Loyola University Chicago Law Journal

Volume 44
Issue 5 *2013 Summer*

Article 2

2013

Behavioral Economics and Investor Protection: Keynote Address

Daniel Kahneman

SPEECH

Behavioral Economics and Investor Protection: Keynote Address

Daniel Kahneman

Transcribed and Edited by Michael J. Kaufman

It is a pleasure to be here. I think it is sort of a mixed pleasure to be embarrassed by introductions, but I find them embarrassing. Anyway thank you very much.

I won't surprise you by anything I say, certainly not those who have read my book. Actually not many people have read my book. Many people are reading my book, which is a completely different state of being. That is a state in which you can remain for a very long time.

I will start with the expression that Amos Tversky and I experienced. We have been branded as prophets of irrationality or as apostles of human stupidity. We never intended that. That is, everything Amos Tversky and I did was based on introspection. Every mistake in human judgment that we describe was a mistake that we found ourselves tempted to commit or in fact that we had committed. We were not particularly humble and we did not think ourselves stupid, but we thought we were studying the human mind in a realistic way.

What we were doing has to do with rationality. We show that people are not adequately characterized as fully rational by a definition of rationality which is completely unrealistic. This is the definition of rationality that is standard in decision theory—rationality is defined essentially by coherent beliefs and coherent preferences and in a complete system of beliefs and preferences that includes the past, the present, and the future. I want to point out that coherence of beliefs is not the same as logical coherence of reasoning. In logical coherence of reasoning, you follow an argument, there are premises and you get from those premises to some conclusion.

Coherence of beliefs—the test of coherence of beliefs—is much, much harder. When you are asked questions about your beliefs one at a time—not in conjunction, not in assessing the relationship between two

different beliefs or two different preferences, but beliefs elicited one at a time—it is completely impossible for a finite mind to achieve coherence. It is an essential characteristic of the mind that it is context-sensitive because beliefs are context-sensitive and preferences are context-sensitive. Therefore, when you illicit a belief in one context, it cannot really be consistent with beliefs that will be elicited singly in other contexts. So, even if there is logical coherence and reasoning (and that is something that I cannot guarantee that there is), there is no logical coherence in beliefs. Therefore, the definition of rationality in decision theory and in extended economic theory is so outlandish that it is not a major achievement to find objections to it.

We did not prove that humans are irrational. I would draw a very important distinction between demonstrating that people are not rational by a particular definition of rationality, especially an absurd one, and claiming that people are irrational. I do not plead guilty; I plead innocent of having ever made the claim that people are irrational. It was made on our behalf more than once. We always regretted it and rejected it. Irrationality conveys something much, much stronger than the view of human nature that Amos Tversky and I ever held.

System 1 and System 2 have been mentioned. They are not my terms. But because the book that I wrote a year ago became part of the language (well not quite part of the language, but they are fairly common to it), let me just introduce System 1 and System 2 very, very quickly by examples of how they work.

If I say 2 plus 2, a number comes to your mind and the number came to mind without effort, without deliberation, without intent. It happened automatically; it is something that happened to you. That number came up as an automatic; it was produced by an automatic operation of associative memory. 2 plus 2 is associatively related to the number 4, and the number 4 pops up when 2 plus 2 are produced. System 1 is the one that is characterized by the thoughts, ideas, feelings, and response tendencies that are produced automatically without any specific intent by a particular context. The context can include goals and intentions, but the goals and intentions do not directly produce the result; they are just part of the context.

System 2 is involved in very different kinds of activities. I always use the example of asking you to multiply 17 times 24. Nothing will come to mind immediately for most people, but several things will actually come to mind: you know it is a multiplication problem, you know whether you could or could not do it in your head given time, you know it is less than 20,000, and you know it is more than 100. There is quite a bit that System 1 actually delivers about that problem. But you

do not know that it is 408. In order to get to that knowledge, you have got to retrieve a program that you learned in elementary school and apply it, and that is a sequential operation. It is not automatic; it is planned and it is intentional.

Another very important function of System 2 is control and inhibition. Many of us have had the experience of not telling someone to "go to hell." When we are refraining from that, that is a System 2 operation. System 2 operations are characterized by the fact that they are effortful and costly, which means that we cannot carry many of them at once. Therefore, System 2 gets tired and gets depleted and when we are tired and depleted, we are more likely to tell people to "go to hell" if we feel like it and to indulge in a variety of other minor vices. An important characteristic of System 2 as I describe it is that it is lazy. What I mean by lazy is that it operates by the principle of the least effort. We do not like by and large to exert mental effort; we avoid it.

And the major theme of my book (you don't have to read much of it to discover that theme) is that I describe the interaction between those two Systems. System 1 proposes associative responses, feelings, and tendencies, most of which are endorsed by System 2, which is to some degree in control, but is tremendously influenced by what happens outside of any control within associative memory.

The key concept to describing the way that System 1 works is another form of coherence. This is not the logical coherence that is invoked in definitions of rationality, but what I call associative coherence. Associative coherence means that our representations of various aspects of the world tend to be consistent with each other, emotionally and intellectually, and we tend to suppress interpretations of the world and ideas that are inconsistent with that model. I speak of the halo effect, a familiar psychological phenomenon, in the book. We have coherent attitudes, and they are highly context-dependent. A word does not have the same interpretation in different contexts. For example, if I say "she approached the bank," most of us will understand "bank" is a financial institution. But, in the context of fishing, "she approached the bank" means something entirely different. We are not aware of our choice about which of the two meanings to accept. That choice is made unconsciously. Associative memory delivers an interpretation.

What we call priming effects are the effects of the contexts in facilitating some interpretations rather than others. I have another example which occurred recently that involves my wife. I am very sorry it did not get in the book because it occurred so recently, but we were out having dinner with a couple, and afterwards we were summarizing our impressions. My wife said of the man "he is sexy,"

and then she said something totally bizarre. She said, "He doesn't undress the maid himself." He doesn't undress the maid himself? It just didn't make sense. What on earth do you mean, "He doesn't undress the maid himself?" Well, it turns out that was not what she actually said. She had said "he doesn't underestimate himself." But in the context of the single word "sexy," that was enough to twist my mind in a direction that produced this very strange interpretation. I would like to point out why that is important. When I heard my wife say this, I thought it was very strange that she had said it. It did not occur to me that I was wrong. That is very symptomatic of the way our mind works. I took it for granted that I heard something correct and it was very puzzling. I did not assume that perhaps I had misheard because in fact she could not say such an odd thing; the idea did not come to me until she actually pointed out that she had said something else.

This associative coherence has many manifestations. One of the more important ones is that our interpretation of the way our mind works—our sense of the way our minds work—tends to be profoundly misguided. We feel that we hold our beliefs and have our preferences because we have reasons. We feel that reasons come first, and they generate or produce our conclusions and our preferences.

We also really believe that we believe in things because they are true. We would not believe them if they were not true. Believing and knowing are very closely associated in our minds, and the idea that the reasons are what drive the beliefs is really part of the experience.

It turns out that that is not why we believe and know things. It is enough to look around you. It is very, very clear that in most significant domains of life, we do not believe what we believe because of arguments. We believe what we believe because people that we like and people that we trust believe the same things. This is really the reason that some people think that the *New York Times* tells the truth and other people think that if the *New York Times* says something, it must be a lie. That is not for argument. People can bring arguments, but what really counts is political identification. Who are the people that we trust?

It is the social determinates of beliefs that come first. Otherwise, the very high correlation in this culture between beliefs about gay marriage and beliefs about global warming should not go together so closely. They go together so closely because of associative coherence, and because the determination of what we believe is primarily social and not by argument. We believe evidence and arguments because it is consistent with our beliefs; it is not the other way around, and that is what associative coherence is.

We take the whole network of ideas that are related and we believe in that network. There is a nice experimental demonstration that I actually cite in my book. It was an experiment done in England, in which people are asked to verify the logical consistency of syllogisms. They are asked: "Does this conclusion actually follow from the premises?" If I remember the example, it says, "All roses are flowers. Some flowers wilt quickly. Therefore some roses wilt quickly." People are asked about that argument; is it a valid argument or not? As this crowd figured out, it is not a valid argument. But a large majority of undergraduates think it is because the consequence is true. And there is an associative relation between the premises and the consequences.

And the consequence being true drives the belief in the argument being valid. What we see here is extremely common in the way we think. It is the essence of the idea of heuristics and biases that Amos Tversky and I developed a long time ago. It is what I call substitution. We answer a different question than the question that was asked. The answer has validity; the answer that comes to mind is something related to truth. It comes to mind because it is associatively related to what we are trying to answer and we do not discriminate. Unless we slow ourselves down by System 2, we are going to blurt out an answer that is incorrect.

My former colleague Shane Frederick came up with an example that many people know about. It is called the bat and a ball example. Where a bat and a ball together cost $1.10, and the bat costs $1.00 more than the ball, how much did the ball cost? What makes this interesting (and is very much like the roses example) is that everyone in this audience had an immediate reaction to that puzzle unless you had read it before, and the immediate response was 10 cents. It is designed to elicit 10 cents. Now 10 cents is false; the answer is 5 cents.

It is very, very easy to find out that the answer is false. Because if it is 10 cents for the ball, it is $1.10 for the bat and $1.20 for the total; so that answer cannot be right. More than 50% of Princeton students fail that test. And the same is true of Harvard and MIT students by the way. What do we learn from this? I think we can learn a very important thing. We learned that these people did not check themselves because if they had checked, they would not have said 10 cents. What we discovered here is our tendency to believe the proposals of System 1 and to endorse them without checking. Sometimes we will not be able to check them, sometimes we could, but System 2 is really lazy and checking is not something that we do as often as we might.

It turns out that System 1 is the origin of most of the mistakes in reasoning and preferences that Amos Tversky and I and many others

have studied over the years. But I want to qualify that I do not want you to emerge with the idea that System 1 is inferior or stupid.

System 1 is often right in the first place. It is mostly right, and our highest skills belong to System 1. Whenever we get skilled at any activity, be it driving or getting a sense of a social situation, it is due to System 1. Let me give you an example. (It is terrible by the way when you write a book. Within approximately a year or a little less than that, you absolutely forget everything that you used to know. This has happened to me. The only thing that comes to mind is written in the book and I apologize for that). But the example that I do give is that I know my wife's mood from one word on the telephone. That's a skill. It's a System 1 response. And it comes from experience. Experience has built an automatic response.

Most of our highest skills have that characteristic. They are automatic. They belong to System 1. We should not have the idea that System 1 is specialized in producing mistakes; it is specialized in producing skilled behavior and it occasionally makes mistakes.

An important characteristic of System 1 is that it does not stop. When it does not know the answer to a question, it generates an answer to a related question which gets substituted for the correct answer. That happens a lot and that is the basis for many mistakes.

I want to emphasize the extent to which very detailed world knowledge is built into our associative memory. Here again is a familiar example, really my favorite I think. There is research in which people's brains are imaged while they hear sentences. They hear an upper-class British male voice that says, "I have large tattoos all down my back." Within approximately one third of a second, the brain shows characteristics of a response of surprise. An incongruity has been detected.

If you stop to think about what is required for this, it is astonishing. You need to identify the voice as an upper-class male British voice, you need to bring the stereotype of upper-class British males, and you need to detect that upper-class British males, so far as we know, are unlikely to have large tattoos all down their back. That is an incongruity. A surprise response is evoked, and System 2 is mobilized to ask: "What is that strange thing?" This a beautifully working system. We are completely unconscious of what it delivers, of the process by which it delivers. We are conscious of the product, but not at all of the process.

I will give you one other good example of answering the wrong question. About twenty years ago or so when there was terrorism in Europe, Americans were leery of traveling to Europe, which actually makes no sense at all because the risk was very small. At that time, life

insurance was sold and trip insurance was sold in the airport. An experiment was run to illicit what was the maximum willingness to pay for people for the insurance policies. One of the policies said that your decedents would be paid $100,000 for death during the trip in a terrorist incident and the other policy would pay $100,000 for death during the trip for any reason. Now, it turns out that the first policy is worth a lot more to people than the second, which is absurd. The reason is that if you ask yourself, "What are people going to do when they buy insurance," they do not have an easy answer to that question. There is something that they *do* know, which is how afraid the question had made them. We are more afraid of dying in a terrorist incident than we are afraid of dying. And that is it. We are not aware of it. There is actually no way when people encounter one of these questions to know that their answer is logically inconsistent with the answer that they would have given to the other question. That is what I meant earlier when I talked about the impossibility of logical coherence. This is the way that our mind works.

Our mind does not work like that all the time. We are perfectly capable of computing. We are perfectly capable of checking ourselves, and sometimes it works and sometimes it does not. For example, the System 2 of most of the population is very poorly aware of statistics. System 1 does not think statistically at all. And a person's System 2 needs to know statistics in order to get to the correct conclusions, and we frequently do not.

Finally, where does this all get us in terms of policy implications? I suggest both in the book and in these remarks the heuristic for predicting the behavior of people. And it is a very general heuristic. It does not get you there perfectly. But the heuristic is to take a careful look at the first response, take a careful look at what comes first to mind. That first response is not necessarily what people will do, or will end up doing, or feeling, or deciding. But it gives you a clue about the direction in which they are being steered. They might resist it, they might change it, they might modify it; but knowing that first direction is extremely important. I would add to this that the prediction of behavior from the first response is stronger when System 2 is lazier. So where people are not used to applying System 2 thinking, or operate under conditions of depletion and fatigue, the heuristic of looking for the first response is a good heuristic.

And what follows from this emphasis on the first response is often called behavioral economics. And I have a complaint. My best friends Cass Sunstein and Dick Thaler published a book, *Nudge*, which is in effect a book about social psychology. But Dick Thaler is a behavioral

economist. From there on, a large chunk of social psychology has been called behavioral economics. It is a mistake actually; it should be called behavioral science because it is not economics at all.

What does follow? What follows is pretty obvious. If people are limited, then one policy implication is that they need protection. They need more protection than they would in the standard rational model. And they need protection in two ways. First, they need protection from their own mistakes, to some extent. Of course that creates the dilemma of how you leave them free, and that is what nudges are about. Second, they need protection from legal, but predatory actions in the market. Legal, but predatory actions basically exploit the laziness of System 2. For example, we are all familiar with what we are willing to sign without reading it when it scrolls down the screen. We would sign our lives away without reading, so people need protection from a lot that is legal. That is one implication.

And then there is another heuristic that follows from social psychology—that follows, really, from the description of the mind as I have offered it. And it is a heuristic for influencing behavior.

If you want to influence behavior, you want the first impulse to go in that direction or you certainly do not want the first impulse to go in a direction that is diametrically opposed to the direction you want to influence. This is what nudges are. Nudges are ways to make it easy for people to think or want or do certain things. And making it easy means that it is compatible with their associative system. It is compatible with their values and it is compatible with their desires.

And that is what I was going to tell you to today. There is more to the story of course. But I think there is enough to get a discussion going.

Thank you.