



Presidential Commission
for the Study of Bioethical Issues

TRANSCRIPT

Carol Ripple, Ph.D.

Associate Director for Education Research and Engagement
Duke University Social Science Research Institute

Raymond De Vries, Ph.D.

Professor of Learning Health Sciences
Co-Director, Center for Bioethics and Social Sciences in Medicine
Professor of Sociology
Professor of Obstetrics and Gynecology
University of Michigan

John Gastil, Ph.D.

Head and Professor
Communication Arts and Sciences and Political Science
The Pennsylvania State University

Meeting 22, Session 4
September 2, 2015
Washington, DC

SESSION 4: FOSTERING AND MEASURING SUCCESS IN ETHICS AND DELIBERATION

DR. GUTMANN: Welcome. Our first panelist is Carol Ripple, Associate Director for Education Research and Engagement in the Social Science Research Institute at Duke University. Dr. Ripple leads the education and human development incubator and oversees evaluation, data infrastructure, and community engagement.

Prior to her time at Duke, Dr. Ripple was Assistant Professor in Yale School of Medicine's Department of Psychiatry, before joining the Annie E. Casey Foundation, where she designed and led evaluations of community-based and foster care programs.

Dr. Ripple has also led evaluations of all manner of state-funded programs as principal evaluator in the North Carolina General Assembly's Program Evaluation Division.

You've had a far-ranging career to date, and we welcome you.

DR. RIPPLE: Thank you. I am really delighted to be here, and particularly excited to talk to you about using evaluation as a tool, particularly for informing ethics education initiatives. You are going to hear more about deliberation, I think, from my esteemed colleagues in a moment.

But the goal here today is really to think about adopting evaluation thinking as we are formulating our education programs, and as we are implementing them. So considering evaluation topics right from the beginning.

So today we will talk a little bit about just the background of what we are talking about when we speak of evaluation, and then we are going to jump pretty quickly into an example taken directly from Nita Farahany's Science and Society

Initiative at Duke.

So evaluation is a particular kind of research, and it has two very broad purposes. One is to inform program development. It can really help give us feedback about what we're doing well and maybe to help us tweak our goals and make sure that we're meeting them. And it can also, then, provide accountability to our stakeholders, to ourselves, to see if we are meeting our goals and to our constituencies, our funders, our participants. And it may seem very obvious to say this, but evaluation requires planning.

Now, researchers are thinking about research all the time. But when we are implementing education programs, we tend to be a little more focused on what we are going to teach, what is the content, how are we going to implement, how are we going to attract our leaders, and evaluation and measurement may be a very distant thought at those beginning steps.

I am reminded of an experience I had with yet another initiative at Duke where this initiative wanted to track whether their participants were going on and being employed in a related field, which spawned a conversation about what that meant. That was actually quite a complemented construct, what that -- what employment in the field actually meant.

But what it came down to is, well, we need to contact our graduates and see what they were up to, and I got blank looks and they said, "Well, how?" Because this thought of how to follow on hadn't been discussed, so there were no emails. Now, it didn't mean that they couldn't get them, but it made it far more complicated. So it does take forethought.

And it particularly takes forethought based on what we want to know

about. Formulating evaluation questions very early is essential just as formulating a research hypothesis helps us understand where our research is directed, what design we need to acquire the correct data, and how we are going to analyze it. So evaluation questions serve exactly that purpose and are really essential.

So another way to think about evaluation is that it is a systematic method for collecting, analyzing, and using data to answer questions about projects, policies, programs, with a particular emphasis on looking at their effectiveness, their efficiency, and development, as we have just talked about. So this is quite broad.

Another way to think about evaluation is it's a form of reflective practice, and it gives us a lens through which we can examine what we are up to and really inform us about how it is working or not, and let the wider public know about our success at doing what we're up to.

There are many, many different purposes for evaluation. Today we are really going to hone in on outcome evaluation, and the central question there is, so what? To what end are we implementing, in this case, ethics education programs? Another really pertinent question is: to what should our program be accountable?

Now, we may have, actually pertinent to this group, the end goal of world peace, but evaluating whether we are achieving that is not really going to help us know what our exact intervention, our exact education approach is actually accomplishing. So, again, having the right question to determine what those outcomes are is essential.

Good outcomes, then, are reasonable. That is, they are associated both with the program intent as well as content. They are actually pertinent to answering our evaluation questions. They are measureable. There is some metric there, whether it's qualitative or quantitative, that relates to informing our question, and tend to be both

short term, very proximal to our education initiative, or -- as well as very long term, as we are getting to this question about, so what?

So let's jump into our example. This is taken, as I mentioned, from Nita Farahany's Science and Society Initiative, which was founded back in 2013 with the goal of examining the integral role of science and law policy, social institutions, and culture.

More recently, just this past July, Science and Society was adopted as the newest of Duke's four campus-wide initiatives. So with this relaunch and focus, it became really important to try to focus the initiative and to really determine where it was going and what its specific goals were -- are.

So I started to work with Nita and her colleagues to frame a tight mission statement for the -- this is for the overall initiative as well as for goal statements. So the mission here is quite broad again. This is for the initiative as a whole, to maximize social benefit from scientific progress by making science more accessible, just, and better integrated into society.

So today we are going to focus on the last goal on this slide, which pertains to improving science communication in order to enhance public understanding of science.

This is a very packed slide. Don't panic. We're going to talk through it, and then we're going to unpack it. At the top here, though, you see the goal that we are interested in. right? So science communication in order to enhance public understanding of science. We can already tell from that goal we need to look long term. We need to look down the road to answer that "so what" question.

So directly below that title we see a range of outcome domains that we

might consider measuring, but I want to start actually at the bottom of the slide. These are the four domains of activities that the science communication part of the initiative is focusing on. So there is training for post-docs and faculty. There are courses offered at the university. There is leadership development as well as workshops for graduate students, and then there are specific activities within them.

But before we get to looking at those long-term outcomes, we need to consider more proximal data: things like, are people showing up? Are they actually getting the dose of education that is intended? Are they coming to workshops? Very short-term, proximal measures of, did they learn what they were meant to right at the end of our education program? Whether it's a workshop or a class. Were they satisfied? Did they get what they thought they were going to do? Is it something that they are going to talk to their friends about?

All of those are really essential in and of themselves, and so depending on those evaluation questions this could be enough, but I know Nita, and I know that's not going to be enough, particularly if we are looking at fulfilling this goal of enhancing public understanding. So we have to look at the longer term outcomes.

Let's step back and take a little closer look at what some of those more proximal short-term data are. These can be very basic. How many learners are showing up or participating? What are their characteristics? We may be interested in particular kinds of demographics that we need to know about. Do they persevere? Are they satisfied? Or we might look at grades to see what their performance is, or have a metric that is really closer to what our content is for the particular education platform.

Data collection for the short-term data tends to be pretty straightforward. It is usually collected before, during, or immediately after our education activity. We

can set up participant surveys to measure before and after learning. We can gather program administrative data that tells us who they are and what they have been up to. Again, though, we can see here, even for these more basic data, there needs to be forethought in order to set up the data systems that are required to have this data on hand.

This is even more true when we get to these longer term outcomes. Those outcome domains are listed here on the left. Are science communication skills increasing? And do they stay higher over time? Are learners actually adopting some of these skills that they learned? And are they talking to other people about them?

Is the broader awareness of the importance of science communication growing? And are communities forming around the importance of science communication?

So on the middle of this slide we can see that there are various ways that we might think about measuring this persistent behavioral change among learners. We might conduct follow-up surveys with them. We might interview them. We might observe them and code their behavior. In looking at dissemination of program effects, we can look at examples here of mentorship in a more qualitative setting. We can observe teaching. We can quantify collaborations, and we can also quantify publications.

When we want to look at awareness of science communication, we can see if our applications and enrollment is up for the program. We can count the number of inquiries, the number of website hits. Maybe tracking social media and seeing what social media chatter is around science communication.

When we look more systemically, we can count the kinds of networks that

are forming around science communication. We can even take account of the conferences and funding opportunities that are there that actually focus on science communication or that actually incorporate them in the call. So there are lots of different ways, again, depending on what we really need to know about from our education programs. And Nita and I are in the process of having these conversations.

So some of these more nuts and bolts aspects here under data collection are the tools that we will need to have at hand, and there are a bevy of tools, depending, again, on these questions, if we want a comparison group to know that -- if our learners are learning more than, say, those who aren't participating. We would need to assemble that comparison group.

We have been talking a little bit about feasibility of evaluation and of assessment today. Some of the feasibility comes into play when we are thinking about how complicated our questions are and how much we need to know and about what.

So evaluation, then, is a tool that really helps us inform program development and tweak what we are doing in education and ethics education. It relies on carefully selected questions. And we have talked about one example here in science communication. The tenets of what we have talked about in evaluation apply to a range of -- whether it's ethics education and beyond.

For example, we might think about the education modules that are on your own website. Depending on the question, you might want to know about, for example, who is engaging, who is accessing those modules? Do they actually -- do your learners complete the series? Where are they tuning out? Where are they tuning in? What are they learning about?

Each of these questions, then, comes with a set of variables that you would

need to collect. So might you have a very brief demographic survey at the very beginning to try to get a sense of who is accessing them. You can do -- you can examine clickstream data that will tell you when they are tuning in and when they are tuning out. And it's not necessarily a bad thing if they're not watching the whole thing, but you would want to know, are they coming in for particular pieces of information?

You could have a few questions at the end of your modules, just to assess satisfaction and to assess basic learning. So, again, depending on the question that we have at hand, evaluation can be used as a very specific tool to help us understand more about what our ethics education programs might be changed to accomplish and what they could actually tell our stakeholders about.

Thank you.

DR. GUTMANN: Thank you very much.

We now welcome Raymond De Vries, who is Professor and Co-Director of the Center for Bioethics and Social Science in Medicine at the University of Michigan Medical School. Dr. De Vries is an expert in the use of democratic deliberation for soliciting informed public opinion in the area of bioethics.

He and his colleagues have developed criteria for measuring the quality of democratic deliberation. With his research team, Dr. De Vries has used this technique to solicit informed public opinion on the ethics of surrogate consent for Alzheimer's and dementia research, the return of incidental findings in the practice of precision medicine, and policies for consent for buy-out bank donations. A great segue from measuring to measuring something very specific to our previous deliberations.

Thank you very much for joining us.

DR. De VRIES: You're welcome. Thanks for inviting me. It is a

pleasure to be here with colleagues who have played such an important role in the theory and practice of deliberation, so I really look forward to our conversation after we are done here.

We are here to talk about what makes for good deliberation. And we have already heard this morning enough about this, but here is an article from about eighteen years ago talking about the wonders of deliberation and providing more legitimate public spirited, mutually respectful, and self-correcting decisions on the part of participants. I think we are all agreed on that.

Yes, but what makes for good public deliberation? And that's what I am here to talk to you about today. I want to start by going over an excerpt from one of our transcripts of our deliberation about the use of surrogate consent for enrolling people with dementia in clinical trials. And I just want to give you a feel for what we are looking at. I'm going to talk about quantitative and qualitative ways of assessing the quality of deliberation here today.

So here is two participants. One is a woman, F, and a man, M. And here is basically the back and forth:

“I can understand why you support the use of surrogate consent, but I think when you get to individual scenarios you have to really stop and think: is there going to be a scenario where the risk is so high we would not, as a society, ever want to have a surrogate make a decision? Now, let me finish. I can see you are trying to answer me. But really think about this. Are we ready, as a society, to say, “okay, it's okay to have a surrogate consent for a relatively high-risk scenario”? You might be willing to push yourself in a higher risk scenario than you might be as a surrogate.”

And then, her interlocutor responds:

“This isn't an argument. This is an exercise from my standpoint.”

Okay.

“If you and I were married, and I was a patient and there was a low risk in research, do you feel that society should let you make that decision if I couldn't?”

And she responds:

“Yeah. Yeah.”

He goes on:

“Let's take it up another notch. I'm the Alzheimer's patient, I can't make a decision on my own. There is a high risk involved in research. Should society allow you, my wife, to make a decision regarding my participation?”

She responds:

“See, that's where it gets fuzzy for me. Because from a societal point of view, it would have to be that proper risk-reward ratio, and I think that's up for societal debate.”

So thinking about some of the quality characteristics of deliberation, if we look at this, just point out a few things. First of all, we see these people disagreeing but doing it in a very civil way. It's like, ‘wait a minute, wait a minute -- I see you are trying to barge in here.’ And his response is also – ‘this isn't an argument. We are engaged in an exercise here.’

You also see that people are taking a societal point of view, which is really a key part of deliberation. Not just what I want, but what works for everyone in society. And you'll see that again here in this follow up: ‘Should society allow you, my wife,’

‘from a societal point of view?’

So here are some things I want us to think about when we think about what makes for quality deliberation. So this is some of the work I have done. I have sent some of these things to you all. I know you don't all have time to read them, but this is what I'm basing my talk on today.

And I have to acknowledge my colleagues. Here is a list of the people who have helped me in this research, and here is the people who have helped fund this research.

So assessing quality. These are the three features I want us to think about -- the structure of the deliberation, the process, and the outcomes. Although I would really like to flip this, and, oddly enough, I want to talk about outcomes first. So talk about outcomes, talk about structure, talk about process.

So outcomes. Does deliberation do anything? That's an essential first question, I think. And I just want to give you a few results from the studies we have done to show you that, yeah, indeed, deliberation does make a difference for people. And I'm mostly talking about this project on surrogate consent for enrolling people with dementia in clinical trials.

So what we did is we brought people together in deliberation. I won't -- there is a lot of detail about how we did that. We can talk about that later. But we asked people before they came to deliberate how they would feel about allowing surrogate consent for a demented person to be involved in a study involving a lumbar puncture, a new drug randomized clinical trial, and there's two more. And if you look at the first survey, which is pre-deliberation, you find in the lumbar puncture thirty-three percent said, "Yeah, I think they would be okay." After deliberation, that number

increased remarkably to seventy-six percent.

And then, we surveyed the participants a month later to look for decay of change. We found still there was a higher number, but it did decay some. The same with a trial involving a new drug; thirty-eight percent before deliberation, seventy-six after. And then some decay, but still a higher number than baseline.

Slightly more risky studies, a vaccine study -- once you vaccinate somebody, you can't really take it out of their body, so you see less at baseline, but, again, a big change after deliberation, and some decay.

And then, the most challenging study, a gene transfer study, we described this study which involves drilling a hole in the brain and placing genetic material into the brain. Seventeen percent say, "definitely allow surrogate consent," afterwards forty-one, and then some decay.

More recently, right in the middle of this study, we did a study on public opinion about the return of -- I call them incidental findings, but now I know they're secondary findings, excuse me -- and we had one, one of the deliberations was, "how do you feel about a policy where we would withhold findings about adult onset diseases in children who are sequenced?"

And you can see, not surprising for an American audience, eleven percent said, "What? You're not going to tell me? No, I don't like that policy." After deliberation, we saw a remarkable change where forty-three percent said, "Yeah. I can see the value of that policy."

So my point is -- those are just examples -- deliberation does something. People change their minds after they deliberate. But is this change of mind based on the informed and considered opinion of the appropriate public? Now we are talking about

quality.

To do that, I want to talk about structure, and that's the "who" and the "what." "Who" is at the table? And we have already heard a little bit about that. Getting a representative sample of people to deliberate, appropriate to the topic you are deliberating on, and "what" is the information that they are provided. So, information -- putting the 'informed' in 'informed and considered opinions.'

These are the characteristics that we would like to see in people when they deliberate. Do they learn new information? Do they understand and apply? Are they using correct information? Are they taking advantage? We, in particular, use onsite experts in our deliberative process. And what is the impact of information? Has the information actually had any influence on their opinions?

And we can measure this quantitatively, and here is from our surrogate-based research deliberation. We ask people before the deliberation a seventeen-item knowledge questionnaire. We have provided them with some information about SBR. And you can see after deliberation their scores on these seventeen items increased. So we have some quantitative evidence that there is some learning going on.

What about more qualitative measures? How are they using this information? So, again, we are looking at these same things, in this case using qualitative measures.

And just a few examples from our transcripts. We recorded all of the deliberations, and we spent a lot of time coding those transcripts, and I'll say a bit more about that in a minute. But here are just a few examples in the surrogate-based research project.

Here is somebody actually using information, talking about, ‘hey, folks, you know, we know that advanced directives don't really work for people about how they want to be treated at the end of life. We really know they are not going to work for somebody saying, "If I'm demented, you can enroll me in research."’

So clearly this person had the idea. A lot of our deliberators at first thought, why don't we just use advanced directives? But gradually they realized, after hearing from the experts, that advanced directives don't work in other areas as well.

Here is another example of somebody using a nice analogy, comparing donating an organ to using surrogate consent to enroll somebody in research. And what I'm trying to show here is you see the deliberators actually using reasoning and not just expressing preordained opinion, or -- they are actually thinking through these things. I apologize for the format. Of course, it's different here than it was on my computer.

This is from the study of return of incidental findings. And, remember, the change was people being more willing to allow results to be withheld on adult onset diseases. And this is kind of a good example. Just look at the first and last sentence. She is saying, "I kind of feel like with all this there are so many unknowns, and the degree of this or that is so unknown."

And then she concludes by saying, "I feel like it could be catastrophic to know these partial maybes. This could happen. You have X probability of developing this condition as an adult." And the deliberators really start to talk about the negative effects of that, which I don't think they had thought about before they came to the deliberation.

So here is impact of information on opinion. Okay. So we have the right people. They're informed, but is it considered opinion? And here we look at process,

and there are four aspects of the process that we have looked at. Are people participating equally? We have heard some about that this morning. Are they respecting the opinions of others? We have heard some about that. Are they adopting a societal perspective? And are they giving reasons for their ideas? So just a few examples of what we found and how we looked for these things.

You can actually quantify, although I would not recommend this to anybody, how much people talk in each group. And we, with the help of our research assistants, which I listed earlier, actually counted the number of words spoken, the number of turns taken in the deliberation. And you can see this isn't perfect. Ideally, it would all be about 15 percent. But you can see, too, that there is no particular case where somebody is just overwhelming in the conversation. So this is one way to measure equal participation. Very tedious, but also very accurate.

How about respect for the opinions of others? We did in our survey after the deliberation ask people, "did you feel respected?" "Did you feel you were listened to?" "Do you feel the process was fair?" And you can see our deliberators at least felt very good about the process that happened. And it would have been a different story if they said, "No, people, I wasn't listened to, I wasn't respected." But in our deliberation, this is what we found.

Here I thought, because it -- I'm making it sound like our research was just this brilliant ideal. We did find some examples where there was no respect, and we actually coded for uncivil interactions. But I have to say, we didn't find many, but this is one example. Not surprisingly, it deals with religion, where somebody said -- remember, we're talking about surrogate-based research. "If your son's death could result in savings millions of lives, wouldn't you love those people more than you

would love your son in order to give him up for others?" And the respondent says, "I don't think I'd be willing to offer my son or daughter to save millions or thousands."

And then this person said, "My God did that." And then sarcastically -- you can't hear the tone, but I have listened to the tape -- the sarcastic response, "Yeah. Well, I'm not God." So you had a little -- there was a moment of tension if you listened to the tape there.

So I don't want to say that, you know, all deliberations are these smooth, respectful interactions, but this does happen.

Adoption of a societal perspective -- I already gave you an example at the outset of my talk, but here is another good example: "It seems that would be a train wreck for our society if we don't allow surrogate consent. We almost have no choice but to have some form of surrogate consent."

Reasoned justification of ideas -- here is somebody using the data that we provided about the danger of the genetic transfer study, and saying, "Look, I kind of like this, but that's a pretty high risk. I'm going to agree with M14, the other participant. I don't think society should allow surrogates to make that decision."

See, what I want you to see is, listening to what people say gives you a sense of, are they using reasons? Are they having a societal perspective?

So, in the three minutes I have left: thinking about deliberations. So those are a broad sketch of what we think to be measures of quality of deliberation, but I still have some nagging questions. Some of you might be wondering if you looked at the stuff I sent, why on earth would I include this chapter from Dan and Jeremy's book on the value of sociology as a method in medical ethics?

It's because I think my discipline -- sociology -- as we were discussing

earlier the value of disciplines, has something to contribute to the way people think about ethical questions. And, in this case, the sociologist always wants to say, "Well, oh, yeah, says who?" So who thinks that this method is better? Why do they think this method is a good way to come to moral reasoning or to come to decisions about ethical policy?

And, of course, as a sociologist, we are always asking, what is the historical situation and social location of that "who"? How does this relate to deliberation? I think, and I -- I have been thinking about this since I was invited to speak here, this to me is a really interesting question: "Why, in the field of bioethics in the United States, have we more or less shifted from expert bioethics to more public bioethics?"

If you look at the history of the Presidential Commissions, from NBAC to the PCB -- trying to remember that acronym -- to PIPSCI, or however you pronounce what you are now -- I think, as a sociologist, it's interesting to see the shift in how this governmental body is thinking about the way we should resolve ethical problems. And I -- yeah, I just find that fascinating. I would love to have a couple weeks to sit down and ponder that and write about it.

And then the next question, especially related to deliberation is, whose opinion -- and even though I have been involved in these, and I publish about these, and I do think they are good ways to develop ethical policy, I still have the question, "is the deliberation delivering public opinion?" What do I mean by that?

This is the ideal, that there is a balance between expert opinion and public opinion, and that they inform each other. The hope is that experts can listen to the public, and we have already heard this morning about how members of the public bring

things to experts that they hadn't thought of, and we have seen that in our deliberations as well. I worry about this. Is expert opinion overwhelming public opinion?

So things I think are worth worrying about, as I close here. There is a problem of self-selection of deliberators. We try very hard. We use random sampling, address-based sampling, to get people to our deliberation. But, of course, the people who say, "Yeah, that sounds like fun," are not like the grumpy people who sit at home watching the Tigers on TV and saying, "I know what's true about America." So there is that self-selection problem, and we have to think about that.

I worry about bias in the information that is provided. We have experts. We try to be unbiased. But, of course, they are coming from a point of view about the -- for example, the value of surrogate-based research in dementia.

And, finally, I worry about how we process the results and how we work them into the policies we create. So these are -- even though I'm an advocate, even though I have done this work -- these are things I continue to worry about that I think -- things that are worth you all thinking about, too.

So thank you for your attention.

DR. GUTMANN: Thank you. You really packed a lot in.

DR. De VRIES: I talk quickly.

DR. GUTMANN: No. Very well. Very well illustrated, too.

Our final presenter today is John Gastil, who is the Head and Professor of Communication Arts and Sciences and Political Science at Penn State. Dr. Gastil specializes in political deliberation and group decision-making, and has demonstrated the ways that deeper values bias how -- bias how we learn about issues and form opinions.

We look forward to your remarks.

DR. GASTIL: Thank you for the invitation, and it is a real pleasure being here. I am excited to share some thoughts with the committee. And the folks that I will thank that have helped me with this presentation are really actually dozens of people that I have gotten to know over the last 25 years, practitioners of deliberative processes, fellow researchers, graduate students, undergrads, and so on.

There really is a large scholarly and professional community of people interested in these issues, and I'm drawing on them like crazy without citation.

As for who I am, I am a bit of an oddball intellectually and historically. Just two parts of me that kind of fit into this presentation to help you understand where I'm coming from. The first is: both of my parents ran for Congress. When my Mom ran in '92, I became her campaign manager and managed campaigns for a few years. I was also raised as a Quaker, reformed, and have always been fascinated by group process, and then now a communications scholar.

So I have this weird mix of this kind of macro-level political science, campaign self, and this much more groovy, kind of hippie-based Quakery self. And what you get when you combine those is a real passion for democratic innovation, but with an eye towards how small group processes can fit into larger institutional arrangements.

So with that personal background in mind, we are going to do four things, really, just two. But, first, I will talk a little bit about the objectives you seem to have for deliberation, and I think there is one that is implicit that hasn't come up. Second, and mostly, we are going to talk about the design of a small group process that can be effective internally. And then, third, institutionally, thinking about how you embed that

small group. And then, finally, just a couple words about evaluation, which my colleagues have already addressed quite well.

So let's get things rolling by talking about this objectives idea. I have heard -- and the question I asked earlier was addressing, you know, what are really your primary objectives? And, yes, we want it all, but we have to prioritize. And here is how I come out of this discussion, having heard everything so far.

First, your primary goal is more reflective ethical and policy judgments made by the public writ large. So you want a more ethical and reflective public. Got it.

Second, you want more ethical public policy and more ethical decision-making in society and institutions generally. So better decisions from an ethical standpoint.

Third, and I haven't heard this, but I think it's there, I think you want a broader commitment to policy implementation and ethical conduct. And I emphasize this because one of the real important findings about group process versus just making decisions either by executive or by some aggregation process of individual opinion, is when you do it in a group, you are more committed to the decision.

And part of what I think you want is more commitment to the follow-through, the painful part. So it's great to make a good decision, it's great to become more reflective, but do you implement? And I think making that more explicit will tweak what you do.

And then, finally, fourth, this has not been as implicit, but all those goals are being perceived as efficiently as possible, with an eye towards the opportunity costs of doing this versus anything else.

All right. So I'll be alluding back to those goals just a little bit, but let's

roll into group design, which is really what the heart of this talk is about. In the last meeting, the 21st meeting's testimony, Jim Fishkin talked about the conditions where people can really engage one another. And we talk about that a lot. We talked about it here today. But that word "conditions" has a lot packed into it. What are those conditions?

And, obviously, we won't go into infinite detail today, but I'm going to talk a little bit about what does make a group an effective setting for deliberation. And I'll start just by acknowledging that a lot of people are afraid of groups. They freak people out. And we have lots of great jokes, things like a camel is a horse created by a committee, which I always thought was kind of strange. I mean, you're never going to see a Clydesdale in the Gobi Desert -- for a reason. So I like to think, you know, the groups kind of got it a little better. No disrespect to the horses.

(Laughter.)

But people worry about groups being -- and you've been in these groups. They are inefficient. They are polarized, in both ways -- all to one side, or off to two sides. You can see group think. You can see conformity, right? All these things. And we've seen them in experiments.

The good news is, when you look closely at those experiments, you find that the conditions that generate those worst outcomes usually aren't the ones you had in mind. And when we talk about group design, we are going to be talking about the very kinds of structural features that avoid those sorts of problems.

And, after all, if groups really were so fundamentally dysfunctional, they would not be as prevalent as they are. Any organization recruiting people from Penn State, for instance, wants people who are not just good writers and speakers. They want

people who are effective at teamwork, because it's done in teams.

Government itself -- I mean, you're sitting right here. This is teamwork, right? This is -- the whole point is to get a group together. Even when we involve citizens in government, the most power -- empowered way we do it is the jury system. And juries generally reach the same decisions as judges, what research has been done on that.

And, finally, we just live and act together in groups. The most important decisions and experiences you have had in your life were probably in groups. Sometimes with a best friend or a partner, but usually in groups. So they are there and they are tremendously important.

The next point I want to make before we get into design features is that there are many, many viable variations on these groups. There is the National Coalition for Dialogue and Deliberation, and the Deliberative Democracy Consortium, which are made up of members and members and members who have come up with all kinds of ingenious ways of structuring groups. They tend to have in common some of the features I am going to describe, but I am also going to describe some important variations.

There is a world of difference between a deliberative poll and a consensus conference, but they both can be highly functional in their own ways. I just want to emphasize that.

I will also bring in at this point one concrete example. I passed on a reading to you guys about this. It's something I have been studying very closely, and I think it shows you a few interesting and unusual features of groups embedded in institutions, and that is the Oregon Citizens' Initiative Review.

This process established by law in the State of Oregon a few years ago takes a random -- a representative cross-section sample of the State of Oregon, about 24 people, and they deliberate for a full week on one ballot measure that is going to appear on everybody's ballot.

At the end of that week, they have written a one-page statement which goes into the official voter's guide in the State of Oregon, so that every Oregonian, as busy as they are, can actually flip to that page and see what would a deliberative body of roughly my peers have to say about this issue. You know, what are the key findings? What are the pros and cons? And it doesn't take long to read.

And the research on this that we've done is very encouraging in terms of people finding out about it, being influenced by it, becoming more knowledgeable. So it's not just that the panel itself is highly deliberative. It actually does appear to be making the initiative process in Oregon more deliberative.

And now Colorado and Arizona have piloted this. Washington is going to do it soon. And Massachusetts is getting close to piloting one of these. So, again, it's just an example of an unusual institutional design that uses a small group to inform a more macro level, potentially deliberative process.

All right. So I teased you about structural group features, and I am only getting to it right now, so let's roll. First of all, the membership. You've got to populate your group diversely. So that's why we do all this random sampling and stratification, and so on. But I can't emphasize enough, if it's a smaller group, you really want to think about cross-cutting differences.

You know, the old idea that you don't want to sort of create a permanent minority. You want everybody to have a stake in different things in different ways. So

that if coalitions form in the group, they are constantly changing, they are fluid. So that's an important feature.

And when you're arranging a group of 24 people, you can actually pull that off. The citizens juries have been doing this since the 1970s, arranging groups actually with advocates on both sides of an issue sitting in the room while the group is formed, so to speak.

The number two dozen is what I tend to study. I have studied other bodies like the Australian Citizens Parliament, about 150 people; the British Columbia Citizens Assembly is comparable. Obviously, deliberative polls are much larger. Twenty-first century town meetings can be thousands of people, you know, connected remotely. But the reason two dozen is a good number is, you know, a 12-person jury gets a lot done, and they are very effective and cost effective. But two dozen is what you need, I think, for the more complex issues that we usually give people, whether it's ethical or policy.

But it is not so big that they can't meet as a coherent entity. And I actually do believe that the group dynamics can be a positive force in these cases, and so that's why that size comes about.

Finally, your membership needs to be kind of motivated and compensated. I strongly recommend compensation for these groups because it is hard work. And in this culture, money -- paying people -- actually tells them what they are going to do is important. As nice as it is that volunteered, it also gets you a more representative cross-section if you can give them some compensation along the way. And I don't mean \$10 a day for jury duty. Like in Oregon it's -- I think the median wage in the state is what you are paid, in addition to all expenses.

All right. How about the structure of the decision and the task? You need

a focused task. Some of these processes fail because it's too diffuse. Deliberative polls often have multiple topics over two and a half days. I suggest one topic, four days, is a really strong design.

The British Columbia Citizens' Assembly is actually, in that case, responsible for drafting a law. Then you need several weekends spread out over time. So you really have to get the task focused and the duration calibrated correctly to make sure that group has enough time but doesn't get bored.

The decision rule is something that is often overlooked. There have been a couple comments about it today, and here is what I suggest. This might seem like a paradox, but it's really not. You want a group like this to be consensus-oriented. You want them to be thinking, where is the common ground? What can we all agree on? Where do our minds meet? But you don't want a unanimity decision rule. You want a majoritarian or plurality rule.

But you really can combine those things, especially if the decision you are asking them to reach can be nuanced. It shouldn't be: do you check Box A or Box B? There should be text they are producing. And when they are producing text, there is more room for kind of compromised negotiation and getting the phrasing just right to bring as many on board as you can.

All right. How about the structure and content of the discussion? Background information is essential. You've heard that there is all sorts of things that are given to people in advance. But it's just as important to have expert witnesses coming into the room, and advocates from both sides who are available for intensive Q&A. I mean, intensive. I mean, they might be there for a couple of hours, because it takes a while to unpack things that they may have trouble expressing to a public

audience because they have never had that job.

In addition, we have heard about the importance of the facilitator. I would just stress, as I do in most of my writing, that there is two important dimensions here. When we say "democratic deliberation," democratic is the social process. That's the respect, the consideration. Extremely important for the facilitator to reemphasize those explicit ground rules.

But also there is an analytic rigor component. That is the deliberation in the sense that you must really weigh and think about tradeoffs. And some processes, like the Oregon CIR, actually have two facilitators. And, to an extent, they divide the labor a little bit. So think about those two dimensions when you are thinking about designing for facilitation.

There are other details about the roles that members can play, and so on. I will just emphasize one last thing here -- face to face. It's the reason that we have all bothered to be here. We talk about the power and scale of online deliberation, and I think that's going to come in some form, but the fact is you are embodied people here. You are motivated to attend to each other, and that really matters. It is part of the thing that makes you attentive and engaged, and there is an emotional component to it that keeps us going. These things can be exhausting, and deliberation -- you probably know this -- is incredibly boring to watch. It is very lively to participate in.

Institutional design -- just a few comments here. This is about how the group fits into the larger institutional context. I already gave you the example from Oregon. But as your chair said earlier, it's important that whatever these processes are that they really be capable, in principle, of impact. So that's impact on policy, on the public generally, and, again, on commitment to implementation.

What I encourage you to look for are institutional niches, or neeshes (phonetic) if you prefer. I don't have a preference. Think about the person holding a policy lever. Look for sweat. Sweat indicates either they are very anxious about the decision they are going to make or they are just tired. This is hard work. In either case, the public can bail them out.

And we have seen things like the Oregon legislature creating the CIR to help them with a very dicey initiative problem they have. They can't do away with it, but they can improve the process.

Other things like the British Colombia Citizens Assembly, you had a political party that knew it had to change the rules for the elections, but knew that no one would believe them if they told them they knew how to do it. It would sound biased. So they had a solution -- to hand it over to the people to write the law. Look for those opportunities.

And then, just a couple examples on opportunities you might miss. It is great if you can just create a process that creates decisions, say, internal to an institution or a government. That's fine. You can have indirect influences, like the Oregon CIR, but don't miss some interesting opportunities like the Base Realignment and Closure Commission, where a deliberative body can actually produce what is a policy that then floats past a legislative body.

In that case, Congress had only one choice. It was to kill it or just to let it go. And that's very different from Congress having to then approve it. So I'd just encourage you to think about those clever ways of arranging the role and power of the body.

Now, I said I'd say a couple comments about evaluation at the end, and I

shall. It is all good news. There have been tons of evaluations of deliberative processes, and on balance the results are pretty encouraging. Multiple designs are effective at organizing diverse publics who deliberate pretty carefully, and generally come up with high quality recommendations. Interested policy makers tend to very much respect the process and sometimes, as you have experienced, actually implement related policies.

Panelists themselves are overwhelmingly satisfied and are often changed in, maybe small ways, but in the long term. And the public itself tends to respect these processes. They love the jury, and actually they love the CIR in Oregon.

The public will attend to the results of these deliberative processes if they indeed are highly motivated. In the case of Oregon, there are voters who are about to vote. You may not have a comparable situation where the public truly cares about these processes, but think creatively.

Thank you again for the opportunity to share these thoughts with you, and I'll be happy to follow up on anything that caught your fancy.

DR. GUTMANN: Thank you very much. Perfectly timed.

We are open for questions. Let me begin, before I take questions, with an answer to a very interesting and I think relevant question, Raymond, you ended your discussion on, which is why -- the sociological question of why now rather than when commissions -- bioethics commissions began? Is there a real shift to an emphasis on public -- bringing the public into deliberation and public deliberation?

And I think, being a political scientist and reading both the sociological literature but also attending to the empirical literature about American politics, and this -- it may have begun in this country, but it is really around the world this has

happened. There are, I would say -- there may be many reasons, but there are three dominant ones.

One is that public trust in experts has diminished dramatically over the last three to five decades. I mean, dramatically, steadily, and it's at an all-time low point -- the public trust. In all kinds of experts, professionals, even churches, which once were very high in trust, leaders have gone down. So public trust dramatically down.

Secondly, public access to information is dramatically up. The public has access to information. You saw it in the medical field about public being able -- you just -- we all do it. You find out you have some condition, you Google it, you get information that two decades ago was far less accessible. So that's the second reason.

And the third reason -- and I just think this is important for our -- the context of our report. That's why I'm saying it. The third reason -- and there is great evidence for all this -- public education, legitimacy, and respect across diverse groups and individuals is -- all three of those things -- education, legitimacy, and respect, across individuals -- has become increasingly essential to a functional democracy. And deliberation furthers all three.

Deliberation furthers education. We heard our representative citizen say it was an education for me. Legitimacy, we heard everybody -- I mean, there is good evidence. And respect -- deliberation furthers all three of these better than non-deliberative alternatives.

It may not be perfect, but it is -- it isn't perfect, and we haven't yet found the perfect way of doing it that's thoroughly affordable, both with time and expense. But any move in a deliberative -- a decent deliberative direction furthers all of those

necessary parts of a functioning democracy better than non-deliberation.

So I think that is why we are -- if you will, we think it is so important to focus on deliberation. And I leave -- I want your reaction to that, and then I open it up for questions.

DR. De VRIES: I think that's -- I think that's a great sociological analysis.

DR. GUTMANN: Yeah.

DR. De VRIES: Which makes -- it just makes it so important, though, that when we are informing the public in these deliberative exercises that we don't present that information in a biased way. That is my concern, that, you know, all the things you mentioned in this area where there's a crisis of trust and legitimacy, I don't want deliberation to be a way to kind of come in the back door and get the public to see the wisdom of expert opinion, rather than giving the information they need and then really hearing their opinion in contrast to expert opinion.

I do have examples of that from our own work, and we heard some earlier, too.

DR. GUTMANN: Well, that's a question -- a follow-up question for you. Am I -- it would be biased if your questioning about surrogate consent assumed that that was -- you wanted surrogates' consent because you wanted surrogates to consent as opposed to whether you are asking people, do you believe this is justifiable process?

And I couldn't tell, in the examples you gave, whether it was bias towards we need more consent to these, and, therefore, you should agree to surrogacy, or were you actually -- did you present it as are you -- do you think it's right for there to be a surrogate, to decide whether to consent or not?

DR. De VRIES: And maybe that's -- this is a particularly tricky issue. To

me, that was -- in participating and not deliberation, that was a fuzzy line.

DR. GUTMANN: Yeah.

DR. De VRIES: So we had an expert from the Mayo Clinic in Alzheimer's research talking about the absolute necessity to have people with dementia enrolled in research. So that alone people are saying, "Oh, my goodness." And you saw some of those comments. We did have Paul Applebaum talk about the history of the abuses that have happened and the need for ethical safeguards.

DR. GUTMANN: Okay.

DR. De VRIES: So we did as best we could.

DR. GUTMANN: That's all you can do. You can only do -- I mean, it -- that's all I wanted to know.

DR. De VRIES: Yeah.

DR. GUTMANN: You really did the best you could to indicate the pros and cons. That's all you can do.

DR. De VRIES: But we did -- well, like as my colleague said, we didn't really have an advocate on the other side, somebody saying, "You know what? We can do this research in the laboratory. We could do it on animals." We didn't have that perspective.

I still think it was a fair deliberation, but these are the nagging questions, as I said.

DR. GUTMANN: Yeah. Although we heard earlier from our -- the science, you know, that you have to be careful not to think every question just has two equally balanced contradictory answers. That is -- that is just an intellectual mistake, right?

Okay. Nita Farahany is on the line. Would you like to ask a question, please?

DR. FARAHAANY: Sure. Can you hear me okay? Because we're having some sound problems.

DR. GUTMANN: We can. Yeah. It's okay.

DR. FARAHAANY: Great. Great. So thank you for this panel. I think it was really helpful. And one of the things that struck me about the panel and the conversations we are having today is the multi-faceted nature and multi-disciplinary perspectives of bioethics, and how you measure success in such a multi-faceted type of thing.

A lot of the public comments we got focused on this issue as well, which is the idea that this is a field and a broad field. So turning to education, and this is in part directed at Carol, since she really spent a lot of time thinking about program assessments, but also the others on the panel, how would you articulate and advance what you think -- or come to some kind of consensus as to what the measurement of success would be in such a multi-disciplinary faceted field.

Or is it simply individual education program by individual education program? If we're thinking about this on a broader policy level, how do we articulate the measurements for success? And how would you measure success in such great diversity?

DR. RIPPLE: So I'll start off. I think something that underlies -- and this was in the article that I circulated, that it really talked about a theory of change that underlies sort of the theoretical approach behind whether it's deliberation or education, and sort of the assumptions that we are making about what we expect to happen.

And another way to frame that is: what are our assumptions about, what are our values going into it? What are we thinking could feasibly happen as a result? And I think that there are very many different levels at which you can examine those assumptions, both from the individual program-by-program level, which doesn't address these higher level questions that you have, but there are also ways I think to glean sort of those higher-level findings from across different approaches, whether it's, again, education or deliberation, that really can get at what you think is reasonable to expect and what your assumptions about human behavior are and what your assumptions about your values are.

And, again, I -- bioethics is not my particular field, but I think this takes a lot of very careful thought about unpacking what some of those assumptions are about the behavior change. I think this is fascinating, to think that this is an earlier child development framework. It's a very different set of assumptions about what you might expect to change and how you want to then think in a developmental context in particularly ethics education from pre-K right up to post-secondary school.

So I think it's still important to keep a very sensitive contextual aspect to that learning, but it is indeed, you know, possible and a good idea to have that larger conversation.

DR. GUTMANN: Well, you had some specific to education, which you developed presumably with Nita on her project. On deliberation, which I've thought about and, you know, practiced, there are -- we have heard from experts in this that there are three -- there are more than three, but I would highlight three just for this sake, really measurable results.

One that you can measure, do people feel more respected? They can say

whether they do or not through -- that, you know, these are uncontroversial issues.

Two, do they understand the issues better? You can do whether they subjectively feel like that, but you can also test -- test that.

And the third one, which is tested in some of these is: do people change their minds in reaction to new information and argument? So there are -- I mean, for education, it's the understanding issues better that is the direct what -- sort of the direct salient outcome that you want from educational programs. And you can break that down.

So we do have -- and there are ways of measuring that. Raymond?

DR. De VRIES: I think John wanted to say something.

DR. GUTMANN: John?

DR. GASTIL: Yeah. I just want to jump in and say the criteria we used for the Oregon Citizens' Initiative Review, we did a report card for the legislature in Oregon. And it really boils down to three things, and you'll see some overlap but a little difference.

First, process integrity. Right? You can do that by both survey and observation. Did it meet the principles of democratic deliberation --

DR. GUTMANN: Sure, sure.

DR. GASTIL: -- internally? But actually I think that is secondary to the quality of the judgment. What did they produce? The one-page statement. Is it factually accurate? Is it coherent? Is it useful to voters?

And as the reviews has been going on now, they happen every two years, they have been talking more and more about, well, when I read the one two years ago, I thought that wasn't -- you know, this has to be useful to the average voter. So they are

really becoming more self-aware about how to craft that.

And then, third, educational impact on the wider public. So now, in the last survey we did, more than a majority of Oregonians are actually aware of this process, and most of those are reading the statement. And they are actually showing knowledge gains.

So that is similar but different. Because the small group process is so focused on its larger impact, you have to have measures of that larger impact.

DR. GUTMANN: Yeah. If that's the focus of --

DR. GASTIL: Exactly.

DR. GUTMANN: -- but we shouldn't forget the impact on the people. And the one thing I would say would be a huge mistake to leave out of any of this, because it is the primary value of deliberation, is, "Do people feel more respected?" It's a value that is embedded in democracy. It's in, you know, taking in -- in every competing, you know, reigning philosophy, whether it be, you know, virtue, character, consequential, or deontological. And deliberation -- that is really important.

DR. GASTIL: Right. And there is a non-obvious side to that when you go from micro to macro. So, for instance, we found that Oregonians, as they become aware of the CIR -- we have some longitudinal data on this -- are actually becoming a little more -- feeling a little more respected by their government.

DR. GUTMANN: Yeah.

DR. GASTIL: External efficacy. And they actually feel a little differently about themselves when they read the statement. So you can -- we call these emanating effects. You can even have some of that, on a thinner slice, at that macro level.

DR. GUTMANN: Good. Good. Dan?

DR. SULMASY: Perhaps, turn the question on its head and ask what you think is the most important item or series of qualities in deliberation or in education that you can't measure.

DR. GASTIL: I would say it's equality of opportunity to participate. We can measure how often someone speaks -- and actually, it's not a problem that everyone doesn't speak equally. But to know that you have the same opportunity as you? You may actually need twice as long to get your point across as she did. It's just impossible to know. So all we can do is ask you. We can just ask you. But that's not the same thing.

DR. DE VRIES: And it's also -- it's also -- I would say it's also hard to know what people bring to the deliberation. And you get some of that. I mean, I was going to put a pitch in for mixed methods, because we can see quantitatively, people change their minds. But then, qualitatively, we can see the kind of reasoning they're using. And I'd make a -- just an Amen to my colleague's statement about the face to face interaction being so important, and we see that in the transcripts.

But we don't always know what people are bringing. You do get to see some of that. So what past experiences, what past opinions do they bring? I suppose you could -- I'm trained in the school that everything is measurable. Although I'm a little critical of that point of view. But I suppose you could measure those things, also, on a survey. We don't -- but you don't know what things would be relevant that people are bringing to the deliberation.

DR. GUTMANN: There's a question that is relevant, that I didn't have a chance to read earlier. So let me read it. Dr. Nicola Jackson? Is she -- great. Who is -- you'll see why this is -- it's DPhil Quality Assurance Agency UK.

“As a member of the quality assurance agency panel revising the UK benchmark statements for bioscience degrees in the UK, we have paid attention to the ethical bases of UK bioscience degrees. Could this be a useful focus for the US presidential commission to strengthen bioethical education?”

And I thought it was important -- Dan, we had a sidebar conversation after the session that provoked that question for you. And Dan Sulmasy had a very good summary of what we thought would be important focus for bioethical -- what should we be thinking that bioethical degrees certify people in? So I'm going to have Dan answer it, and that obviously raises the question of, how do you measure it? But we don't -- but first you have to know what you're trying to measure. So Dan, if you could say a few words about that?

DR. SULMASY: Yeah, I guess building on the idea that probably we have close to a consensus that bioethics is not, at least at this point, a discipline. And a principle that we have of regulatory parsimony, that we have enunciated before. That we were hesitant to think about accrediting of bioethics programs. But we did think it would be important to call for more uniformity, instead of self-regulation, about that.

And maybe we can make some comments, but then the field itself would also have to do some work on its own quality standards for education. And then we sort of also had a little bit of a discussion that one of the things about taking bioethics not to be a discipline but a field -- and Ray recognizes from the book -- is this sort of sense that it is an opportunity for -- particularly for higher education, which often talks the talk of inter-disciplinarity -- to walk the walk of it.

Because bioethics presents this great opportunity for sociologists and lawyers and philosophers and theologians to all talk together about these -- about the --

DR. GUTMANN: And scientists.

DR. SULMASY: -- these questions. And scientists, I'm sorry, too.

DR. GUTMANN: To really -- and doctors. No, so affirmatively, we think it's a positive that bioethics is a field, not a single discipline. It would be narrowed artificially not to have a legal and sociological and philosophical perspective, and biology and neuroscience. But it still needs a set of criteria for what students should learn, what should be taught, and what measurable outcomes there are in understanding and appreciating. And I -- by measurable -- I don't mean necessarily measurable by multiple choice tests.

It could be -- there are different ways of measuring. And so this has been extremely helpful. If we had -- we will have more time to continue this in another way, but I want to just end this part of our session by giving enormous thanks. You've done a great job, to John, Ray, and Carol.

Really, thank you very, very much.

We're going to break for five minutes and reassemble for our round -- oh, five minutes, reassemble for a round table.

DR. MICHAEL: [inaudible]

DR. GUTMANN: And Nelson, who I had on my list, but we ran out of time -- which is my fault -- Nelson will begin with a question in the round table. Thank you all very much. Five minutes.

(Pause.)