RECORD, Volume 31, No. 2*

New Orleans Health/Pension Spring Meeting
June 15–17, 2005

Session 14PD
Behavioral Economics: The Art and Science of Making Choices

Track: Health

Moderator: JOHN W.C. STARK

Panelists: ELLEN PETERS†
JOHN W.C. STARK

Summary: The concept of improved information for improved decision-making is a cornerstone of the consumerism movement that is such a hot topic in today's health care marketplace. But how does that information really influence the choices that people make? What other factors impact decision-making and behavior? Panelists at this session discuss some of the research behind understanding decision-making and how this type of research can be used to enhance the modeling of consumer choice. Attendees gain an awareness of the field of behavioral economics and its potential application in improving assumptions for modeling health care purchasing.

MR. JOHN W.C. STARK: I work for Anthem Health Plans of Virginia which is part of WellPoint. We are going to talk about decisions, how to make decisions in the face of risk and how people behave when they make decisions today. My co-panelist is Ellen Peters, who's a psychologist. I probably couldn't do her justice, so she'll tell you more about herself when she comes to the podium.

During my career, I have worked in a variety of areas, including HMO, individual, large group, pricing, forecasting and budgeting. I have seen a lot of things that we are going to talk about today.

†Dr. Ellen Peters, not a member of the sponsoring organizations, is research scientist at Decision Research in Eugene, Ore.

You might ask how I became interested in this topic. In the course of working with the Risk Management Section, I was reading an article in *The Economist*, and there was a survey on risk. One of the articles in the survey had to do with how individuals dealt with risk. It was fascinating, and it described a lot of behaviors that I had seen in the course of my work. As I began to read a more about it, I became more interested in this area. Having said that, let's do a little Q&A first.

Why should we care? As actuaries, numbers are our life. One of the reasons we should care is that the perception of actuaries has been that we're very narrowly defined as numbers people. From the standpoint of professional relevance, incorporating some of the tenets of behavioral economics into our work would enhance our value. I am not aware that any other group is doing this.

Another is that our training has covered utility theory and the rational consumer, which are nice constructs that make the formulas mathematically tractable. However, experimental results contradict this theory. Some of the experiments seem fairly simplistic, but the researchers take a lot of pains to make sure that the controls are in place. Enough experiments have been done so that you can look at them and say that they're very simple, but they reinforce each other.

Finally, we deal with more behavioral issues than we probably realize. For example, since risk is part and parcel of our business, it is important to understand the behaviors that drive decisions in the face of risk. Here is another example. How many times have you tried to present a result and you've just gotten blank stares? Here the risk is that people in your company may make a decision with less than complete data. Finally, when we talk about health care utilization—since I'm a health care actuary, a lot of the examples are going to come from health—a lot of what we try to do in terms of reducing utilization and promoting wellness deals with changing behavior. These are some things that I think are important in terms of behavioral economics for actuaries.

I am going to talk about decisions and assumptions. As I said before, one of our main jobs is to help people make decisions. How do we do this? We do an in-depth, detailed and complex analysis. If you leave out two or more of those adjectives, you have not done an actuarial analysis. Then we present it, either as a written report or in person. The question is: How do people use our results? Do they understand them the way that we understand them?

Ellen is going to talk about how people look at numerical results. People process numerical information, especially probabilities, in very different ways than you might think. As a result, people can misuse our results, or they can just ignore them. I do not have to say much about assumptions because we make assumptions in the course of our work all the time. Some are material; some are not.

Having said that, let's look at some behaviors. Antiselection relates to the behavior of people who are trying to do the best for themselves and the worst for our companies. I know that in our area we spend a lot of time trying to figure out how to price for antiselection and how to deal with it. In the small group market in life insurance, underwriting can help you deal with this, but when you get more into the group market, you have to figure out how to price for it or use some other mechanism.

As far as changes in member utilization, our tools are benefit design and medical management. We try to use those, to change behavior, to get people to use more cost-effective settings, and to use the appropriate services. For example, some years ago we had very low co-pays for outpatient services versus inpatient to get people to go to more cost-effective settings.

We try to change provider behavior. Again, we have reimbursement mechanisms and medical management. Think about trying to change the behavior of doctors. This is a group that is more independent than we are. If you deal with provider negotiations, you know this is very difficult.
When you think about changing behavior, what tools do we have as actuaries? We only have a couple: benefit designs in policies and money. We can give more money or take away money; we can make our company easier or harder with which to do business. Are those the best tools to change behavior? I do not know. Sometimes it seems like they are the best tools; sometimes it seems like they're woefully inadequate.

Lapses are related to antiselection and are important when you start talking about cash value and extended term in life insurance. There are more. Commissions and incentives are very closely related. Incentive programs work great. You just need to make sure that you are using the correct incentives to result in the desired behavior. The incentives I am talking about are not only management incentives; they are provider incentives and commissions for your sales force. How many times have you put a product in the market and you sell more than anticipated or find that providers are doing certain types of services that are not what you intended? In these situations, the incentives are working; just not the way they were intended to work.

Any of you that deal with disease management and have had to do some pricing on it know that this is all about changing behavior, including getting people to take their medicine, go to the doctor, and so on.

Then there are multi-choice products. How do people decide what products to take? Are they as astute as we think they are?

What is our basis for all of these behavioral assumptions? Experience data is one of the cornerstones used in selecting assumptions. However, experience data merely reflects the outcome, not the thought processes of the members making the

decisions. Another technique that we use is logic. For example, when we have meetings to select or validate assumptions, someone will say, "If I were doing this, this is how I would pick a benefit." Frankly, I think we need a better foundation. I think behavioral finance provides at least some insights, if not some tools.

Let's take an example. Let's talk about gender differences in decision-making. Who makes the decision to purchase insurance? When you go to the doctor, how do men and women decide to go to the doctor? As a man, if I can't lift my head off the pillow, it's time to start thinking about going to the doctor. Maybe I'll go the next day. Who decides when children should go? Again, this varies all over the place. Who's developing products and policies? Are the answers to these questions consistent? Probably not. Gender-based differences in decision-making have been documented, but the reasons are still hazy. One school of thought is that the differences are not necessarily biological; they relate more to political and social values. These are the kinds of things to think about in terms of behavior.

What are some other uses? You heard about risk management today, and this is another important factor. Operational risk is mainly people risk, as is reputational risk which has been prominent in the past year or so. Many of you are involved with Sarbanes-Oxley. That was all about behavior and reputational risk. Situations involving reputational risk are rare (hopefully). As a result, there are very few data points. Also, each situation is unique. So it would be difficult if not impossible to develop a trend or to develop a model for your company. Finally, when a company's reputation is in jeopardy, things get emotional real fast and you're in crisis management before you know it.

Strategic risk is similar. I think we have a very short-term view these days, especially public companies, because we are always worried about what is going to happen in the next quarter. It is easy to take your eye off of what's going to happen in the next year or so. Again, when a strategy is not working (e.g., you feel like you are going to miss your numbers for Wall Street), things can get emotional very quickly.

For the rest of the presentation, I'm going to talk about some things that I think are interesting and illuminating. Now, as an actuary, some of the things I've experienced in my company make me want to scratch my head. Sometimes I think I am on the wrong planet because I have a certain way of thinking about things, and it seems right to me but not to other people. These topics will help show the parts of human nature with which we're dealing.

The first one is anchoring and adjustment. This has to do with how people process and integrate information. The first thing you have is an anchor which is somebody's first approximation. So far, so good. Next, you apply adjustments when you get more information. That sounds reasonable, except those adjustments are very crude and very simplistic. See if this rings a bell. You're in a meeting. The CEO or CFO says, "We're thinking about putting this new benefit in."

The CEO or CFO turns to the actuary and says, "Can you give us an estimate? We won't hold you to it. We just need an idea." Judging by your reaction, clearly all of you have been in that spot. A day or so later, you are thinking about this. Your estimate changes, in the wrong direction, of course, because you're trying to help people out, and all hell breaks loose. What has happened is that the anchor was very strong, and the new information that you considered important, other people didn't consider that important. This is something to be aware of as you are dealing with folks.

Here's one that's going to be near and dear to your heart. What about those pesky little probabilities? They are a pain in the neck, aren't they? I'm going to talk about two phenomena which are very similar. The first one is a little more scientific and is called the law of small numbers. What happens in this case is that there was an event that occurred several times. The person making the statement generalized, and, voila, we have a conclusion on our hands that we need to deal with. In some of the literature this kind of behavior has been documented, and I have seen this plenty of times.

The next one is called availability. This one is pretty interesting. What is availability? Either you can imagine it, or it's something that has happened and you can remember it easily. It goes back to experiential things. See if this sounds familiar. You're sitting in a product development meeting talking about the latest cost of health care, and somebody says, "I took my son to the emergency room last week, and it cost a bundle." All of a sudden we're off and running about the expense of emergency room care. It made a big impression and it was easy to remember. On the other hand, how many of you have been in meetings talking about risk and someone says, "I just can't imagine…"? That person's judgment is going to be based on whether or not he or she thinks it's realistic. That's very dangerous because it causes you to ignore risk.

Have you ever had one of those little back-of-the-envelope calculations people want you to verify? Again, this is something that helps people integrate information. They are using simple techniques to avoid very complex calculations. Say you're in a meeting. One of your colleagues turns to you and says, loud enough for everyone to hear, "I just used a little quick math to come up with how much our new medical management initiative will save. What do you think of this calculation?" Several things go through your mind. The first is that you've heard this talk, and you know about anchoring. So if you don't refute this, it will be the standard in the company. Next, you look at the calculations. They look kind of logical, but you know there's something wrong. Finally, you know that it's going to take time to review plus the explanation is going to be complex enough so that you will lose everybody. These are dangerous situations, and as these happen, if you're dealing with people you've known a long time, you can probably say that you need to think about it, but sometimes you're really on the spot.

Standard theory in utility theory says that the utility of a gamble is linear in all the probabilities, but there have been experiments that show that the difference between getting information that will take you from 99 percent to 100 percent (a 1 percent difference) is much more valuable than taking you from 10 percent to 11 percent. Going from still being uncertain to being almost 100 percent means a lot more. In this case, the willingness to take a chance on an event is a function not only of the level of uncertainty, but also the source.

Let's talk about risk-seeking behavior. Risk aversion is assumed to be the norm, but experiments have shown that people exhibit risk-seeking behaviors in several areas. First, there's a preference of a small probability of winning a large amount versus a sure thing by taking the expected value. See if this one rings a bell. There's also a preference for a probability of a large loss versus a small certain loss. How many times have you been in meetings when you say, "We're not going to make our forecast unless we do such-and-such"? You know the probability is small. What's the first thing that happens? "Let's do it. We'll make our numbers." The next thing is loss aversion. Losses loom larger than gains. They are not symmetric.

The next topic is framing. In theory, no matter how a problem is presented, the standard theory says that you should be able to cut through it all, and you will come up with the same answer or the same analysis, no matter what. Experiments have shown this not to be true. Remember the old saying, "It's not what you say; it's how you say it." There's a lot of truth to that. Now, as an actuary, I've had to modify the following statement: "The truth will set you free" to "The truth will set you free if you say it right." An example is using numbers versus percentages. The statement "Ten people out of 100 will get cancer" will have more meaning to some people than "You have a 10 percent chance of getting cancer." This is all about knowing your audience. I'm sure that in presenting results this is something we've all run up against.

My last topic is bounded rationality. Can you bound rationality? Are you rational up to a point? Actually, this is an alternative to maximization in utility theory. Going back to utility theory, one assumption is that we maximize at every possible opportunity. You try to get the *best* deal on a car, make the *best* decision about bills and so on. Well, some experiments have been done, and the results seem pretty reasonable. People really work to get a satisfactory conclusion. When you think about it, optimization takes a lot of time and effort. You can't do it all the time, and sometimes it's not worth doing. Sometimes we're talking about very minimal things. Also, you have problems with complex calculations. As human beings, there's a certain point at which we don't want to make them.

What does this bounded rationality mean to us as actuaries? First, consumers make a lot of choices during the course of a day. Next, health insurance is extremely complicated. Based on these two facts, can our members optimize their choice of health coverage? I know that we sit in meetings and talk about how people are going to use benefits. I feel that as actuaries, product managers and

underwriters, we attribute some of our knowledge to our customers, and it's probably not true. Have you ever priced a health product? Sure. Have you ever had to use your benefits? Do you suddenly feel like you're way on the other side of the table? I think that's another way to try to look at things. One of the questions I have is: Are we making our products too complex? Do we need to simplify? Some of the complexity in our products is meant to protect and guard against risk, but is that really necessary?

This is the end of my part of our talk. My goal was show that behavioral economics has some interesting ideas that we as actuaries may be able to use.

DR. ELLEN PETERS: As John said, I'm a psychologist. I was originally trained in engineering and marketing. I switched over to psychology about 10 years ago because I was interested in how people process information and how that makes a difference in how they judge and how they decide. In particular today, I'm going to be talking about the processing of numbers. How do people understand numbers? How do they use those numbers or not use those numbers? How does that make a difference in how they judge and how they decide?

I'm going to take you on a brief tour of findings in my research literature that John and I thought might relate to what you do as actuaries. My aim is to help you understand how the general public understands and responds to numbers and to risks.

Your training covers expected utility theory. So does mine. I'm going to talk today, though, about bounded rationality and about experimental results that illustrate the shortcuts that decision-makers use when they process information. Your work deals a lot with behavioral issues, and it deals a lot with assumptions. You make a lot of assumptions in the work that you do. One of the assumptions that you make, from my standpoint, is that consumers actually understand the information that you give them. Not only do you assume that they understand that information, but you assume that they're then able to use it in order to make better choices. We're going to look at that today.

As it turns out, simply providing information to people isn't always enough. You have to pay attention to how that information is presented. Let's take a quick look at a real-life example about the simple provision of information not being enough for good decisions. For 30 years, the State of Nebraska allowed workers to choose either a traditional pension plan or a 401(k) plan that was managed by the individual worker. Workers who chose the traditional pension plan earned reasonable average returns. However, those who chose to manage their own 401(k) plan earned average returns that were substantially less. In 2003, Nebraska eliminated employee choice from its 401(k) plan.

To make good decisions, you do have to have information. It has to be available, timely and accurate. In addition to that, people have to comprehend it. They have

to comprehend not just what the number is but what it means. What's the meaning that underlies those numbers? They have to be able to determine meaningful differences between options and to weight factors in order to match their own needs and their own values, because your values and needs might differ from mine. They also have to be able to make trade-offs. They have to decide, "Okay, I'm going to go for a little less on this because I want a little more on this." Ultimately, they have to choose.

There are a number of potential barriers to using information effectively in choice, however. For example, information in retirement choices can be insufficient or uncertain because the worker doesn't know things like the path of future earnings, his or her age at death and the interest rate he or she can earn across time. A decision-maker might simply not comprehend the information that's in front of him or her.

Let me give you an example of that. We've conducted a series of studies on health insurance choices. In one particular study, we examined one of the basic building blocks of choice, simple comprehension, with two groups of subjects. One group was a group of employed-age adults (less than age 65). The other was a group of Medicare beneficiaries, adults who were age 65 or older. I think our oldest subject was probably about 100 years old. With these two groups of subjects, we looked at a comprehension index that reflected the average number of errors made across age on 34 fairly simple decision tasks that involved interpretations of tables and graphs.

Let me show you an example of the simplicity of some of these tasks. Subjects got the information that you see in Chart 1. They got information on four HMOs, with two pieces of information on each of those HMOs. They are simply asked which HMO requires the lowest co-payment for a visit with a primary care doctor.

Chart 1

Example of Decision Task

	HMO A	HMO B	HMO C	HMO D
Monthly Premium	$50	$75	$48	$63
Copayment for office visit with primary care doctor	$10	$5	$15	$10

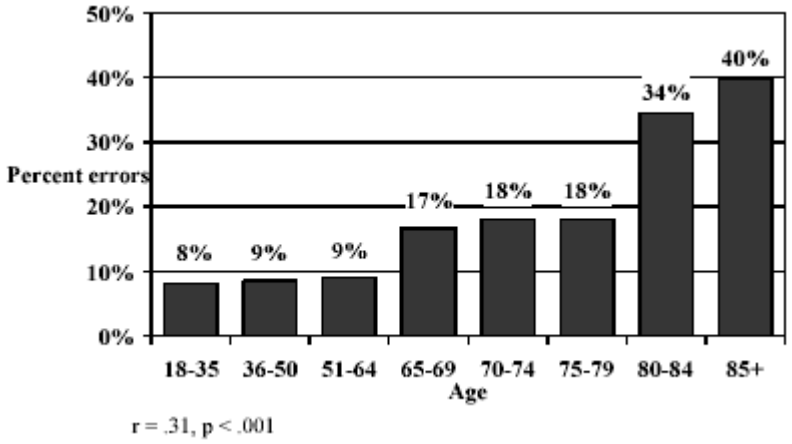
1. **Which HMO requires the lowest copayment for a visit with a primary care doctor?**

You know that answer. I know that answer. It turns out that not everybody knows the answer to this question. Not everybody is able to answer this simple question correctly. That comprehension is strongly related to age as well.

What I'm showing you on Chart 2 is the percentage of errors that are made on average across those 34 decision tasks. Again, it involves fairly simple interpretations from tables and graphs of numbers. What you can see is that there are problems of comprehension in every age group. The youngest subjects, age 18-35, made 8 percent errors on average, but that proportion of errors goes up substantially with age. Those older adults, people who are 85 years old and older, actually made an average of 40 percent errors on these fairly simple graphs. They didn't understand the numbers that were presented to them.

Chart 2

Problems with Comprehension and Large Age Differences



Sometimes there's a simple lack of motivation. There's a lack of motivation to try to understand information, to understand maybe your own values or to make a choice. Maybe one way to solve the motivation problem in choices, though, is just to provide more options to people. If you provide more options to people, then maybe at least one of them will be attractive. So that's good, right? Well, Iyenger & Lepper conducted a study in upscale grocery stores in Menlo Park, Calif. They invited passing customers to taste-test some jam. In one store they presented six jams at their booth. In another store they had 24 jams that customers could taste-test as they passed by the booth. They thought that maybe more would be better. What they found was that more people, indeed, stopped by their booth to try a sample when there was more variety. However, while the variety is initially more attractive and more people actually did taste-test when at the 24-jam store than the six-jam store, it was demotivating to actual purchase behavior. In the six-jam store, 30 percent of the shoppers actually purchased a jam compared to only 3 percent of those in the 24-jam store. Having more choices seemed to be demotivating.

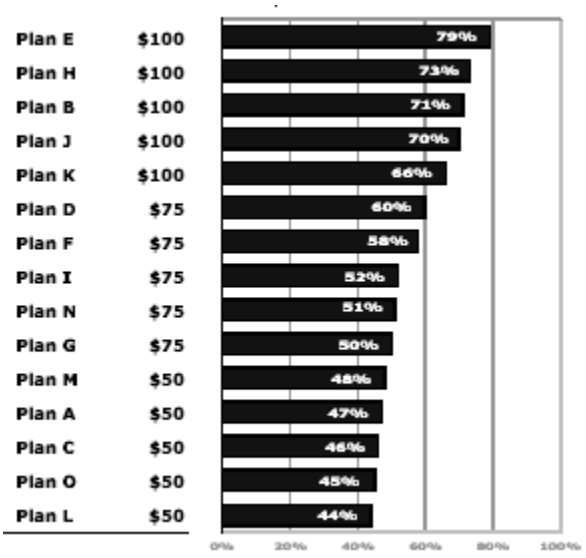
In another study on health insurance choices with Paul Slovik and Judy Hibbard at the University of Oregon, we again looked at this idea that more is not necessarily better. Subjects were shown 15 health-plan options with just two pieces of information: their monthly premium cost and consumer satisfaction ratings for each of these different health plans. The task was simply to choose a health plan. In this case, having more choices seemed to lead people to choose worse, because it took effort to try to figure out the best plan for the cost that the person wanted to spend. This was actually an experimental study. Half of our subjects were shown information ordered alphabetically, just from A to O, as shown in Chart 3. But for half the subjects, we ordered the information for them, as shown in Chart 4. We ordered it by quality within cost strata. Making this simple change helped people to use the quality data in choice. More subjects were able to make a quality maximizing choice just through this simple torque in how we presented the information.

Chart 3

Premium Cost & Consumer Satisfaction Ratings
(ordered alphabetically)

Chart 4

Premium Cost & Consumer Satisfaction Ratings
(ordered by quality within cost strata)



A fourth potential barrier to using information effectively is that sometimes decision-makers will actually comprehend information—they'll know what the number is—but they won't understand what it means to them. Let me give you an example of that. In another series of studies, we've been examining the concept of evaluability. Evaluability is the ease or precision with which an attribute can be mapped onto a good/bad scale. How easy is it to tell how good or bad this thing is so that I can then use that information in a choice? What we found is that evaluability matters as decision-makers try to figure out how much they value dollar amounts and how much they value other types of numerical quantities.

In this particular case, we looked at how people evaluate the attractiveness of a simple bet. Subjects are asked to rate a bet on an attractiveness scale from zero to 20, where zero is "not all attractive" and 20 is "extremely attractive." There are two different bets, and each subject sees only one of the two bets. So again, this is an experimental study. Group 1 rates a bet that gives you seven chances out of 36 to win $9. The subjects rate that bet on that 0-20 scale. A different group of subjects rates a slightly different bet. Group 2 rates a bet that gives you seven chances out of 36 to win $9, like Group 1, but the remaining 29 chances out of 36, you'd lose five cents. Group 2, in other words, has an objectively worse bet. They should rate that as less attractive, right? That would be the rational thing to do. Yet what we find on that 0-20 scale is that Group 1 rates the bet about a 9, while Group 2 with the objectively worse bet finds it more attractive, about a 14. Why is that? How can

that be? In drawing meaning from the numbers, the presence of that little small loss makes a difference. It's not that the decision-makers didn't comprehend the $9. They know what $9 is. It's that the meaning of the $9 was different in the presence of the five-cent loss.

We explained the bets example using the concept of evaluability. The seven chances out of 36 are easily evaluable. We know how good or bad it is. We know it's not a very good chance. It's a pretty poor chance, in fact. But how good or bad is the $9? It depends on the context. It makes the valuation of that $9 more difficult for people. Subjects faced with a $9 no-loss bet rate the $9 as neither good nor bad. They're neutral about it. They have no particular feelings about the $9. However, those subjects who are faced with the $9 loss bet, in the context of that small nickel that they might potentially lose, rate the $9 as quite good. In comparison to the small loss, we think about the $9 as coming alive with feeling. It suddenly becomes very positive. It becomes evaluable. The decisions that your customers face are much more complicated than a simple bet. Yet even with a little simple bet like this, there are barriers to how decision-makers understand the meaning of the information that's given to them.

Let's look at another example of evaluability. This is actually a 1998 study that was done by Christopher Hsee at the University of Chicago. This, again, is an experimental study. There are three groups of subjects here. The first group of subjects sees two music dictionaries side by side, Dictionary A and Dictionary B; they're in the joint evaluation condition. One group of subjects sees both Dictionary A and Dictionary B, and they price how much they'd be willing to pay for Dictionary A and then for Dictionary B. A second group of subjects sees just Dictionary A and never sees Dictionary B, and they tell us how much they value Dictionary A. A third group of subjects, again in a separate evaluation condition, sees just Dictionary B and never sees the information on Dictionary A at all. Again, remember, we're looking at this idea of evaluability. How good or bad is the number that's presented to people?

What we find is that in the separate evaluation condition, when subjects see only Dictionary A, they see that it was published in 1993 and that it has 10,000 entries. It's a used dictionary, and it's like new. That's good. There are no defects in this dictionary. They tend to price that dictionary fairly high, at approximately $24. Subjects in a separate evaluation condition where they see only Dictionary B see that it was published in 1993 and that it has 20,000 entries. How good or bad is that? I don't know. But you know what? There's a defect. The cover is torn, and that's bad, and subjects use that "badness" as a way of helping them to price this dictionary. They actually price it at approximately $20, lower than Dictionary A. But of course, people who see the dictionaries side by side and who evaluate both Dictionary A and Dictionary B price the dictionaries based on the most important attribute. That's the number of entries. They're able to do that because of the comparison process. They can see that the cover is torn on Dictionary B, but they can also see that 20,000 entries is much better than 10,000 entries. That's what

they price higher in that joint evaluation condition. They price Dictionary A at approximately $19 and Dictionary B at approximately $27.
It's a little hot in New Orleans, so let's talk about ice cream. This is another example of evaluability. In this example, subjects, again, are in either a joint condition where they see two different ice cream vendors' offerings of ice cream or they see just one of the two (they don't see the other one). In the separate evaluation condition, subjects see that Vendor H has eight ounces of ice cream. I'm not sure how to value eight ounces of ice cream. I don't know how good or bad it is. But you know what? That cup is underfilled. I don't like that very much. The subjects tend to price that at $1.66, fairly low. In the separate evaluation condition, subjects see that Vendor L has seven ounces of ice cream, but that cup is overfilled. That looks like a lot. That's really good. The subjects tend to price Vendor L's ice cream higher, at $2.26. However, in the joint evaluation condition, when people have the two vendors side by side and they're able to do a comparison of the numbers and draw meaning from that comparison, suddenly Vendor H's ice cream is valued more than Vendor L's, at $1.85 and $1.56 respectively. This is what's called, by the way, a "preference reversal" in my literature. There's a reversal of preferences depending upon how you present information to people, in this case whether it's a joint evaluation condition or two separate evaluation conditions.

Results of these experiments, as well as many others, lead us to believe that information in decision-making is processed using two different interacting modes of thinking. One is deliberative, and one is more experiential. The deliberative mode of thinking is analytical. This is what you actuaries do—logical thinking. You work your way consciously, logically and sometimes slowly through information, in order to figure out the right thing to do. It's based on fairly recent evolutionary history. The experiential mode is also used, though. In an experiential mode of processing information, we primarily use affect. We use feelings. Thinking through things experientially is based on intuitive and holistic processing. It's based on our experiences. It's based on what we've learned throughout life. Processing in this system tends to be fast, and it can be less than conscious because we can be seized by our emotions.

Both these modes of thought interact with one another. What we think about influences how we feel. How we feel influences what we think about. Both modes of thought are important to making good decisions. Let's first talk about the deliberative mode a little more. In your field, as well as in mine, the traditional high-reason view of decision-making has to do with deliberation. It has to do with thinking hard and maximizing utility. The problem is that we have limited capacity to process information, to manipulate it and to use it where, as John mentioned earlier, we are boundedly rational. An implication of our bounded rationality is that we tend to rely on information as it's given to us. Because of that, we show what are called "framing effects."

Here's an example of a framing effect experiment. Subjects in this study (subjects were patients as well as physicians) were shown two options for treating lung

cancer, both surgery and radiation. They're shown the information, though, in one of two frames. One is a survival frame where, for example, 90 of 100 people would survive surgery, 68 would still be alive at one year, and 34 would still be alive at five years. One hundred of 100 people would survive radiation, 77 would still be alive at one year, and 22 would still be alive at five years. A different group of subjects was shown the information in a mortality frame. Now this is normatively equivalent information. So, if 90 people survived surgery, then 10 must have died during surgery. Thirty-two would have died at one year and 66 would have died at five years following surgery. Similarly, zero people would die during radiation, 23 would have died at one year, and 78 would have died at five years. The second group of subjects was given this mortality frame. How the information was framed ended up making a big difference to this very important treatment choice. Surgery was less attractive in the mortality frame (58 percent chose surgery) compared to the survival frame (75 percent chose surgery). People made different choices based upon how that information was framed to them.

Also because of bounded rationality, decision-makers use other simplifying heuristics, such as anchoring and adjustment. I want to give you an experimental example of anchoring and adjustment. This is a study that we did on medical errors. We asked employed-aged adults to estimate the number of fatalities from medical errors in the United States each year. The average figure that these employed-aged adults gave us was close to the figure that had been estimated by the Institute of Medicine. However, the estimates depended considerably on an anchor that we gave them. Prior to asking them to make their actual estimate, we either mentioned that appendicitis causes about 400 deaths per year in the United States, or we told them about kidney disease, which causes about 40,000 deaths in the United States each year. The median estimates given by subjects depended greatly on whether we gave them the 400 anchor (subjects estimated 3,000 deaths) or the 40,000 anchor (subjects estimated 15,000 deaths).

We use our deliberative abilities in making decisions. However, because of bounded rationality, we don't necessarily make decisions in ways predicted by expected utility theory. But then, in addition to processing information deliberatively, we also process information based on our feelings. We process information experientially. An affective and experiential view of decision-making suggests that our affective feelings guide decisions, and they guide perceptions of information, like perceptions of risk information. In fact, affect acts as a source of information and as a source of meaning in decisions, so that consideration of possible bad outcomes will lead to unpleasant feelings in the gut that lead us away from choices that are bad for us. That's, in part, how we choose affectively and experientially. Consideration of possible good outcomes leads to pleasant feelings in the gut, and that acts as a beacon of incentive, drawing us toward those choices that might be better for us. Some of the examples that I've been using and some of the further examples that I will be telling you about all suggest that without affect, information lacks meaning. We need to be able to have an affective reaction to information, including numbers, for those numbers to make a difference to how we judge and decide. Decision-

makers need to be able to consider information carefully using a deliberative mode, but they also need to be able to understand and to be motivated by the meaning that underlies that information.

I want to make sure I defined "affect" for you. What I mean by that is just the positive and the negative feelings that you have about an object or an attribute or an event. They're very subtle feelings that are experienced as you think about some object. For example, it might be your feelings about sunshine—just the brief, subtle, positive feelings that come up when the word appears on the screen—or your feelings about a funeral. They are very subtle positive and negative feelings that are experienced.

In my group we've developed something called the affect heuristic. The affect heuristic predicts and explains how good and bad feelings about hazards and technologies are associated with perceptions of their risks and their benefits (not their actual risks and benefits, but what people perceive as their risks and their benefits). In your work you look at risk as being derived through analytical means. The affect heuristic predicts instead that decision-makers will use their affect toward an activity or a decision option in order to guide their perceptions of its risks and its benefits. In the real world, risk and benefit are positively correlated. Things that are high in risk have to end up being high in benefit, or they're not going to be around in the real world for very long. If they're high in risk and low in benefit, they just don't last. They go away in our marketplace today.

However, in people's minds, risks and benefits are actually negatively correlated. The relationship between risk and benefit is such that things that are perceived as high in risk tend to be perceived as lower in benefit. Things that are perceived as high in benefit tend to be perceived as lower in risk. The strength of that inverse (negative) relationship between perceived risks and perceived benefits depends on the degree to which that hazardous activity is judged to be good or bad. Things that are judged as good—things that I have positive feelings about—tend to be perceived as high in benefit and low in risk. Conversely, things that are judged as bad tend to be perceived as low in benefit and high in risk. You can see this relation in medical and nonmedical exposures to radiation and chemicals. For example, medical sources of exposure to radiation are perceived as good overall, while nonmedical sources tend to be perceived as bad. It's the same radiation, but the medical sources of exposure have more favorable benefit risk ratings than do the nonmedical sources.

Why should we rely on our feelings? It may be that we know our feelings better than we know our thoughts, because they're simpler. Even though our feelings aren't always clear about what to do and what not to do and what's right and what's wrong, in the end, feelings may be simpler than the distracting multitude of logical reasons, and so people may rely on them. In fact, I would encourage you to think about affect as a source of information. We talked before about this idea that unpleasant feelings can act as an alarm. When considering something that's bad for

you, unpleasant feelings can develop in the gut, sound an alarm and lead you away from those bad choices, while pleasant feelings can act as a beacon of incentive. Affect can be a quite helpful guide in judgments and decisions. It doesn't always help, though. Sometimes a reliance on feelings can lead us to pay attention to irrelevant feelings, like what our mood is right now, for example.

In an experimental study that we conducted in Sweden, we induced a positive mood experimentally in half of our subjects, and we induced a negative mood experimentally in the other half of our subjects. Then we offered all subjects the opportunity to purchase a real lottery ticket with their own real money. The lottery ticket was worth a real 2 percent chance to win a real 1,000 Swedish Kroner, which is about $130. People were very excited about this; these were college students in Sweden. What we found was that those subjects in whom we induced a positive mood were willing to pay almost 25 percent more for the same lottery ticket than people in whom we had induced a negative mood.

The presence of affect has a variety of consequences, including an influence on judgments like perceptions of risk or likelihood. It turns out that strong positive and negative events (things that we get emotional about) were sensitive to possibility rather than to probability. Rottenstreich and Hsee conducted a study in 2001 where they asked subjects how much they would be willing to pay to play one of four lotteries. They asked subjects, "How much would you be willing to pay for a certain (100 percent) chance to win $200?" They were willing to pay about $200. They asked a separate group of subjects, "How much would you be willing to pay for a certain (100 percent) chance to meet and kiss your favorite movie star?" This is the affect-rich condition. People get excited about this. They get to think about their favorite movie star. They had two other groups of subjects who were told that they had just a possibility, a 1 percent chance, to win $200 or a 1 percent chance to meet and kiss their favorite movie star.

At certainty, subjects were willing to pay far more for the cash than for the kiss. They were willing to pay about $30 for the kiss when it was a certainty. That was the median amount they were willing to pay. However, when there was only possibility, when there was only a 1 percent chance, subjects were willing to pay far more for the kiss than they were for the chance at cash. The attractiveness of the kiss lottery was not much affected by the probability of winning. So with strongly affective goods, we're not as sensitive to changes in probability values. We tend to show what's called "probability neglect." It may be that this neglect of probability with strongly affective goods is part of what drives our country's mania for lotteries. One man says, "I have six kids I have to take care of. That lottery winning will buy a lot of everything. Maybe I can quit my job." This, by the way, was said when the odds were one in 80 million.

With strongly negative events, such as terrorism, probability neglect may also happen. People's attention may be so focused on the bad outcome itself that they fail to attend to the fact that it's actually very unlikely to occur. They show

probability neglect. Similar results are found in intuitive toxicology. When exposure to chemicals can produce dreaded high-affect effects, like cancer, the public tends to be insensitive to different exposure levels. So the public, with its stronger affective reactions, in comparison to toxicologists, responds such that if large exposures to some chemical are bad and risky, then small exposures also are bad and risky. They're relatively insensitive to differences in the actual probability of harm.

How we feel about possible events seems to strongly influence the risks that we perceive, what we identify as good and bad influences, how we process information and ultimately how we make decisions. Affect seems to convey meaning upon information that it doesn't necessarily have on its own, and without affect, information appears to lack meaning, and it tends not to be used in judgments and decisions. Affect, in fact, is a key ingredient of rational behavior. It can also, however, lead to poor decision-making.

Now I want to talk about these dual processes of affect and deliberation and how they also seem to balance one another. So while affect can inform deliberation, and deliberation can influence how we feel, they also seem to be in balance with one another. Let me give you an example of this. There was an experiment done at the University of Iowa where subjects came into a room and were told that they were going to do a study about memory and a change in environment. They told the subjects, "Oh, by the way, you're going to be in the condition where we're going to give you something to memorize here, and then we're going to make you switch to a different room and change your environment and you tell us what that number is that we asked you to memorize originally in the first room. We want you to report it back to us in the second room."

Now, we psychologists sometimes don't actually tell subjects exactly what we're studying. On the way out of the room, each subject was given a number to remember, and the subject was supposed to walk down to a second room. However, half the subjects were given a large, seven-digit number to remember. It takes a lot of deliberative capacity to keep that seven-digit number in mind so that you can report it back in the next room. You're going to use up a lot of your deliberative capacity, and there's not much deliberative capacity left over to maybe control for affect that might appear in front of you. The other half of the subjects were given a two-digit number. It's a fairly small number. Then all the subjects, one by one, walked down a hallway to a second room. In the hallway was the real experiment. They were told, "Thank you so much for taking part in our study. We really do appreciate it. As a token, just a small token, of our appreciation, we'd like to give you a snack. Would you like to have this luscious chocolate cake, or perhaps you'd prefer this sort of dried-out fruit salad? Which one would you prefer?" Subjects simply picked up a tag of whichever one they preferred, and then they went on to the second room where they would report back their number and presumably get their snack.

What these experimenters hypothesized and found was that those subjects who had memorized a very large number and had low capacity left over to do more deliberation and perhaps to control for emotional inputs were much more likely to choose the chocolate cake than those subjects who had remembered a very small number. When deliberative capacity is reduced, a decision of the heart becomes more likely.

What about when deliberative capacity is increased? A set of researchers at the University of Virginia did a study where they increased deliberation. They had undergraduate subjects passing by a table out on the quad, and they told subjects that they could choose a poster and take it back to their dorm room. "In payment for being given this poster for free, we're going to need to call you back in a few weeks." Subjects chose either a humorous poster or an art poster, but subjects were in two different conditions. Half of the subjects gave reasons for their choice prior to making the actual choice. They increased their deliberation by being required to tell the experimenters the pros and cons of their choices. Now according to an expected utility model, they should make better choices. Since they thought harder about it, they're going to make better choices and they'll be more satisfied later with their choice. The other half of the subjects merely chose a poster, and then they moved on.

The subjects went home. They put the posters up in their dorm rooms. Three weeks later, the researchers gave them a call back and just simply asked, "Is the poster still up, and how satisfied were you with your choice?" What they found was that the greater-deliberation-and-the-reasons condition led to worse choices. Subjects who gave reasons were less satisfied three weeks later than the subjects who merely chose. It appeared that greater deliberation reduced the input of affect, which happened to be important to making a satisfying choice about something aesthetic, like a poster that you're going to put up in your room. So these dual processes of affect and deliberation seem to exist in balance with one another. As you increase the input of one, you may decrease automatically the input of the other.

I'd like to switch gears. I don't know a lot about actuaries, but I do know you are good at math. I thought I'd talk a little about math and how other people understand mathematical concepts and numbers. In particular, I'm going to talk about the ability to understand and use probability and mathematical concepts. I'm going to talk about it in a way that was developed in the cancer literature by Isaac Lipkus and some of his colleagues. Actuaries are very good at understanding probability numbers, but what about laypeople?

I have a measure of numeracy that I use. I ask a series of questions of younger adults, and I ask some of the questions of older adults as well. Ordering the questions from the easiest questions down to the hardest questions, the first question is: Which of the following numbers represents the biggest risk of getting a disease: 1 in 100, 1 in 1000 or 1 in 10? This is not a trick question; this is a very

simple question. When I ask my undergraduate subjects, about 96 percent of them do get that correct. However, with the older adults, only 76 percent were able to get this correct. The older adults are people who are 65 and older. They probably range in age, again, from 65 to about 95, if I had to guess. I don't know the exact representation here.

There are other questions as well: If Person A's risk of getting a disease is 1 percent in 10 years and Person B's risk is double that of A's, what is B's risk? So among the undergraduate subjects, 83 percent got it correct. Only 66 percent of the older adults got that correct.

You see the same pattern throughout all the questions. Some people get every one of them wrong. Older adults tend to understand these kinds of probabilistic concepts less well than college students. The very hardest problem is: In the Acme Publishing Sweepstakes, the chance of winning a car is 1 in 1,000; what percent of tickets win a car? Fewer than half of the undergraduate subjects that I tested got that correct. I didn't ask my older adults because they were doing so much worse on the easier questions that I didn't want to ask them all the questions.

What does this say about the assumptions you might make concerning what your customers understand and what they don't understand? Many choices are numeric. All of the choices, I believe, that you deal with involve numbers. For example, in retirement choices, the three main choices are the quantity of money to save, the proportion of savings to allocate to different areas and retirement age. Imagine that you're a senior and you have to make a Medicare prescription card choice. At one point there were 72 options available, each one with different prices on different medications, and the prices were allowed to change weekly. What about choosing a treatment for cancer? In this day of informed choices, you'd be presented with the risks and benefits of multiple treatment options and asked to share in the decision-making process with your doctor. What happens to people who are less numerate and who don't understand these kinds of numbers? How might information processing differ, depending upon how those numbers are presented and how people draw meaning from the numbers?

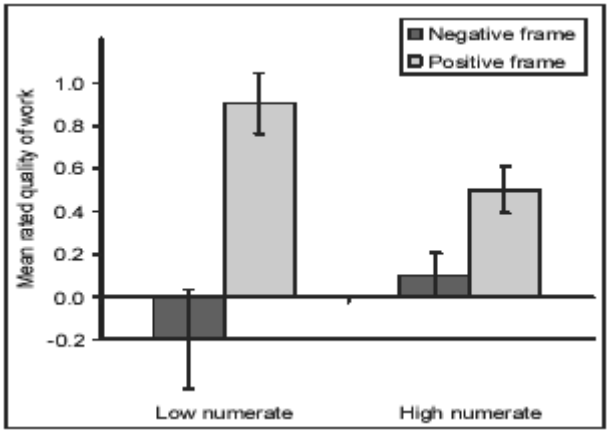
We started looking at this question with respect to decision-making. We compared people who are lower in numerical ability with people who are higher in numerical ability. The ability to understand numbers will influence a variety of things. It will influence how much a decision-maker transforms numbers from one format to another, so it will influence framing effects. John talked about this earlier. We also talked about framing effects with the experiment on framing treatments for lung cancer in terms of survival versus mortality. We described the impact of frames as being due to people relying on information in the form that it's given. It may be, however, that people who have greater numberability may be more likely, or at least more able, to transform numbers from one format to another. We did an experiment to look at this.

This experiment was done with college students. Subjects were given either a positive frame or a negative frame. In the positive frame, "Emily" was described as getting 74 percent correct on her exam. Then those subjects were asked to rate the quality of her work. In the negative frame, a different set of subjects was told that "Emily" received 26 percent incorrect. Again, subjects were asked to rate the quality of her work.

In Chart 5, I show low numerates on the left and high numerates on the right. The bars represent the mean rated quality of work, depending upon whether the information was provided in the positive frame or the negative frame. A bigger framing effect is represented by a larger difference between the positive-frame bars and the negative-frame bars. As you can see, the low numerates show a much larger framing effect, compared to the high numerates. This suggests that the high numerates may not rely on information in its given form as often. You are high in numeracy. You probably automatically transform information from 26 percent incorrect to 74 percent correct, and you can do that quickly and easily. Not everybody can, and, as a result, some people rely much more on information in the form that it's given.

Chart 5

Numeracy and Attribute Framing



Let's talk about another example. John mentioned earlier this idea of communicating risk information in a probabilistic way, as 1 percent, or in a frequency format, as 1 chance out of 100. We looked at this idea. We looked at whether probability and relative frequency are the same or different in terms of communicating risk. We told subjects about a patient, Mr. James Jones, who has been evaluated for discharge from an acute mental health facility. A psychologist has done an assessment of Mr. Jones, and among the conclusions reached in the psychologist's assessment is the following (here we split subjects into two different groups). Subjects are either shown a probability condition "Of every 100 patients similar to Mr. Jones, 10 percent are estimated to commit an act of violence," or

they're shown the frequency format "Of every 100 patients similar to Mr. Jones, 10 are estimated to commit an act of violence." The only difference between what subjects were shown is whether they were shown the number "10" or the number "10 percent." Every other word is identical.

Past studies have shown us that risk information shown in a frequency format elicits more affect. It elicits images—in this case, of a dangerous patient—and it drives greater risk perceptions compared to the same information shown in a probability format. In this experiment, we hypothesized that the high numerate would be more likely to transform the information from one format, whatever the high numerate was given, into the other format, so that the given format wouldn't make as much difference to his or her risk perceptions, compared to the low numerate, who is going to be greatly influenced by the given format. That's exactly what we found.

Subjects were asked to describe the riskiness of Mr. Jones on a scale ranging from one to six, I believe. Low numerates perceived Mr. Jones as being significantly riskier when that risk was described in frequency terms; the average perceived risk for the low numerates was 3.9, compared to the average perceived risk of 2.5 when that risk was described in probability terms. The format of the information, however, didn't make much of a difference to how high numerates perceived that risk. When that risk was described in frequency terms, the average perceived risk for the high numerates was 3.8, compared to the average perceived risk of 3.7 when that risk was described in probability terms.

The ability to understand numbers will also influence how much numbers are used versus how much irrelevant sources of information influence choices instead. We did a study that we call the jelly bean task. This is also called a ratio-bias task. Subjects were shown Bowl A and Bowl B. Bowl A has 100 jelly beans, and nine of them are colored. They're told explicitly that underneath the bowl there are 9 percent colored jelly beans. Bowl B is a smaller bowl. It only has 10 beans, and one of them is colored. Again they're told explicitly that underneath the bowl there are 10 percent colored jelly beans. Then they're told, "If you pick a colored jelly bean, you're going to win $5."

From which bowl would you prefer to pick a jelly bean? Anyone here want to pick from Bowl A? I'm betting no. In fact, what we find is that very few people who were high in numeracy were likely to choose from Bowl A, but among those subjects who were low in numeracy, more than six times as many wanted to choose from Bowl A. Why would that be? It turns out that the number of beans in Bowl A was attractive. This is what subjects will say to me. "I know that nine is smaller than 10, but there are nine beans there. I can see them. I feel like I have a better chance of winning. There's only one winning bean in Bowl B." So they're more likely to pick from Bowl A. They're lured by this irrelevant source of information. When we asked subjects about their feelings, we asked, "How clear are your feelings about the goodness or badness of the 9 percent chance in Bowl A?" What we see is that those people who are low in numeracy have a less clear idea of the goodness or badness of that 9

percent chance. They don't understand the meaning as well of the 9 percent probability.

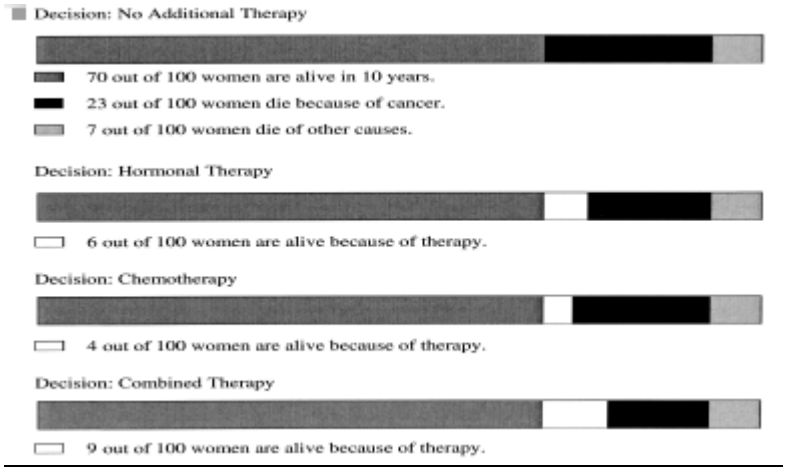
The ability to understand numbers seems to influence how much feeling is drawn from numbers. Let's look at another example. It turns out that it's not the low numerates who will always be the less rational ones. You might remember the bets example earlier. Group 2 gets an objectively worse bet, but they actually rate the bet as more attractive. What we hypothesized is that this effect would interact with numeracy and that it would actually be the high numerate who would show the less rational effect. Why would that be? It's because they're good with numbers and they'd be more likely to draw meaning from the comparison of numbers, while people who are low in numeracy perhaps wouldn't.

As we hypothesized, the high numerates are the ones who drive this surprising effect. They're the ones who find the objectively worse bet more attractive, rating it at 14.4 but rating the certain $9 bet at 9.1. In particular, the high numerates are the ones who feel the best about the $9. If I ask them to rate their affect toward the $9—how good or bad is the $9—it's the high numerates, who have the comparison to the five-cent loss, who rate that $9 as being particularly good. The $9 comes alive with feeling to those people because they're more facile with numbers in numerical comparisons.

So does numeracy matter? It turns out that it does. Here's an example in a more real-world context. A decision aid called Adjuvant is used by oncologists down at Duke University. The purpose of Adjuvant is to help breast cancer patients to better understand the risks associated with different therapies. These are women who have had surgery, by the way. Chart 6 is the output from the actual decision aid. It shows the survival rates depending upon whether you get no additional therapy, hormonal therapy, chemotherapy or if you choose to get a combined therapy of hormonal and chemotherapy.

Chart 6

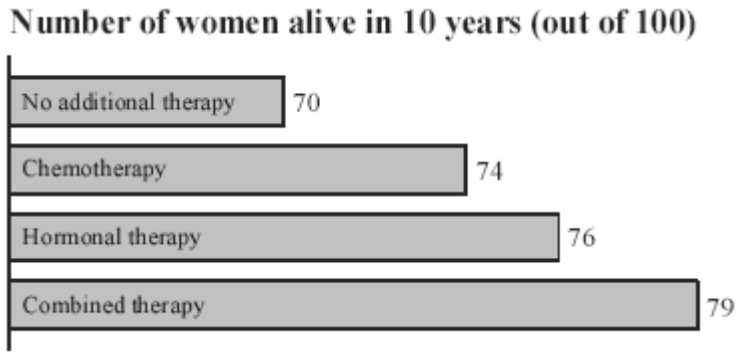
Adjuvant Decision Aid



We showed high- and low-numerate women this information, and we asked them comprehension questions. How many people would survive if they received hormonal therapy? Which therapy has the greatest likelihood that you'll be alive in 10 years? We asked simple comprehension questions. But since I'm an experimentalist, I only gave half my subjects this format. I also gave them normatively equivalent information, because I thought there was lots of redundant information in how Adjuvant presented risk information to the patients. I simplified the format. I looked at an improved format that showed the normatively equivalent risk information, as shown in Chart 7, and I asked the same comprehension questions of my low and my high numerates.

Chart 7

Improved Format of Decision Aid



What I found was that with Adjuvant, comprehension was fairly low, but it was higher for people who were higher in numeracy. They better understood what was going on. But if presented the normatively equivalent information in a simplified and improved format, comprehension for everybody rose quite a bit. I had six questions, and everybody got an average of 5.3 of those six questions correct. In fact, the low numerates improved so much that they looked just like the high numerates on comprehension.

The ability to understand numbers influences a variety of ways that people process information and how they choose and decide and how they understand risks. Numeracy ability varies, and it makes a difference to decisions. How you present information and how you present numbers can influence how people feel about those numbers and how much they comprehend. Affect seems to act as a source of information in both affect and deliberation. Both modes of thinking matter in judgments and decisions, like perceptions of likelihood and risk.

In terms of implications, less may be more. If it's not critical to show some information to your consumers or if it's not critical to present some choice options, you may, as an information provider, want to choose less because more complete information may lead to less comprehension and less motivation to act and decide. More choices and more information are not necessarily better, and they may be worse. Second, you can't present "just the facts" because "just the facts" doesn't exist. How you present data will influence the choices that people make around you. It suggests that you need to make thoughtful and defensible choices of how you present that data in order to support good, informed choices. Finally, in communication about choices, numbers are just that. They're just numbers, but we can provide more meaning to those numbers through affect and other means. How we present numbers will influence how decision-makers feel about them and how they understand them, even when it comes to very familiar numbers like dollars.

MS. LORI WEYUKER: I have a question for both of you with respect to the media. When the media presents numbers on TV, for example, often they present numbers that are just in gross and maybe not per capita. An example is the recent increases in the price of gasoline. I think this can be very misleading, as opposed to comparing how much the price of gasoline has increased versus inflation, for example. Do you think the media is doing this as a strategy to influence the public or because they're unaware?

DR. PETERS: There's more than one answer to that question. One answer is that I think people in general tend to be unaware that how you present information does make a difference to how people perceive it. A second answer, though, is that—I believe John mentioned this earlier—there is this idea that negative things loom larger than positive things. The media's job is to sell their media; that's what they do for a living. Presenting information in a negative light may actually get them more readership, because it's more attractive to people in some senses. They notice negative information more than they notice positive information.

MS. MARILYN KRAMER: In consumer-directed health plans we're asking consumers to make two kinds of choices, a benefit design as well as a choice of providers (making choices about a tiered network, for example). We're asking them to do two very tough things for an average consumer. What does this mean in designing plans for the average consumer?

MR. STARK: I think about this more in the multi-choice situation that we've been dealing with for years where we offer PPOs, HMOs and so on and ask people to make similar choices. I believe we're asking too much of people. If you're conscious of money, you may just choose the lowest premium. But how do you define a good network? If you're at a cocktail party, what you'll find out is that the best network is the one that has my doctor. That's another way people make decisions, not necessarily on whether the benefits are best for them. As Ellen said, there's a lot of affect. There's a lot of emotion around what goes on and how people make choices. I don't think that we, as actuaries, take into account how strong that is.

DR. PETERS: We're doing some experiments right now where we're looking at this choice between consumer-directed health plans and PPOs. One of the things that we're planning on looking at within the experimental studies is how you can you present information in ways that might make the consumer-directed health plans easier to understand. Given that you're doing it, how can you at least present that information in ways that make the numbers more accessible to people? We'll see how well it works.

MR. STARK: In this age of Sarbanes-Oxley, litigation, reputational risk and media scrutiny, insurers need to be very careful. If we don't do a responsible job of presenting information and we use it for bias, we could get in a lot of trouble.

MS. KARA CLARK: Ellen, do you have a sense of what percentage of the population is considered highly skilled in numeracy? How does that vary by population segments or even in different countries around the world?

MS. PETERS: There's very little information on it. The numbers that I do know don't have to do with highly skilled. There has been some research done by Irwin Kirsch that suggests that about half of the American public is not numerate enough to be able to understand the tables and charts that are presented in our media today. That doesn't have to do with highly skilled; it has to do with low skill. I also know that numerical ability shows a decline with age. I don't know whether that's a cohort difference in terms of there are just different skills that have been taught at different age groups, or whether that's linked with cognitive decline. But there is a substantial difference in people's numerical abilities from younger age groups to older age groups. I don't know the answers in other countries. I've never seen any data like that. This is a fairly new area of study in the United States.