Risk Analysis, Vol. 29, No. 8, 2009 DOI: 10.1111/j.1539-6924.2009.01241.x

Communicating About Xenotransplantation: Models and Scenarios

Wändi Bruine de Bruin,[1]* Ümit Güvenç,[2] Baruch Fischhoff,[1,2] Christopher M. Armstrong,[3] and Denise Caruso[3]

Xenotransplantation entails using organs from genetically modified animals as a way to solve the shortage of human organs for transplantation. As with other novel technologies, if xenotransplantation is to be judged fairly, proponents must explain its complex, uncertain, and unfamiliar risks and benefits. Xenotransplantation's risks include the possibility of a recombinant virus infecting human transplant recipients, potentially causing an epidemic of an unfamiliar disease. Using materials vetted by scientific experts, we communicated the variables and relationships determining this risk in three formally equivalent formats: (a) a graphic model, (b) scenarios structured by the graphic model, and (c) both the model and the scenarios. Participants were randomly assigned to receiving one set of materials. They rated them as equally clear and studied them equally long, suggesting similar ease of cognitive processing. Compared to participants receiving the scenarios, those who received the graphic model better identified causes and effects of the risk, and saw less risk of xenotransplantation. Participants who received both the model and the scenarios generally showed intermediate responses. The study demonstrates a general procedure for developing and evaluating formally equivalent graphic and scenario communications regarding highly uncertain risks. In this application to xenotransplantation, presenting a graphic representation improved people's understanding of the risk.

KEY WORDS: Mental models; risk perception and communication; xenotransplantation

1. INTRODUCTION

Transplantation is often the only effective therapy for patients with end-stage organ failure. However, a shortage of human organ donors limits the chances of receiving one. Over the past five years, some 6,500 patients died annually while waiting for a donor organ.(1) Many other patients are just getting by, with badly impaired organs and decreased quality of life.

The transplantation of genetically engineered animal organs to humans, or *xenotransplantation*, is one proposed solution to the shortage of human organs.(2) The threat of organ rejection, however, limits the feasibility of this biomedical technology. Transplanting whole organs from genetically modified pigs to primates has not yet been successful enough to warrant human implantation.(3) Transplantation of porcine islet cells to human patients with type 1 diabetes may be closer to widespread clinical application.(4)

As xenotransplantation becomes more feasible, policies are needed to manage the attendant risks,

[1] Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, USA.
[2] Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA, USA.
[3] Hybrid Vigor Institute, San Francisco, CA, USA.
* Address correspondence to Wändi Bruine de Bruin, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA; tel: 412-268-3237; fax: 412-268-6938; wandi@cmu.edu.

0272-4332/09/0100-1105$22.00/1

the most worrisome of which is zoonotic disease.[5] Interactions between farmers and pigs have led to zoonotic outbreaks throughout history. The more intimate contact created by tissue transplants could open novel avenues for infections, most notably the possibility of porcine endogenous retrovirus (PERV) "jumping" species. If that happened, a human transplant patient might become infected with an unaltered or mutated PERV or a recombinant virus that includes elements of porcine and human viruses. Were that infection readily transmissible among humans, it could introduce an epidemic of a novel virus dangerous to humans.[6,7] Additional risks include those of transgenic pig carcasses contaminating the food supply, transplant recipients passing along novel retroviruses to children conceived post-transplantation, and healthy pig organs being killed rather than rejected, through their contact with the viruses and bacteria in a human body.[8] Additional ethical issues include surveillance of patients and their sexual partners, allocation of limited resources, and animal welfare.[9,10]

As with other technologies raising novel risks and ethical concerns, policy decisions regarding xenotransplantation will be shaped by the opinions of a public that is currently largely ignorant of the prospect. Recognizing this ignorance, surveys of public attitudes toward xenotransplantation have found it necessary to explain the risk of cross-species transmission.[11–15] Receiving such information has typically made lay people feel less favorable toward xenotransplantation.[12,13]

How well such survey studies predict actual public responses to xenotransplantation depends on how well they anticipate the information that the public will eventually receive. It is difficult to predict those information flows with any confidence. One can, however, question whether a survey has presented a fair and adequate picture of xenotransplantation's risks and benefits.[16–18] Studies designed to improve the public's understanding of other technologies have found that doing so can decrease or increase acceptance. Learning about carbon capture and sequestration has led people to like it less,[19] while learning about using nuclear energy sources in spacecraft has led people to like it more.[20]

These last two studies used Morgan *et al.*'s *mental models* approach to create communications,[21] seeking a clear and fair picture of the processes creating and controlling the technologies' risks and benefits. Morgan *et al.*'s mental models approach builds on psychology's long tradition of mental models approaches, characterizing people's intuitive theories of complex domains in their natural formulation.[22–25] It is designed to capture beliefs regarding complex, open systems, like those that determine many risks (in contrast, say, to intuitive representations of syllogistic reasoning). It has been applied to diverse health risks, such as domestic radon,[26] breast implants,[27] vaccines,[28] and pandemic influenza.[29,30]

The first step of Morgan *et al.*'s mental models approach[21] involves an *expert assessment*, summarizing decision-relevant scientific knowledge regarding the processes determining a technology's effects. Typically, these take the form of graphic models akin to influence diagrams,[31,32] with nodes reflecting variables and arrows the relationships between them. Second, a *lay assessment* characterizes what members of the target audience currently believe. A comparison of the two assessments reveals what people still need to learn in order to make more informed decisions about the risk. When people already have a basic mental model of the facts, communications may focus on closing critical gaps, as well as on improving behavioral and emotional skills needed to act on that knowledge.[21,33] However, when people know little, communications must convey the big picture of the decision-relevant variables and the relationships among them. Xenotransplantation presents a communication challenge of the latter kind.

The study reported below provides a general approach to creating and evaluating fair and comprehensible communications, applied to xenotransplantation. It contrasts the effectiveness of two basic presentation modes—graphic models and scenarios—deriving both from the same expert assessment, so as to ensure the equivalence of their information content.[30]

The *graphic model* incorporates the main features of an influence diagram, along with text briefly describing the presented variables and relationships. Formal influence diagrams can produce quantitative estimates of risk (if their data needs are met). For communicating basic understanding, however, it is enough to treat the graphs as qualitative formal models, whose elements reflect a rough quantitative screening for decision-relevant facts. Such precision ensures that the key facts have been assembled and their relationships defined.[30,34,35]

The *structured scenarios* describe specific instantiations of the variables and relationships in the expert assessment or graphic model, adding concrete detail and context. Unlike ordinary narratives, structured scenarios explicitly include each variable and relationship, trying to eliminate any need to read between the lines.[30] As with any narrative, a structured scenario has so much detail that it has a near-zero probability of happening exactly as described. Therefore, a set of scenarios is presented to explain how the complex pieces fit together, telling different stories of what might happen.

In medical decision-making research, graphic displays have been used to communicate health risks to patients and other members of the general public, in terms of absolute risk magnitude, relative risk, cumulative risk, uncertainty, and interactions among risk factors.[36] Influence diagrams and other graphical displays showing how risks emerge and can be controlled have typically been presented only to professional audiences such as technical experts and policymakers.[30] Risk communications aimed at lay audiences tend to present such information in textual form.[30] We know of no studies that systematically compare lay responses to formally equivalent graphic and scenario-based communications.

In terms of basic psychological principles, graphical and narrative communication methods each have potential advantages and disadvantages. Graphic models have the advantage of making it immediately apparent what the relevant variables are and how they are related.[37] However, as models become more abstract and complex, lay people may experience cognitive overload, more so than experts accustomed to navigating such displays.[38,39] Narrative scenarios have the advantage of providing compelling stories, while leading people through the elements of the expert assessment. However, such stories can make specific outcomes unduly easy to imagine, leading users to overestimate their likelihood without realizing how the presented conjunction of multiple events lowers the overall probability, possibly to near zero.[28,40,41]

In the present study, we compare graphic models and structured scenarios explaining the risks of a xenotransplantation patient becoming infected with a recombinant virus. In a between-subject design, an educated sample of lay people received information covering the same variables and relationships, either in the form of a model, three scenarios, or both sets of materials. Presenting both forms might build on each method's strengths, allowing recipients to triangulate on the two perspectives,[30] or it might produce confusing information overload.[39]

We compared these communication strategies in terms of recipients' (a) ease of processing, (b) depth of understanding, and (c) evaluations of xenotransplantation. Ease of processing was measured in terms of time spent answering questions and ratings of the material's clarity. Depth of understanding was measured with performance measures expected to favor each strategy. That is, we predicted that model recipients would be better than scenario recipients at identifying the causes and effects of key variables and at understanding how changes in one variable would influence another variable. By contrast, we expected scenario recipients to be better at producing additional scenarios. We also predicted that scenario recipients would see xenotransplantation as more risky and view it less favorably, given the greater concreteness of the narratives that they received.[28,40,41]

2. METHODS

2.1. Participants

In total, 150 participants were recruited through advertisements on an online bulletin board targeting students at Carnegie Mellon University. One was omitted from the analyses because his written responses were unrelated to xenotransplantation. Participants ranged in age from 18 to 56 years (mean = 23.0; $SD = 7.4$). They were 45% female, 52% organ donors, and 79% pork eaters, with 15% having at least some graduate education.

2.2. Materials

Participants first received a brief written introduction to xenotransplantation, including three possible pathways to viral infection: (1) from an unaltered pig virus in the pig organ, (2) from a pig virus that mutates into an infectious virus, and (3) from a new infectious virus created by recombination of a pig virus and a human virus. They then received detailed information about recombination, communicated by means of the model, the scenarios, or their combination.

Participants in the *model* condition received an influence diagram (Fig. 1) constructed with input from an expert panel, along with a separate sheet providing definitions for the variables presented in the nodes. The model reflected existing scientific knowledge, which was initially assembled from the research literature, then reviewed by experts in biology, public health, and related domains. Its core shows that the probability of recombination (node 15) depends on the probability that a patient's cell is infected with a pig virus (node 9) and a human virus

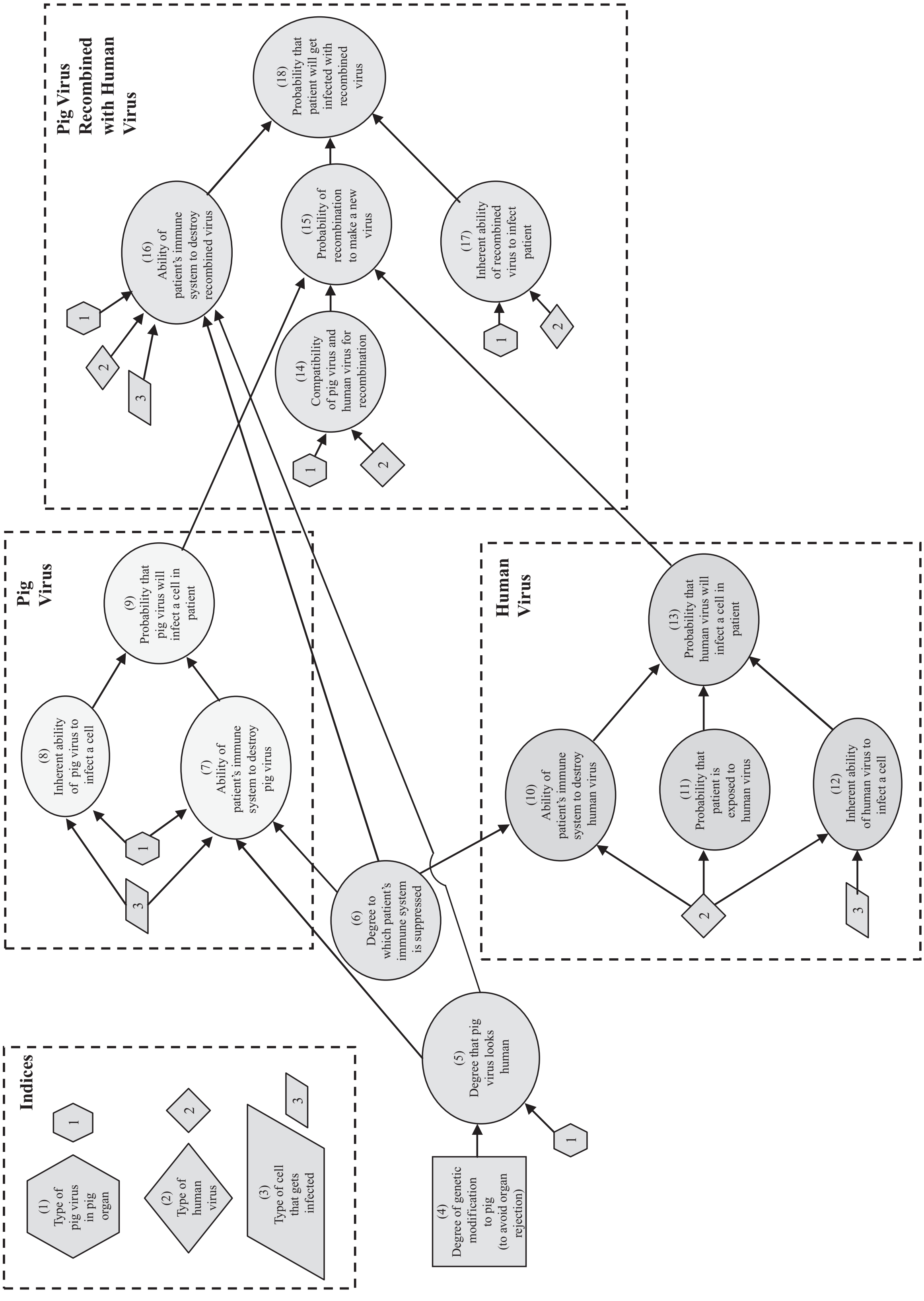

Fig. 1. Influence diagram showing recombination of a human and porcine virus.

Table I. One of the Three Scenarios

Nancy has Parkinson's disease. She suffers from shaking muscles, slow movement, poor balance, a fixed facial expression, and speech problems. Nancy developed the disease in her late 40s. Over the years, her symptoms have gotten worse. Nancy is now 53 years old.
Parkinson's disease slowly damages the part of the brain that controls muscle movement. Some doctors believe that Parkinson's can be treated by putting cells from the brains of pigs into the patient's brain. This procedure is currently being tried on patients. Nancy is one of them.
Because pig cells are so different from human cells, there is a risk that the human immune system will reject them. Genetic modification can be used to make pig cells look less "foreign." However, the immune system is less likely to attack brain cells than other cells in the body. This is why Nancy receives normal pig cells that are not genetically modified.
To be on the safe side, Nancy's doctors are still giving her mild levels of drugs to suppress her immune system. A weaker immune system should reduce the chances that her immune system will reject the pig brain cells.
Yet, Nancy is worried when her doctor explains that there is always a chance that the pig brain cells may be infected with a pig virus that may not be detected with tests. Moreover, in such a case, her immune system may not be able to destroy these viruses because the immunosuppressant drugs make her immune system a little bit weaker. The doctor tells Nancy not to worry too much. Because the pig viruses will come from pig cells that were not genetically modified, her immune system is very likely to recognize them as "foreign" and destroy them. In addition, even if an infection occurs, it may not cause a problem.
Unfortunately, it turns out that the pig cells in Nancy's brain harbor a class of pig viruses called PERV. Although PERV is normally dormant in pigs, the stressful conditions of a transplant can activate them. This is what happens in Nancy's case, and her weak immune system is not able to destroy the virus when it leaves the pig tissue to try to infect the human cells.
PERV does not have a strong natural propensity to infect human cells, but studies have suggested that in rare circumstances they could. Because Nancy's PERV can infect human cells, and her immune system fails to destroy her PERV, a small number of Nancy's brain cells end up being infected with PERV. Fortunately for Nancy, PERV does not make her sick.
While PERV has so far not proved to be a danger to Nancy, there is the concern that it can recombine with a human virus to create a new, dangerous virus. HERV is a type of virus that all people have. Like PERV, HERV is dormant. Because HERV is part of people's genome, it is impossible for the human immune system to destroy them. Therefore, like everyone else, Nancy has HERV.
HERV and PERV are very similar and it is possible for them to recombine if they meet in the same cell. Because PERV has already infected some of Nancy's brain cells that also have HERV, and the two viruses are very similar, they may recombine to form a new hybrid virus. In fact, this happens in Nancy's case.
This new virus could be very dangerous if it is more infectious and stronger than the otherwise harmless HERV and PERV. Fortunately for Nancy, her new hybrid virus is not infectious. As a result, she does not get an infection. Her transplant works out as she has hoped, and the symptoms of Parkinson's are fading.

(node 13). The probability that the patient will get infected with the recombinant virus (node 18) depends on the probability of recombination (node 15) and on the ability of the patient's immune system to destroy the recombinant virus (node 16). That ability lessens the more that the patient's immune system is suppressed (node 6) and the more the pig virus has been genetically modified (node 4) to have human properties (node 5), both changes that may be made to avoid organ rejection.

Participants in the *scenario* condition received three scenarios incorporating the model's variables and relationships, with added context and detail. The scenarios, which the expert panel also reviewed, told of (1) Rudolph, who received porcine pancreatic cells for his type 1 diabetes, (2) Edith, who received a porcine lung, and (3) Nancy, who received porcine brain cells to mitigate Parkinson's disease (Table I). Güvenç[42] presents the full scenarios and their mapping onto the model. Each variable appeared in each scenario, with different values. Edith's was the only scenario in which a human and a porcine influenza virus recombined and caused an infection. The patients in the other two scenarios did not become infected with a recombined virus.

The *combination* condition presented all materials used in the model and the scenario conditions.

2.3. Procedure

Participants were randomly assigned to the model, scenario, or combination condition. Half of the combination group received the model first and then the scenarios; half received these materials in the reverse order. Participants had 15 minutes to study the materials, before answering questions.

Participants kept the materials before them when answering the questions, as in an open-book examination. Before starting to answer, participants were asked to "please look at the clock and write down what time it is." This question was repeated at the end of the questions, allowing us to assess how long participants took to complete the survey.

The first four questions asked for ratings of the *clarity of the materials* on a scale ranging from 1 (extremely unclear) to 7 (extremely clear).

These questions included "how well did this material present the topic to you?" and "based on what you know now, how clear, do you think, is the risk of infection to a transplant patient for scientists?" The next question asked participants to make *policy judgments about xenotransplantation*, on a rating scale from −3 (it should definitely be banned) to 3 (it should definitely proceed), with its midpoint, 0 ("I don't care"). The questions asking participants to rate the clarity of the materials and to make policy judgments about xenotransplantation were repeated after participants had completed all other questions and, presumably, thought about xenotransplantation in greater depth.

After a general warm-up question (not analyzed here), 14 open-ended questions asked participants to *identify direct causes and effects* of specific variables. For example, a question about causes asked, "What are the factors that most directly affect whether a patient's immune system can destroy a pig virus?"; a question about effects asked, "What factors are most directly affected by the degree that the pig virus looks human?"

Next, the survey posed seven questions, asking participants to predict the *direction of influences*, or how changing one (or more) variable(s) would influence another variable. For example, the first question asked: "How would an increase in the compatibility of a pig virus and a human virus affect the chances that they will recombine to make a new virus?" A 7-point rating scale was labeled at 1 (it would decrease a lot), 4 (no predictable difference), and 7 (it would increase a lot).

Participants then read a scenario beginning, "Suppose your mother is a patient who has been waiting hopelessly for a kidney transplant for more than a year" and ending with the mother agreeing to receive the kidney of a genetically modified pig. Participants were then asked to *write a scenario* "about what might go wrong after she receives the pig kidney, based on what you have learned from the study material."

Participants then produced *probability judgments*, on a scale ranging from 0% (no chance) to 100% (certainty). They judged the probability that (a) a recombination event will successfully occur between a pig and a human virus in a transplant patient; (b) a pig virus and a human virus that are biologically compatible for recombination will infect the same cell in a transplant patient and will recombine to form a new virus; (c) a pig virus and a human virus that are biologically compatible for recombination will infect the same cell but will fail to recombine to form a new virus; (d) a transplant patient who has received an animal organ will get infected with a recombinant virus; (e) a transplant patient will get infected with a recombinant virus, as described in a detailed story.

Finally, participants answered demographic questions about their age, gender, education level, organ donor status, and pork consumption. Participants were paid $15 and left their email address to be eligible for an additional award of $50 that would be awarded to participants with the 10 best scores.

2.4. Analyses

When several dependent variables address a common topic, we report average responses and Cronbach's alpha,[43] a measure of internal consistency. Dependent variables based on open-ended responses were coded by two independent judges, one of the co-authors and an undergraduate research assistant. Kappa (κ) scores reflect interjudge agreement, corrected for chance. Pearson correlations reflect relationships between overall scores based on these codes (explained below). The analyses used the codes produced by the co-author.

One-way analysis of variance (ANOVA) was used to examine most differences among the three conditions (model, scenarios, and combined). A repeated-measures ANOVA compared clarity ratings and judgments about xenotransplantation before and after answering all other questions. Effects that passed $\alpha = 0.10$ were followed by *post hoc* Tukey tests, comparing each of the three conditions.

3. RESULTS

3.1. Experimental Check

The three conditions did not significantly differ in terms of participants' age, $F(2, 146) = 51.21$, $p = 0.63$; or percentages of females, $\chi^2 = 3.70$, $p = 0.16$; organ donors, $\chi^2 = 0.67$, $p = 0.72$; pork eaters, $\chi^2 = 4.83$, $p = 0.09$; or those having received a post undergraduate education, $\chi^2 = 1.98$, $p = 0.37$.

3.2. Ease of Processing

3.2.1. Time

Participants spent just over half an hour answering the survey questions in each condition, $F(2, 140) = 0.80$, $p = 0.45$. Means appear in Row 1 of Table II.

Table II. Mean (SD) for Dependent Measures

Measure	Graphic Model	Structured Scenarios	Combination
Ease of processing			
Time needed			
Minutes	32.8 (6.02)	33.5 (7.04)	31.8 (6.70)
Clarity of materials (1–7)			
Rated at the beginning of the survey	4.24 (0.89)	4.50 (0.84)	4.59 (0.98)
Rated at the end of the survey	4.23 (1.13)	4.56 (0.93)	4.70 (1.17)
Depth of understanding			
Identifying direct causes and effects			
Sensitivity (% listed correctly)	0.80 (0.25)^s	0.25 (0.11)	0.68 (0.29)^s
Specificity (% correctly unlisted)	0.98 (0.04)^s	0.92 (0.03)	0.96 (0.03)^s
Direction of influence			
Percent correct	0.72 (0.18)	0.68 (0.17)	0.73 (.14)
Scenario generation			
Number of words	62.9 (30.1)	69.3 (34.6)	62.8 (26.3)
Number of variables	4.45 (2.74)	5.62 (3.07)	5.39 (2.97)
Evaluations of xenotransplantation			
Risk judgments (0–100%)	34.2 (16.1)	47.0 (16.6)^m	42.3 (18.5)
Policy judgments (−3 to 3)			
Before survey	0.35 (1.36)	0.86 (1.29)	0.22 (1.61)
After survey	0.33 (1.46)	0.50 (1.42)	0.10 (1.63)

m = significantly higher than model condition; s = significantly higher than scenarios condition.

3.2.2. Clarity of Materials

Internal consistency in responses to the four questions about the clarity of materials was $\alpha = 0.72$ at the beginning of the survey and $\alpha = 0.78$ at the end. We used each participant's mean rating at each time in subsequent analyses (Table II, Rows 2 and 3). There was no significant interaction between the materials provided (models, scenarios, or both) and the timing of the ratings (at the beginning vs. the end of the survey), $F(1, 145) = 0.50$, $p = 0.48$. There was a marginal difference in ratings of the materials, $F(2, 145) = 2.62$, $p = 0.08$, with the model rated as marginally less clear than the combination, Tukey 95% CI = (−0.40, 0.18), $p = 0.07$, but no significant difference between clarity ratings for the model and scenarios, Tukey 95% CI = (−0.29, 0.18), $p = 0.26$, or for the combined and scenario conditions, Tukey 95% CI = (−0.12, 0.18), $p = 0.80$. There was no significant difference between clarity ratings given at the beginning versus the end of the survey.

3.3. Depth of Understanding

3.3.1. Identifying Direct Causes and Effects

As reflected in the Kappa statistic, the two coders showed high agreement in deciding whether specific variables were mentioned in participants' open-ended responses ($\kappa = 0.97$). Based on these codes, we computed two test-diagnosticity scores for each participant.[44] *Sensitivity* reflects the proportion of cause and effect variables that were correctly identified, among those that should have been identified. *Specificity* reflects the proportion of correctly identified cause and effect variables, among those that the participant listed. Both scores avoid rewarding participants for simply listing many variables. The values of these scores range between 0 and 1, with higher values reflecting better performance.

Scores computed from the two judges' codes were highly correlated, with Pearson correlations being 0.85 for sensitivity, 0.74 for specificity. Cronbach's α for the test-diagnosticity scores computed across the 14 questions about causes and effects revealed high internal consistency ($\alpha = 0.94$ for sensitivity; $\alpha = 0.90$ for specificity). As a result, we computed mean sensitivity and specificity scores across questions (Table II, Rows 4 and 5). Two separate ANOVAs revealed a significant effect of the presented materials on both sensitivity, $F(2,146) = 77.61$, $p < 0.001$), and specificity, $F(2,146) = 40.09$, $p < 0.001$. As predicted, both scores were significantly better for those receiving the model than for those receiving the scenarios, Tukey 95% CI = (0.43, 0.65), $p < 0.001$ for sensitivity, Tukey 95% CI = (0.04, 0.07), $p < 0.001$ for specificity.

Scores for the combination condition fell in between. For sensitivity, scores were significantly better for the model than for the combination, Tukey

95% CI = (0.01, 0.23), $p < 0.05$, and for the combination than for the scenario, Tukey 95% CI = (−0.53, −0.32), $p < 0.001$. Specificity scores were better for the combination than for the scenario, Tukey 95% CI = (−0.06, −0.03), but scores for the combination and the model were similar, Tukey 95% CI = (0.00, 0.03), $p = 0.22$. Performance in the combination condition was significantly better than in the scenario condition, Tukey 95% CI = (−0.53, −0.32), $p < 0.001$ for sensitivity, Tukey 95% CI = (−0.06, −0.03), $p < 0.001$, for specificity.

3.3.2. Direction of Influences

Responses were scored as correct (0) or incorrect (1). Cronbach's alpha, which reduces to KR-20[45] for dichotomous items, was low ($\alpha = 0.13$). The overall score did not significantly differ between participants receiving the three sets of materials (see Table II), $F(2,146) = 1.30$, $p = 0.28$, nor did any individual item included in the overall score ($p > 0.05$). Overall, for about 70% of our questions, participants could discern the direction in which a variable's value was influenced by changes in the value of another variable.

3.3.3. Scenario Generation

The number of words in each scenario (Table II, Row 9) was derived with the "word count" function in Word. Two independent judges coded whether each model variable was present in each generated scenario, with the Kappa statistic showing their high agreement ($\kappa = 0.88$). There was also a high correlation ($r = 0.90$, $p < 0.001$) in the number of different model variables that the two judges identified in the scenarios. Contrary to our prediction, ANOVAs found no significant effect of the presented materials on the length of participants' scenarios, $F(2,146) = 0.76$, $p = 0.47$, or number of variables described, $F(2, 146) = 2.15$, $p = 0.12$. On average, participants produced scenarios with about 65 words, mentioning 5 of the 18 variables in the expert assessment (Table II, Row 10).

3.4. Evaluations of Xenotransplantation

3.4.1. Risk Judgments

The five probability judgments showed good internal consistency ($\alpha = 0.76$). Table II shows mean judgments (Row 6). An ANOVA revealed a significant effect of presented materials, $F(2,145) = 6.97$, $p < 0.01$. As predicted, scenario recipients saw much higher probabilities of problems than did model recipients, Tukey 95% CI = (−21.09, −4.62), $p < 0.001$. Probability judgments in the combination condition were in the middle, not significantly different from those in either the model condition, Tukey 95% CI = (−16.35, 0.04), $p = 0.05$, or the scenarios condition, Tukey 95% CI= (−3.37, 12.77), $p = 0.35$.

3.4.2. Policy Judgments

There was no significant main effect of presented materials on judgments about xenotransplantation across ratings provided before and after completing the survey, $F(2,145) = 1.72$, $p = 0.18$. Across conditions presenting different sets of materials, these ratings were significantly lower after answering the questions than before, $F(1,145) = 7.49$, $p < 0.01$. There was a marginal interaction between presented materials and before versus after ratings, $F(2,146) = 2.72$, $p = 0.07$, with the decline in ratings being significant for the scenarios condition $t(49) = 3.40$, $p < 0.01$ but not the others ($p > 0.10$).

4. DISCUSSION

Although models are often used to communicate to experts, they are uncommon in communications targeting lay people.[30] Here, we empirically compared the responses of educated lay people to three strategic alternatives for communicating the complex, uncertain risk of a xenotransplantation patient becoming infected with a recombinant virus: (1) a graphic model, (2) three structured scenarios, and (3) both model and scenarios. The model and scenarios presented the same variables and relationships, based on the expert assessment of Fig. 1.

We evaluated participants' responses to these communication materials in terms of (1) ease of processing, (2) depth of understanding, and (3) evaluation of xenotransplantation. Model recipients did at least as well as scenario recipients on every measure.

Ease of processing was measured by the time needed to answer questions about the materials and by ratings of their clarity. There were no significant differences, suggesting that recipients worked equally hard to understand the three kinds of material and felt equally (and moderately) positive about them. Thus, materials covering the same variables and relationships appeared equally accessible, in all three presentation modes.

Depth of understanding was measured by how well recipients could (a) identify the direct causes and effects of key variables, (b) understand the effects of changes in variables on contingent variables, and (c) produce scenarios describing xenotransplantation risks. Although these measures were chosen to capture expected strengths of the model (a, b) and scenarios (c), the performance of model and scenario recipients differed on only one measure. As predicted, model recipients could better identify the causes and effects of focal variables, presumably because the model showed these relations explicitly, whereas scenario recipients had to extract them.

Despite being better able to identify cause-effect relationships, model recipients were no better at identifying how a change in a focal variable influenced the value of another variable. Two possible methodological issues may obscure this test. One is that our questions may have inadvertently communicated the direction of influence, overriding whatever participants had taken from the materials (i.e., "How would an increase in the compatibility of a pig virus and a human virus affect the chances that they will recombine to make a new virus?"). The second is that performance on the seven items in this set showed low internal consistency, making it a poor measure for the underlying construct.

There were no significant differences between the model and scenarios conditions on a measure that we had expected to favor scenario recipients: generating a new scenario describing the risk of a xenotransplantation patient getting infected with a recombinant virus. Participants in the model and scenarios conditions used as many words and included as many variables.

As predicted, scenario recipients gave higher probability judgments for xenotransplantation patients getting infected with a recombinant virus, compared to model recipients. That occurred even though only one of the three scenarios described a xenotransplantation patient getting such an infection—and the fact that other studies have found that reading multiple scenarios with different endings tends to reduce judged probabilities.[41] Perhaps reading that concrete scenario made the risk easier to imagine, compared to receiving just the abstract model. However, because the actual probability of these events is unknown, the accuracy of these increased judgments cannot be evaluated.

After reading the materials, scenario and model recipients made similar policy judgments about the acceptability of xenotransplantation, while the scenario recipients saw it as marginally less acceptable. Thus, the scenarios seemed to have a stronger effect after participants thought about the risks in greater depth.

Overall, our results suggest a slight advantage of a graphic model over narrative scenarios. The model helped with identifying direct causes and effects, without requiring additional time, reducing the judged clarity of materials, or affecting the ability to produce risk scenarios. Without accepted scientific estimates of the presented risks, the accuracy of these risk judgments cannot be evaluated. However, model recipients' superior performance suggests that their lower risk judgments might mean that this technology seems more appealing the better it is understood.

Participants who received both the graphic model and the scenarios performed at least as well, in all respects, as those who received just the scenarios. Thus, the chance to see things from both perspectives outweighed any additional cognitive load.

4.1. Limitations and Implications

As detailed as they were, our materials did not provide a comprehensive picture of xenotransplantation. Rather, they focused on one risk central to policy making: a patient getting infected with a transmissible recombined virus, creating the possibility of a novel epidemic. Fully informed decisions will also need to consider other patient infection pathways, as well as the public health measures that could and would be taken to reduce an infection's spread. Our materials also did not address ethical concerns, like those raised above.[9,10] We believe that the methods used here to create and evaluate communications could be applied to communicating these issues as well. Although we would predict similar results regarding risk-related topics, ethical ones may be different. Speculatively, because of their vividness, scenarios may evoke deeper ethical concerns, whereas models may suppress them.

Participants in this study were primarily Carnegie Mellon University students, who are younger, better educated, and more numerate than the general public. In terms of *what* people think, those who are less familiar with new technologies might find xenotransplantation less acceptable. In terms of *how* people think, we expect less-well-educated individuals to show lower performance overall. Less numerate individuals, who also tend to be less well educated, have been found to trust verbal information more than numeric information, with the opposite being true for more numerate individuals.[46] Whether that result would hold for graphic models, which are abstract but not numeric, is a matter for further research, as is the question of whether the combined condition would provide more of a cognitive burden or greater opportunities to find information in a useful format.

The public's judgment of xenotransplantation, and other innovative technologies, will depend on the information that it receives. We offer a method for creating authoritative communications designed to complement recipients' mental models for such complex, open systems, as well as for evaluating their success. Further research is needed to assess the ability to generalize this result to other audiences and technologies. In its absence, the best strategy may be to provide both formats, hoping that individuals will choose what works best for them.

ACKNOWLEDGMENTS

The research reported here was supported by National Science Foundation Grants SBR-95-21914, SES 0433152, and SES 0350493. We thank our expert panel, Alwynelle Ahl, Margaret Cary, Linda Hogle, Sally Kane, Kim Waddell, and Steven Weber, as well as our research assistants Charlotte Fitzgerald and John Wylie, for their contributions to the project. The views expressed are those of the authors.

REFERENCES

1. Yang YG, Sykes M. Xenotransplantation: Current status and a perspective on the future. Nature Immunology, 2007; 7:519–531.
2. Sim KH, Marinov A, Levy GA. Xenotransplantation: A potential solution to the critical organ donor shortage. Canadian Journal of Gastroenterology, 1999:13: 311–318.
3. Cascalho M, Platt JL. The future of organ replacement: Needs, potential applications, and obstacles to application. Transplantation, P 2006; 38:362–364.
4. Elliott RB, Escobar L, Tan PL, Muzina M, Zwain S, Buchanan C. Live encapsulated porcine islets from a type 1 diabetic patient 9.5 yr after xenotransplantation. Xenotransplantation, 2007; 14:157–161.
5. Boneva RS, Folks TM. Xenotransplantation and risks of zoonotic infections. Annals of Medicine, 2004; 36:504–517.
6. Blusch JH, Patience C, Martin, U. Pig endogenous retroviruses and xenotransplantation. Xenotransplantation, 2002; 9:242–251.
7. Weiss RA. Transgenic pigs and virus adaptation. Nature, 1998; 39:327–328.
8. Caruso D. Intervention. Confronting the Real Risks of Genetic Engineering and Life on a Biotech Planet. San Francisco, CA, Hybrid Vigor Press, 2006.
9. Anderson M. Xenotransplantation: A bioethical evaluation. Journal of Medical Ethics, 2006; 32:205–208.
10. Melo H, Brandâo C, Rego C, Nunes R. Ethical and legal issues in xenotransplantation. Bioethics, 2001; 15:427–442.
11. Einsiedel EF, Ross H. Animal spare parts? A Canadian public consultation on xenotransplantation. Science and Engineering Ethics, 2002; 8:579–591.
12. Julvez J, Tuppin P, Cohen S. Survey in France of response to xenotransplantation. Lancet, 1999; 353:726.
13. Lundin S, Idvall M. Attitudes of Swedes to marginal donors and xenotransplantation. Journal of Medical Ethics, 2006; 29:186–192.
14. Pearce CE, Thomas APM, Clements DAV. The ethics of xenotransplantation: A survey of student attitudes. Xenotransplantation 2006; 13:253–257.
15. Persson MO, Persson NH, Ranstam J, Hermerén G. Attitudes toward xenotransplantation—Patients waiting for transplantation versus the general public. Transplant International, 2001; 14:334–342.
16. Hagelin J. Public opinion surveys about xenotransplantation. Xenotransplantation, 2004; 11:551–558.
17. Wright Jr JR. Alternative interpretations of the same data: Flaws in the process of consulting the Canadian public about xenotransplantation issues. Canadian Medical Association Journal, 2002; 167:40–42.
18. Wright Jr JR. Public consultation on xenotransplantation. Transplantation, 2004; 78:1112–1113.
19. Palmgren C, Morgan MG, Bruine de Bruin W, Keith D. Initial public perceptions of deep geological and oceanic disposal of carbon dioxide. Environmental Science and Technology, 2004; 38:6441–6450.
20. Maharik M, Fischhoff B, Morgan MG. Risk knowledge and risk attitudes regarding nuclear energy sources in space. Risk Analysis, 1993; 13:345–353.
21. Morgan MG, Fischhoff B, Bostrom A, Atman CJ. Risk Communication: A Mental Models Approach. Cambridge, MA: Cambridge University Press, 2002.
22. Gentner D, Stevens AL (eds). Mental Models. Hillsdale, NJ: Erlbaum, 1983.
23. Johnson-Laird PN. Mental Models. Cambridge, MA: Harvard University Press, 1983.
24. Rouse WB, Morris NM. On looking into the black box: Prospects and limits in the search for mental models. Psychology Bulletin, 1986; 1100:349–363.
25. Newell A, Simon HA. Human Problem Solving. Englewood Cliffs, NJ: Prentice-Hall, 1972.
26. Bostrom A, Fischhoff B, Morgan MG. Characterizing mental models of hazardous processes: A methodology and an application to radon. Journal of Social Issues, 1992; 48:85–100.
27. Byram S, Fischhoff B, Embrey M, de Bruin WB, Thorne S. Mental models of women with breast implants: Local complications. Behavioral Medicine, 2001; 27(1):4–14.
28. Downs JS, Bruine de Bruin W, Fischhoff B. Parents' vaccination comprehension and decisions. Vaccine, 2008; 12:1595–1607.
29. Bruine de Bruin W, Fischhoff B, Brilliant L, Caruso D. Expert judgments of pandemic influenza risks. Global Public Health, 1:1–16.
30. Fischhoff B, Bruine de Bruin W, Güvenç U, Brilliant L, Caruso D. Analyzing disaster risks and plans: An Avian Flu example. Journal of Risk Uncertainty, 2006; 33:131–149.
31. Clemen R. Making Hard Decisions. Belmont, CA: Duxbury, 1997.
32. Fischhoff B. Scientific management of science? Policy Sciences, 2000; 33:73–87.
33. Bruine de Bruin W, Downs JS, Fischhoff, B. Adolescents' thinking about the risks and benefits of sexual behavior. In: Lovett MC, Shah P. (eds). Thinking with Data. Mahwah, NJ: Erlbaum, 2007; 421–439.
34. Horvitz E. (ed.). Special issue on graph-based representations. Decision Analysis, 2005:2.
35. Howard RA, Matheson JE. Influence diagram retrospective. Decision Analysis, 2005; 2:144–147.

36. Lipkus IM, Hollands JG. The visual communication of risk. Journal of the National Cancer Institute, 1999; 25:149–163.
37. Larkin J, Simon H. Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 1987; 11:65–99.
38. Chase W, Simon H. Perception in chess. Cognitive Psychology, 1973; 4:55–81.
39. Sweller J. Cognitive load during problem solving: Effects on learning. Cognitive Science, 1988; 12:257–285.
40. Carroll JS. The effect of imagining an event on expectations for the event: An interpretation in terms of the availability heuristic. Journal of Experimental Social Psychology, 1976; 14:88–96.
41. Schoemaker PJH. Multiple scenario development: Its conceptual and behavioral foundation. Strategic Management Journal, 1993; 14:193–213.
42. Güvenç U. A Scenario-Based Integrated Risk Assessment & Communication Method. Doctoral dissertation, Carnegie Mellon University, Department of Engineering & Public Policy, 2005.
43. Cronbach LJ. Coefficient alpha and the internal structure of the tests. Psychometrika, 1951; 16:297–334.
44. Altman DG, Bland JM. Diagnostics tests 1: Sensitivity and specificity. British Medical Journal, 1994; 308: 1552.
45. Kuder G, Richardson MW. The theory of the estimation of test reliability. Psychometrika, 1937; 2:151–160.
46. Gurmankin AD, Baron J, Armstrong K. The effect of numerical statements of risk on trust and comfort with hypothetical physician risk communication. Medicine of Decision Making, 2004; 24(3):265