ABSTRACT. From the point of view of the analysis of everyday reasoning about causality, the psychologists’ research has shown the relevance but also the insufficiency of Suppes’s definition of causality. Neither the probabilistic definition of causality nor the theory of rational choice can fully explain the data obtained with experiments on counterfactual reasoning and decisions. We have tried to show a link between judgements of causality and decision making. It seems necessary to analyze the type of decision on which an action is based in order to judge the causal role of this action. The data obtained by psychologists in different domains, such as perception and thinking, support both the realistic and the constructivistic positions discussed by philosophers in relation to the nature of causality.

INTRODUCTION

Suppes (1970) has made a central philosophical contribution to the probabilistic theory of causality (cf. Spohn, 1983, p. 75). In his first definition (1970, p. 12. Def. 1) he claims that:

– given two events, $B$ and $C$, where $C$ is chronologically following $B$, $B$ is a cause of $C$ if and only if the occurrence of $B$ raises the probability of the occurrence of $C$.

In this paper we will discuss this definition in relation to descriptive accounts of how people make causal attributions and causal judgements. The general question is about the possibility that the probabilistic definition of causality can capture the everyday sense of ordinary causal claims, as it was suggested by Suppes.

From this point of view, the above definition is very broad and has to be further specified in order to differentiate $B$, a genuine cause of $C$, from other events $B_1$, $B_2$, $B_3$, ..., $B_n$. In fact it is possible that the occurrence of these other events $B_1$, $B_2$, $B_3$, ..., $B_n$ also raises the probability of occurrence of the ensuing event $C$. The occurrence of these other events can be only statistically correlated with the increase of the probability of the occurrence of $C$. In that case $B_1$, $B_2$, $B_3$, ...,
$B_n$ can be considered as pseudo-causes and subdivided into spurious causes and indirect causes.

In the case of spurious causes we have a constant correlation between the occurrence of one event, e.g. $B_1$ and $C$, even if $B_1$ is not a genuine cause of $C$ (Spohn, 1983, p. 77). For example, a barometer always falls before a thunderstorm but that behavior is only an indicator of low pressure and not the cause of the thunderstorm. We could also say that it is a diagnostic and not a causal event in relation to the thunderstorm. Quattrone and Tversky (in Elster, 1985) have experimentally studied this distinction, showing that people behave as if a diagnostic event could also be causal. For example, their subjects ‘cheated’ on a medical exam by selecting actions (e.g., holding their hand in very cold water for an extended period of time) that they believed to be diagnostic of favorable outcomes (e.g. strong heart) even though they must have known that their actions could not possibly produce the desired outcomes.

Indirect causes seem also to follow the initial definition. The occurrence of $B$ increases the probability of the occurrence of $C$, but this relation is mediated by one or more intermediate events, $B_1, \ldots, B_n$, correlated with $C$. For example, imagine that the Bank of Italy decides to raise the Bank Rate ($B$). The banks cut down credit to companies ($B_1$) and the companies reduce investments ($B_2$). The fall in investments is followed by an increase in the unemployment rate ($C$). We can believe that this outcome was determined by the decision to raise the Bank Rate, nevertheless, between the first cause $B$ and the final outcome $C$ we have two other potential causes, $B_1$ and $B_2$. Both $B_1$ and $B_2$ were caused by the previous event and caused the following event.

Let us ask an economist which was the genuine cause of the increase in unemployment. He/she can answer in counterfactual terms: ‘If the Bank of Italy had not raised the Bank Rate, unemployment would not have increased. As a matter of fact, the higher Bank Rate caused the reduction of credit to companies, this reduction determined the ensuing fall in investments, and this fall caused the increase in unemployment. In conclusion, the decision taken by the Bank of Italy caused the increase in the unemployment rate.’

Suppes’s seminal work on the probabilistic definition of causality was followed by epistemologists and philosophers of science (for a review see Galavotti, 1983). In the last decade, since the fundamental work of Tversky and Kahneman (1980) and Kahneman and Tversky (1982)
the issue of causality has been studied in an experimental way. In order to examine the mechanisms behind our everyday naive judgements of causality the paradigm most utilized has been based on counterfactual reasoning.

CAUSALITY AND COUNTERFACTUAL REASONING

Let us imagine that the occurrence of $B$ causes the occurrence of $C$. We can express this belief with the conditional statement: 'If $B$ then $C$'. In this case the counterfactual reasoning in everyday life can take the form of a *modus tollens* syllogistic inference: 'If not-$C$, then not-$B$'. In a world in which the effect-event $C$ does not occur, the event-cause $B$ also has not occurred. The negation of $C$ implies the negation of $B$ if, and only if, there are no other causes of $C$. On the contrary, it often happens (as we have seen in the example of unemployment) that there is more than one event as a potential cause for a final outcome: 'If $B, B_1, B_2, \ldots, B_n$, then $C$'. In such a case, it is necessary to specify a hierarchy of importance among the potential causes.

Consider the unemployment example. An economist could suggest that the genuine cause, the one triggering the whole sequence of causes and effects, was the first decision, i.e. the decision of the Bank of Italy to raise the Bank Rate. Do other economists share this judgement about the most relevant cause of unemployment? In order to ascertain this point, it is sufficient to present them with the sequence of events and then to ask them to complete the following sentence: 'There would not have been an increase in unemployment if \ldots'. If other economists give the same judgement as the first one, the most frequent answer should be: '\ldots the Bank of Italy had not decided to increase the Bank Rate'. If, on the contrary, most economists disagree with the first one, we could have answers of this type: '\ldots the companies had not decided to stop investments'.

The counterfactual judgement can be used as a tool to investigate experimentally the beliefs about the causes of a final outcome of a sequence of events. When we want to know which one of the three events ($B\!-\!B_1\!-\!B_2$) was judged to be the most relevant cause to the final outcome, we can present a sample of economists with a counterfactual judgement. We can ask them to read the following story: 'A month ago, the Bank of Italy decided to raise the Bank Rate, the raising was
followed by . . . and so there was an increase in unemployment’, and then to list, in order of importance, the different ways in which the events in the story could be changed so as to produce a different outcome. In this way it becomes possible to ascertain the most important cause, the so called ‘genuine cause’, and also a hierarchy of indirect causes.

Let us imagine that we have now carried out the above experiment on the causes of unemployment. The majority of economists believe that there would not have been an increase of unemployment if the Bank Rate had not been increased. We can interpret the economists’ answers in two ways. On the one hand, we can hypothesize that event B has been judged the most relevant cause of C because the economists, in their organization of knowledge, have this strong belief: ‘It is a Bank Rate increase that usually leads to an increase in unemployment’. On the other hand, we can hypothesize that B was chosen simply because, in the list of potential causes, it was the first one. The first cause is the most relevant not for its specific content, but because it is the first cause that triggers the economic processes leading to the increase in unemployment.

Suppes’s definition of causality seems independent from content. In fact, the definition is based on the ratio of events occurring many times. According to Suppes, B is a cause of C if and only if, after the occurrence of B, C becomes more probable. In order to define the notion of cause it is sufficient to know the frequencies of C in relation to the situations in which B occurs and to the situations in which B does not occur. If the content of causes and effects is irrelevant we can use formal factors such as the frequency and the order of occurrence. In this last case we can hypothesize that the economists indicated the Bank Rate increase as a cause only because it is the first event of the sequence and not because they believe that this factor is the most relevant cause in the economic process.

More generally, we could say that the Suppes approach to causality can be used to discuss the empirical data obtained by experimental psychologists. These data show that people are very coherent and systematic in defining a hierarchy of importance when presented with a list of potential causes. So it is necessary to detect the criteria utilized in order to build this hierarchy of importance. If this hierarchy were based on the content of causes, it would be plausible to imagine that economists of different schools of thought, having different beliefs, could answer with different orderings. On the contrary, the Bank Rate
increase could be judged as the most relevant cause simply because it is the first one on the list, i.e. the one triggering all the economic process. In this last case the hierarchy of importance could be explained by a formal factor, i.e. the first position on the list of potential causes.

A second way to interpret the judgements of causality in formal terms, i.e. independently from the contents of causes and effects, is based on the assumption that our judgements are based on the computation of relative frequencies. This hypothesis is nearest to Suppes’s probabilistic definition, as it is based on the same factor. Suppes’s seminal work was followed by many researchers and, recently, has inspired some experimental work carried out by psychologists. In which way can frequency have an influence on the mental identification of causes? Following Suppes’s definition, we can hypothesize that frequency has a role in the sense that we search the list of causes for exceptional events and not for normal events (here ‘normal’ is defined generally, according to both a statistical and a normative meaning). For example, in the context of their ‘norm theory’, Kahneman and Miller (1986, p. 148) noted that: “The why question implies that a norm has been violated” and “requests the explanation of an effect, defined as a contrast between an observation and a more normal alternative”. They specified that a cause must be abnormal in the sense that it is not “a default value among the elements that the event (to be explained) has evoked” (p. 149). In order to understand the definition of ‘default event’ an example could be useful. Consider the way in which people tend to explain the peculiar behaviors of cars observed on the road: they refer to the drivers being young, elderly, or female. The default value for a car driver appears to be middle-aged male, and driving behavior is not explained by it. A default presumably involves a plurality but does not necessarily involve a majority. Being a non-default is therefore a less stringent requirement for a cause than being rare, even if often the default value corresponds to the majority. (Is it so in the case of American drivers?)

Kahneman and Tversky (1982) were the first ones to show experimentally that people, in making counterfactual judgements about causes, take into consideration normalcy vs. exceptionality of events. This idea has inspired further research (cf. Einhorn and Hogarth, 1986; Hilton and Slugoski, 1986; Turnbull and Slugoski, 1988) and has led to analysis of the organization of people’s knowledge. For example Hilton (1990) assumes that: “...in explaining an event to a competent adult, we would refer to individuating features of the case which cannot be presupposed
from general world knowledge, such as abnormal conditions, and omit to mention . . . ‘what’ can be presupposed”. (p. 67).

COUNTERFACTUAL REASONING AND SUPPES ON CAUSALITY

Suppes’s definition of causality anticipates the intuition of psychologists who have studied how people individuate causes in counterfactual contexts. We have seen that Suppes’s definition is formal and based on the concept of statistical frequency. If the individuation of causes has nothing to do with the content of causes, we could hypothesize the formal factors to be the ordering of occurrence of potential causes and their exceptionality. The role of these factors can be experimentally studied through counterfactual reasoning. Wells et al. (1987), to study the factor ‘order’, have utilized the counterfactual paradigm and a story of four events and a negative outcome. Kahneman and Tversky (1982) and Gavanski and Wells (1989), in order to study the ‘exceptionality’ factor, have utilized the counterfactual paradigm and a story in which the same events were presented either as exceptional or as normal.

Wells et al. (1987) presented their subjects with a scenario about a young man (William) who was going to a store to buy a stereo system on sale. His progress was impeded by four minor misfortunes: a fine for speeding, a flat tyre, a traffic jam and a group of senior citizens crossing the street. William arrived at the store 35 minutes after the sale had started only to find that the last stereo system had been sold just a few minutes before. Subjects read one of four versions of this story. Each version had a different ordering of the event sequence, and the four events were presented in each of the four possible positions in the scenario. Subjects were asked to list the different ways in which the events in the story could be changed to produce a different outcome of the story. It should be noted that the events of the story were linked in a causal way and that the modification of each event was sufficient to undo the outcome. The results showed that subjects tended to change the first event.

Kahneman and Tversky (1982) utilized a story in which Mr Jones was killed in a traffic accident when driving home after work. In one version of the story, Mr Jones left work earlier than usual, but he took his usual route. In the other version, Mr Jones left work at his usual time but took a different route home. The subjects’ tasks was to complete
a sentence that began: ‘If only . . .’. Even though in the two versions
the outcomes were the same, subjects produced different and specific
changes in the two cases. In the first version, most of them modified
the time variable. (If only he had left work at his usual time . . .) In the
second they modified the route. (If only he had taken his usual route
. . .) Kahneman and Tversky concluded that subjects undid “the accident
by restoring a normal value of a variable rather than by introducing an
exception” (1982, p. 205).

Gavanski and Wells (1989) confirm the results of Kahneman and
Tversky (1982) by showing that people will change events towards
normalcy to undo exceptional outcomes but will change events towards
exceptionality to undo normal events. Subjects read stories describing a
student (good vs. poor student) studying for an exam (pass vs fail exam).
With a $2 \times 2 \times 2$ design, the events were described as exceptional vs
normal. The results show that mental simulation processing is governed
by the correspondence, in terms of normalcy vs exceptionality, between
outcomes and prior events. Moreover, Gavanski and Wells (1989)
argue that “negative outcomes naturally trigger counterfactual thinking
because people are motivated to understand how to avoid such outcomes
for themselves in the future”. Gavanski and Wells (1989, p. 323)
conclude that: “counterfactual processing is basic to people’s perception
of causality. Causality is attributed to events that are perceived as
readily mutable”. On the one hand, the first event in a causal chain is
more mutable than the subsequent event(s) because it is perceived as
not constrained by previous events (at least events of the story). On
the other hand, normal events are less mutable than exceptional ones,
because the former are perceived as constrained by other causes (e.g.
social habits, rules of everyday life, rules of nature, and so on) and the
latter are, by definition, events which occur in spite rather than because
of these constraints. Thus, in both cases, the mutability of an event
appears to be inversely related to the number of causal conditions which
constrain its occurrence.

The results of this experimental tradition show that Suppes’s formal
definition of causality seems to correspond, at least on some occasions,
to our everyday use of the notion of causality. In fact, from these
experiments it comes out that – when we have more than one potential
cause – we utilize the perception of subjective frequency (normal vs
exceptional) and the ordering of the events rather than their content in
order to find the most relevant cause. Now we will try to show that
these conclusions were reached because the first factors studied were just the formal ones emphasized by the epistemological debate. On the contrary, Legrenzi et al. (1984) tried to show that it is impossible to explain our judgements on causality unless we take into account the organization of knowledge.

MENTAL MODELS AND CAUSALITY

Legrenzi et al. (1984) presented their subjects with a tragic story (a collision between a train and a car at a level crossing). The protagonist of this story (the level crossing keeper) was stopped by an exceptional event (a bridge falling down) while on his usual way to the level crossing. The subjects, as in the original experiment of Kahneman and Tversky (1982), had to complete a sentence that began: ‘If only . . .’. Despite the exceptionality of the main event of the story, a bridge falling down, only a minority of subjects (6%) changed this event. The majority (60%) preferred to change the usual time the keeper left his house towards exceptionality (‘If only he had left home earlier that night . . .’). But this result can also be interpreted according to the conclusion of Gavanski and Wells (1989), i.e. that the more mutable event is selected as the cause. In fact, from the keeper’s point of view (that is also the subjects’ point of view) it is easier to change the usual time of leaving the house rather than the fall of the bridge, as this event occurred independently of the keeper’s will. This interpretation can be extended to the results obtained by Girotto et al. (1991).

Girotto et al. (1991) presented the subjects with a story about Mr Bianchi. Bianchi, a bank employee, is going home after work. His progress is delayed by three misfortunes, not depending on his will, and by an intentional decision. When he arrives home, he finds his wife lying on the floor suffering a heart attack. He tries to help her but it is too late.

The results show that the events corresponding to the intentional decisions of the protagonist are more mutable and are indicated as the more relevant causes of the dramatic outcome. This factor, that we could call controllability from the protagonist’s point of view, prevails on other factors as ordering and exceptionality.

These results, replicated by Ferrante et al. (1992), show that Suppes’s definition of causality is not sufficient to cover our everyday judgements
on causality. The ‘controllability’ factor prevails on other factors even if it interacts with them. The general principle explaining all these results is that an event is selected as the cause when it is more mutable than the other events. This general principle applies not only to results obtained by manipulating exceptionality and ordering factors but also controllability. In fact for the protagonist of a story, it is easier to change his own decisions. Moreover, previous results show that subjects, in reading the story take the protagonist’s point of view. That means that subjects change the event more easily mutable for the protagonist of the story. In fact, works on the comprehension of narratives show that readers take the character’s perspectives; they follow the character’s thoughts, activating mental images of the same things that the character is thinking about (Bower and Morton, 1990, p. 47; see also Morrow et al., 1989). This easy ‘projection’ on the protagonist is also triggered by counterfactual tasks as we know from our qualitative results. When presented with a story we take the protagonist’s point of view and we check the type of decision taken by him. If no event of the story is the result of the protagonist’s intentional decision we select the cause of the outcome according to factors like ordering or exceptionality. On the contrary, if we find some voluntary decision on the part of the protagonist, we select this event as the cause as it is the more easily mutable event.

In order to understand how, in everyday life, people give judgements on causality, we have to understand how people make decisions. The descriptive theory of subjective causality seems linked to the descriptive theory of decisions.

CAUSES AND DECISIONS

According to the theory of rational choice we are consequentialist in the sense that decisions are determined by an assessment of the potential consequences and their perceived likelihood. On the contrary, in a cognitive perspective, decision-making is based on the reasons connected to a specific decision. For example in a recent experiment of Tversky and Shafir (1992), people chose to take a vacation in Hawaii whether they passed an exam or not, but decided to postpone the vacation in the disjunctive case, i.e. when the exam outcome was not known. If they passed the exam the vacation was seen as a time of celebration; if they
failed, the vacation was seen as a consolation. In the disjunctive case, lacking a clear reason, people postponed the decision.

Shafrir and Tversky (1992) have analyzed many situations in which the disjunctive effect occurs. These situations show that we decide to evaluate the reasons for the specific choice under examination, and not compare the expected utility of each possible option. This discovery is very relevant in relation to judgements of the type: 'The action $B$ caused the effect $C$'. In fact the evaluation about the causality of action $B$ on the outcome $C$ is often done by considering the circumstances in which $B$ has been decided.

In general terms we can say that we have the disjunction effect when people prefer $x$ over $y$ when they know that event $A$ occurs, and they also prefer $x$ over $y$ when they know that event $A$ will not occur, but they prefer $y$ over $x$ when it is unknown whether $A$ will occur or not. This effect (a paradox according to the classic theory of decision) becomes easy to understand if we assume that decision making is based on reasons built for a specific decision. In the same way we can explain the results of a recent pilot experiment (Legrenzi et al., 1993) in which subjects have to ask for information in order to decide whether or not to do $X$. For example the subject is told to request information in order to decide whether or not to go to the cinema. When this task is presented in absence of any context, it is very difficult to induce subjects to ask about not-$X$. On the contrary, subjects ask about not-$X$ when the context makes $X$ very implausible, for example when subjects are told that they are for the first time (and just for one night) in Rome. These results cannot be interpreted by means of the theory of rational choice. In fact we would expect people to compare the utility of $X$ and not-$X$, but this comparison is never really made. In the absence of context they ask about $X$ as they do not take not-$X$ into consideration. When there is a context that makes $X$ implausible (i.e. without any reason for doing it) they do not take $X$ into consideration and they ask about not-$X$.

The general conclusion is that we build a mental model focused on the specific decision under examination without taking possible alternatives into consideration. The decision making mechanism, based on this focussing effect, is related to the judgements about causality. In fact, in order to build a hierarchy of events as potential causes, we must judge how each event was decided. The relation between subjective causality and decision making is corroborated by the results obtained with the elaboration of the answers to another story, where all the events were
decided by the protagonist, Mr Bianchi (Ferrante et al., 1992). The results obtained through the usual counterfactual task, show that the hierarchy is built on the type of decision leading to each event. The more a decision for an action is taken under constraints (i.e. not freely), the less that action is seen as a cause of the dramatic outcome (for example to stop in order to recover from an asthma attack is a less free decision than to stop for a drink).

CONCLUSIONS

We have seen that the selection of causes among potential causal events can depend on our evaluation of the type of decision (intentional vs constrained). Both the probabilistic definition of causality and the theory of rational choice constitute two descriptive but also prescriptive models. If we analyze people’s reasoning in everyday life we discover that neither of these models explains all the experimental results.

We can also add that the experimental results allow a psychologist to raise the question about realism of causes. This question concerns the nature of causes: when we say that \( B \) causes \( C \) do we say something about the world or about a mental model of the world? Philosophers such as van Fraassen (1980) and economists such as Hicks (1979) have discussed the difficulties of a realistic interpretation of causality. According to these scholars, causal relations are established inside an interpretation of the world and they are not induced by the world itself. The work of psychologists, in different domains, support both these interpretations. On the one hand, the results obtained by Michotte (1954) on perceptual causality seem in favour of a realistic interpretation of causality (see also Miller and Johnson-Laird, 1976, pp. 96–100). On the other hand, the results about judgements of causality discussed here, show that these judgements not only depend on a mental model but also on a mental model built by a specific person with a specific knowledge organization. In fact, to evaluate whether a decision taken by that person was free (or not) depends on the knowledge organization attributed to that person. For example, if we apply the counterfactual paradigm to the analysis of errors in human-computer interaction, we discover that we treat errors made for different reasons in a different way, and that we infer the reasons from the type of knowledge that we attribute to the operator (Bagnara et al., 1992).
In order to emphasize the mentalistic nature of judgements about causality we can recall the results of an experiment by Wells and Gavansky (1989). They show that the protagonist’s mental consideration of an alternative can have an influence on the judgements about the causality of his/her actions. For example, a woman was described as having died from an allergic reaction to a meal ordered by her boss. When the boss was described as having considered another meal without the allergic ingredient, people were more likely to change the boss’s decision. Consequently, the causal role of the decision was judged to be greater than when the alternative meal was also said to contain the allergic ingredient. The two stories were identical except for the protagonist’s thought, which was not communicated to anyone but was crucial for causality attribution.

In accordance with all the experimental data presented in this paper, we can say that the probabilistic definition of causality cannot be considered psychologically as a real definition. Moreover, we tried to show a link between judgements of causality and decision making. It seems necessary to analyze the type of decision on which an action is based in order to judge the causal role of this action. Lastly, the data obtained by psychologists in different domains (perception vs. reasoning) support both the positions taken by philosophers on the nature of causality (realistic vs. constructivistic).

In conclusion, psychologists’ work in the last decade has shown the importance of Suppes’s definition also from a psychological point of view. Nevertheless, the present research has shown that everyday intuitions about causal links are very rarely based on computing different ratios of frequencies.

Paolo Legrenzi,  
Department of Psychology,  
University of Trieste,  
Trieste, Italy

Maria Sonino,  
Department of Psychology,  
University of Padua,  
Padua, Italy
REFERENCES


Legrenzi and Sonino's presentation of psychological studies on causality raises a number of key issues that I think it is important to comment on. Before I turn to these comments, however, I want to clear up one minor confusion in their characterization of my theory of probabilistic causality. It does not in any sense rest on a relative frequency interpretation of probability. I explicitly point out in my 1970 monograph that any of three interpretations of probability can be used: relative frequencies, probabilities given by some particular theory whether psychological or physical in nature, and subjective probabilities.

Intentional Actions as Causes. Legrenzi and Sonino cite a good deal of evidence that from the standpoint of psychological judgment of causal factors, the treatment of intentional actions as causes is very different
from that of nonintentional events as causes. These data could be related nicely to the very early history of physics in which many causes that we now treat as nonintentional were originally given an intentional cast. A good example is the early treatment of forces as due to the action of various kinds of spirits – former ancestors, local gods, etc. We can of course take a direct approach and say that the theory of causality of the sort I was concerned with is entirely focused on scientific causes. Misplaced judgments due to lack of adequate scientific consideration should not be part of the scientific theory of causality, although they may well be part of the scientific study of human judgments.

Before saying more about this important issue directly, I want to draw a parallel to the situation in the study of expected utility. Starting with the Bernoullis in the eighteenth century there is a long history of the normative theory of choice in uncertain conditions, leading up to the modern standard theory of expected utility given its classic formulation in Savage’s influential book of 1954.

In the last several decades a large number of psychological experiments have been performed to show that in making actual choices under uncertainty, various aspects of the standard expected utility theory are systematically violated. Kahneman and Tversky, themselves leaders in the psychological theory of causality judgments, have developed a prospect theory. A number of other alternatives are on the market. A good survey and the explicit development of rank-dependent utility theory is to be found in Luce (1991). These carefully developed models that are formally inconsistent with the standard subjective expected-utility model of choice certainly catch features of actual choices that are easily replicated in a variety of experimental settings. What is less thoroughly understood and documented is to what extent they catch features that are characteristic of real-life decisions on which we have enormous documentation, for example, those concerning financial transactions in markets of a great variety of kinds. In spite of the lack of really satisfactory documentation on this point, there is little question in most people’s minds but that the standard model is violated in ordinary financial transactions a great deal of the time. One classic response to this remark is ‘if that is the case, then any smart operator can make money from the violations of the expected utility model.’ This line of argument is similar to that concerning the much discussed question of the efficiency of various markets, for example, commodity markets, stock markets, etc.
There are a few cases where documentation is pretty good and show that indeed systematic deviations do occur. A classic case is the documentation of paramutual betting on horse racing. Broadly speaking, the evidence is overwhelming that bettors tend to overestimate the probability of winning on horses for whom the odds are long, and underestimating the probability of winning on horses for whom the odds are short. However, this distortion is usually not enough to overcome the track ‘take’ so as to make it really profitable to spend one’s days at the tracks betting on the short-odd horses. In a similar vein, in situations where large financial rewards or losses are at stake, it is likely that the deviations from the expected utility model are hard to take advantage of.

Given such standard deviations, but not of overwhelming magnitude, the question naturally arises, what are we to think about and to teach as the correct normative model of judgment? I continue to be conservative about this. We should accept as a normative model the expected utility model but recognize at the same time where it fails in predicting behavior and what the reasons are for some of those failures.

I take the same attitude towards intentional actions as causes. I must say that what is missing in the review of the literature by Legrenzi and Sonino are any studies about the overestimation or underestimation of causes that are intentional actions in terms of their subjective probabilities. This is a different question, of course, than their mutability, the point focused on in their analysis and in most of the experiments. In any case, we should be enlightened by these experiments and teach accordingly. It is, in fact, one of the great lessons of the design of random experiments in this century, that we need to protect ourselves against unconscious as well as conscious biases. Most of us familiar with medical data do not for a moment trust the subjective judgments of doctors in terms of skillful assessments of probabilities. If I want to know if a new vaccination is really effective, I want to see data from a real experiment, not just seek the subjective judgments of experts. And so it goes on and on. The purpose of such psychological research, in my judgment, is to alert us to the fallible character of our own causal judgments and to pinpoint especially those places where we need to be on guard and whenever possible, when the questions involved are critical, to conduct appropriate scientific experiments.

Given my own Bayesian predilections, the experiments surveyed by Legrenzi and Sonino provide a sobering influence on too subjective
a view. It is clear, as I pointed out long ago in my 1956 article on subjective probability, one can have a prior which from the standpoint of available information seems absurd. In the same way one can allow judgments about intentional actions, as opposed to causal events that are not intentional actions, to severely distort one's Bayesian prior. What the important work of psychological understanding brings is a methodology for improving the way we think about our Bayesian priors, more generally about the information available to us, and the selective part of it that we consider important. In my own view it is this direction that we should follow in building better normative models, namely, rich ones that have appropriate safeguards and checks as part of their structural features but recognizing always that these safeguards and checks cannot be absolute because still other information may be available to change the Bayesian prior. We can also be distrustful of our prior beliefs and recognize that without good experimentation it is very likely that in many subtle matters it will be impossible to come to a firm belief as to what is the actual state of affairs.

Causes and Mental Models. Legrenzi and Sonino state the point that in making or explaining our judgments on causality we must take into account the organization of knowledge, or put another way, we must use the mental models available to us. This I accept and would only add another formulation, namely, we use in advanced scientific cases the scientific theories available to us. Our mental models, if we are dealing with electricity, had better be guided by the modern theory of electricity and magnetism and not by the theory implicit in classical folk psychology. Without question, context is inescapable and the context itself must be fixed by selective attention drawn from scientific theories, informal mental models, or in unfortunate cases, blind prejudice. Again, it is the purpose of the modern theory of experimentation to learn how to reduce the elements of this context that are mistakenly identified as causally influential.

In this connection it would be useful to have a psychologically sophisticated overview of the continuing mental models that represent serious deviation from known scientific results in many different domains, as opposed to those mental models that are modern and recent, much influenced by results that were not known several hundred years ago. Certainly a very good example of this would be the mental model of diseases and their spread.
**Freedom of Decision.** I agree with much that Legrenzi and Sonino have to say about freedom of decision but I want to emphasize a point of my own which I developed earlier in my 1987 article on freedom of decision. The point is that the more general theory of constraints on decisions is that whenever possible we want to be faced with as wide a set of choices as we can. There are many different reasons for wanting the set of choices not to be closely constrained: (i) uncertainty over what will actually happen as a primary reason, (ii) unwillingness to make computations of an elaborate kind when future information that fixes the course of events much more closely will make such computations unnecessary, (iii) skepticism about the constancy of one’s own desires over a long period of time and therefore an unwillingness to commit to a course of action too far into the future. Classical subjective expected-utility theories have tried to take account of such matters by using the standard economic concept of discounting the future at some definite discount rate. I do not think that such apparatus is adequate to the problem of freedom of decision. It does not, for example, cover the very different cases of wanting choices in the marketplace to be as varied as possible given budgetary constraints, nor does it take proper account of the other factors just mentioned above. For this reason I do not take too seriously the paradox engendered by the experiment of Tversky and Shafir (1991) about the trip to Hawaii and having a different attitude when the result of the examination was not known. I take this example to be within the freedom-of-decision paradigm which I think can constitute a line of modification suitable for changing the standard normative theory. It is another example of increasing the structural complexity of the model, which seems required to account for our naturally deep concerns with freedom in decision making.

**REFERENCES**


