunches concerning the answers. The 0.5 objective probabilities on which they chose to bet may well have been handicapped to them (relative to the alternatives) by appreciably *more* than the equivalent of twenty-five cents. Therefore the group may on the whole be said to have shown a distinct preference for the objective probabilities.

All tests of the second experiment, and most tests of the first, were undertaken by asking the subjects to make their decisions in writing. They had much time to reflect.

I hope that the tests I have undertaken are "illustrative," and that in this sense they are not entirely useless. It would, of course, be unreasonable to claim that the significance of the phenomenon to which Ellsberg and I are calling attention has been proved by these experiments at Yale. All I would claim is that "I have seen the thing happen." I have seen it happen in circumstances in which one would expect it to happen on a priori grounds; and I have seen this mode of behavior become the characteristic of a mere minority in circumstances which on a priori grounds seem very unfavorable for detecting the phenomenon.

YALE UNIVERSITY

Downloaded from http://qje.oxfordjournals.org/ at Central Queensland University on May 26, 2015