Questionnaire Design

Arthur Lupia: This session is about visuals. So on the panel study, which is obviously Internet-based, you have the opportunity to present respondents with visual stimuli during the interview. But, of course, during the face-to-face we talked this morning about turning the laptop around for self-administration. So you have not only the video option, but when have the respondent self-administering *[inaudible]*— do you have presentation? ***(I've got two.)*** Jon has a presentation that will begin this conversation.

Jon A. Krosnick: Okay, we're going to show you some examples of things that we're considering doing with the panel respondents and with the laptop turned around to the face-to-face respondents. And this is to stimulate you to think about what could we do visually using this paradigm that might be interesting.

The first thing that I want to show is Keith Payne's affect misattribution paradigm. If you think this is obscure social psychology, I will tell you that the Federal Register a month ago posted a proposed study to be done by the Food and Drug Administration using this technique to find out how direct consumer TV ads for drugs affect people's attitudes.

And so the idea here is to measure attitudes without actually asking people what they are — on the grounds that you might not believe what they say. And so the paradigm goes like this. I bring you into the lab, I sit you in front of a computer — or I put you in your living room in front of my computer — and I tell you, you're going to see a series of flashes on the screen that first you're going to see what we'll call a "real-life picture" very briefly, and then you're going to see a Chinese pictograph.

So this picture appears, and then this one takes over for it. Your job is to judge how pleasant or unpleasant — excuse me, if the Chinese pictograph is pleasant or unpleasant.

Then after the Chinese pictograph disappears, there's going to be what's called a "masked" slide that will appear. So you'll only get to see this for a hundred milliseconds; you get to see this for 75 milliseconds, and then there's a blank screen for 125 milliseconds in between. Then you simply after you see this, you say, "Was this Chinese pictograph pleasant or unpleasant?" Okay, very simple. Just ignore the first picture, and judge the second picture. Fine, simple. You do like ten trials of this.

The logic is, as I just said, to ignore the real life picture and judge the pleasantness of the Chinese pictograph —the notion is that the Chinese pictograph is a *nothing* stimulus. You have really, essentially no basis for saying it's pleasant or unpleasant. But by putting before it on certain trials pictures of an object — whatever your reaction is positive or negative to that object —spills over onto your interpretation of the pictograph. This picture makes you feel good or bad, and then you think it's the pictograph that seems more pleasant or less pleasant. So let me just show you some results.

Here's a study in which they exposed people either to pleasant pictures like beautiful nature scenes and couples hugging romantically, or unpleasant pictures like open-heart surgery and a bucket of snakes or neutral pictures like squares on a background *[inaudible]*.

As you can see, and don't worry about "warning/no warning" because it makes no difference, that the proportion of pleasant judgments of the pictograph is higher when the picture is pleasant, lower when it's neutral, and even lower when it's unpleasant. So there's this kind of spillover of the emotional reaction. So you might say, "Okay, well, let's do this with George W. Bush." You don't ask people, "Do you like or dislike him." You just ask them to rate the pictograph a bunch of times and see what you get from them.

As it turns out, they did this with Bush and Kerry, and what you can see is that for people — this is pre-election — who said that they intended to vote for Bush, when Bush's picture preceded the pictograph, they were more positive, and when Kerry's picture preceded the pictograph, they were more negative. Whereas for Kerry voters when Kerry's picture preceded the pictograph, they were more positive towards the pictograph and happened to be equal for neutral and negative *[inaudible]*.

So, this is evidence that you can pick up what people tell you is their opinion, right, because they've told you explicitly, but this is in some sense a validation of that. And the notion is well, maybe the proportion pleasant judgments is a more accurate reflection of their real opinion of Bush and Kerry — even than their self-report. So that's one thing we could try, we have no evidence to support that.

Let me now actually transition us to show you a different way of doing this, which is more complicated. And this is the implicit association test that Skip mentioned earlier. Here is the beginning of it. It says, "You have opted to complete" — you can do it in different ways.

This is the presidential preference IAT. "You will be evaluating George W. Bush in comparison to another randomly selected president or group of presidents. You will complete three tasks — two brief questionnaires and an IAT in which you sort words and pictures into categories as quickly as possible. You should be able to complete the task in less than ten minutes. When you finish, you'll receive results *[inaudible]."*

And the instructions are, "In this task you'll be presented with a set of words and images to classify into groups. The task requires that you can classify images as quickly as you can. *[inaudible]* takes about five minutes. So on the left-hand size, these are good words — joy, love, wonderful— bad words — agony, terrible, horrible — Presidents— Thomas Jefferson and George W. Bush. So I have to keep my index fingers on the "e" and the "i" key, and two labels at the top will tell me which go with which key, and I need to make the judgment as quickly as possible. *[inaudible]*

So if it's Thomas Jefferson, I'd push the right key; and if it's George W. Bush, I'd push the left key. Now if good and bad words show up, I push the left key if it's good words, and the right key if it's bad words. Isn't this fun? *[inaudible]*

Now, here's where it gets hard. Notice if it's Bush or good, I push the left. If it's Jefferson or bad, I push the right. Now, this is the key thing. What we're looking for is the match and the mismatch. On the next set of trials after this one, good and bad are going to switch. So if Bush and good go together in my mind and Jefferson and bad go together in my mind, I will be quick at this and I'll be slow going the other way. But if Bush and bad are connected in my mind, I'll have a hard time remembering which finger to use — supposedly.

[inaudible] Here's the last one. You see how I did. So now it's going to be—did it switch? *(No.)* No it didn't switch, okay, it's the same. *[Pause] [inaudible] Okay, now Jefferson is switching sides. Sorry I didn't to take so much time on this. I want you to see the result.* Okay, fine. So now we've got Jefferson and good. This is the reversal. This is the important one. Let's see what happens. *[Pause]* In the interest of time, *[inaudible] you get the idea, [inaudible]* what you do is you look at whether the speed for the accurate trials is quicker when Bush is paired with good or when Bush is paired with bad. *[inaudible]*

So this is just showing you what Skip talked about earlier briefly, just so you can see it. We usually say, "What job or office does Dick Cheney hold?" but we could instead say "What job or office does Dick Cheney hold?" and more people get this right. Different people get this right. We could ask, "What job or office does John Roberts hold?" and we could show his picture as well. We could also say, "Who is this?" And I don't know if people have done that yet, and it would be sort of interesting to give that a try. *[inaudible]*

Now, the last thing of course that we could do — someone mentioned this — is we could show ads and look at how people *[inaudible]* with the ads *[inaudible]*. So, I just wanted to give you a feel for the kinds of things that the computer allows us to do to start you thinking about if we were to do ads, what ads should we do. What questions should we ask after they see an ad to evaluate reactions to the ad, and anything else you'd like to suggest?

Unidentified Speaker: [inaudible] doing the split sample, it said who is the Supreme Court Justice? [Inaudible]

Jon A. Krosnick: Well, it said, "Who is John Roberts?"

Unidentified Speaker: Oh, who is John Roberts. Right, right. [inaudible] Instead of adding the picture, do them as a split sample [inaudible] then just show the picture, "Who is this?" without the name to see if they're more likely to associate it. [inaudible]

Jon A. Krosnick: Yes, nobody has done that yet, and we should try that. The goal here is to try to find who's getting more political information. There are some people who know these things because they're being exposed to a lot of it, and some people who don't, and so which would you be more interested in? Which would you believe more — the one with the picture and the name, the one with the picture only, the one with the name only?

Frank Newport: Do you mean just in terms of name ID and what office they hold — ? **(Yes.)** I think the question would be, would you gain information if you said "tell me your reactions to Dick Cheney" — whether you showed his picture or not? With the assumption being that if you showed his picture then gathered reaction to that candidate, it would be more accurate, more reflective or you would learn more. That would be the question that I would ask you, I guess. Would that be the advantage?

Jon A. Krosnick: Well, what do you think? That's a great idea. In other words, we ask feeling thermometers all the time about Hillary Clinton, you know, and why don't we put Hillary Clinton's picture on the computer screen with that?

Frank Newport: You know, I go both ways when we do research on a news anchor. It would be important to show it, but most people — I would not think, for example, follow me through on this when you were asking the horserace choice who you're going to vote for. I would not think you would want to put the pictures of the two candidates, because in the voting booth scenario there are no pictures.

And a lot of times when we react to candidates, we're reacting without looking at their pictures. The interesting question would be how much do you gain by getting a reaction to a visual image in terms of whatever the goal is as opposed to not. For a news anchor, I think it would be important if you're doing marketing research. If you're going to watch or not watch Katie Couric when she comes on the air, you're replicating that situation by showing the picture, but in this setting we're not — your reaction to a politician and voting for them isn't always in a visual setting where you see they're busy, so to speak. I don't know. *[inaudible]*

Jon A. Krosnick: Well, a lot of the campaign you see them on TV, right?

Frank Newport: Yes, but when you're ending up voting you don't see them.

Unidentified Speaker: And you don't see them in a static image. The problem with this is that the image is content. It's the same way as like asking a question with a lot of words in it, explaining or describing the situation to where you're getting an immediate controversy over the content with the choice of content you put in. In a static picture, you know, you can see it as negative; you can see it as positive. It depends on the expression on their faces.

Frank Newport: Or like drugs! *[inaudible]* Yes, absolutely. It will be very dependent on the picture. You know, that picture of Cheney was unflattering.

Unidentified Speaker: Yes, or that baby. I couldn't tell if the baby was laughing or crying, and whether it makes a big difference.

Jon A. Krosnick: Well, the thing about that study is that in such a tiny fraction of milliseconds, you don't need to know. Your eyes and brain will decide what it thinks it's doing, you know? Let's go back, though. Let's just take for a moment the hypothesis that we were interested in considering the possibility of showing the picture. And you said, absolutely right, that which picture we used would matter.

So we could actually randomly assign each respondent to one of 50 different pictures of Cheney.

Unidentified Speaker: Or you could show them more than one. [inaudible]

Unidentified Speaker: A collage!

John Curtice: Right, a collage. But wouldn't the interesting thing be to do — I mean, the questions at the moment you asked verbally, arguably most related to — either a picture or a video of the candidate are the ones that triggers emotional reactions—fear, afraid, hopeful, et cetera. Wouldn't the interesting experiment be whether or not the relationship between candidate preference *[inaudible]* voting intention be greater so far as these effective reactions are concerned, if you showed the images as opposed to simply asking the words?

Because, I mean, I would expect, obviously it does depend on the nature of the picture. But in theory, what you're picking up are indeed people's emotional reactions to seeing that lovely person or that horrible person or whatever. You ought to be able to measure it more accurately through *[unintelligible]*.

Arthur Lupia: That's a good idea, but the more appropriate, I think, instrument for that is to have — I mean, it's hard for us to kind of have a one-off experiment in the middle of the ANES because so many people are depending on using it. That's the perfect kind of thing for TESS, right? I mean, TESS is designed for let's say we have this idea about which image or would it have the emotional response. TESS is the place for one-off experiments where you have to have the idea, somebody else figures out how to implement it.

Unidentified Speaker: It's almost like an aided and unaided question. ***(Exactly.)*** *So you have the unaided without the picture, and then you have the aided with the picture. And I think in terms of Supreme Court Justice or lower level — they know who they are if they see their face, but they don't know their name. But I don't agree with using different pictures, because it's almost like comparing apples and oranges. If you're going to use a picture, it should be the same picture for everyone. I mean, use the same picture that you're showing everyone. Or if you're going to use separate pictures, then I guess you would split the sample on that.*

Arthur Lupia: The fellow who did the initial experiments on that is now at Princeton, and so some of you guys can drive and see him. These were with the knowledge questions, but the biggest gains were with — you probably know the punch line if you think about —people who tend to get their political information from television. If you got most of your political information from the newspaper — *(Gain in what?)* — Your ability to correctly identify, or to correctly answer the question.

Unidentified Speaker: But that doesn't sound like it's of any interest to the ANES. You don't care how — ?

Arthur Lupia: Well, you know, I think that the interest is that a lot of people actually use the political knowledge questions to try and assess awareness or something like political IQ and things of that nature. And experiments, I think, like this call into question a lot of the common interpretations of the political knowledge questions. What are you really tapping there?

Jon A. Krosnick: I mean, Scott is really the reigning leader in this area. So the question is if you could go back to the '40s and '50s and the old Gallup questions that are essentially quizzes about politics, you'll get the wrong answer if you don't know the information; but you'll also get the wrong answer if you don't understand the question. The idea here was to say well, what if you asked what job or office does William Rehnquist hold? A bunch of people could say, "I don't know." But then if you were to show them the picture, they would say, "Oh, I know who that guy is; he's the Supreme Court guy."

The failure to show the picture caused our numbers to be lower than they should be. Now, if that's particularly true among heavy television viewers, it's not surprising, right? Now, you could argue that they're seeing his face on TV, but they're not actually paying any attention. But seeing the face actually allows them to say, "Oh, yes, that's the Supreme Court guy," then they do have some information. Thus we have underestimated knowledge. I would actually love to hear, you know, your thoughts—

Harold Clarke: I think it's worse than that, though, because what you're saying is it's heterogeneity, the accuracy of the measure. And then you're going to go ahead and use that somehow *[unintelligible]* with how political sophistication interacts. *[Portion unintelligible]* It's a little worse. It's not just hitting the margins.

*Unidentified Speaker: Right, precisely, and the correlation between the demographics — **(They change.)** One thing that Prior's work probably indicates is that maybe there's not quite as much of a class skew in knowledge as we thought. Maybe women are not significantly less politically knowledgeable than men as some friends/researchers have argued—*

Jon A. Krosnick: That would be you? *[Chuckling]* ***(Yes.)*** Because you were a victim of the surveys that existed at the time that you did it. *(That's right.)*

Unidentified Speaker: I can think of two ways in which pictures would be helpful that aren't on policital grounds. One is on "favorable/unfavorable." I don't know if you ever do that, but we all the time talk about whether to [inaudible] to somebody and not help to identify them, but a picture would be a lot less likely to skew the results it would seem to me than [inaudible] would, and yet it would help you.

I mean, a lot of times I feel like people have an opinion about it, but in order to try to [unintelligible] tell who it is, you wind up affecting the results. A picture is less likely to do that. I think the same thing would be true, and it probably would affect you as much as it affects us, but asking questions about news events where you could show a picture of some [unintelligible], they actually do have an opinion about the event, but to try to describe the event, you would probably affect the outcome of the situation.

Unidentified Speaker: Why would you use a picture and not a video. It would work for TV.

Jon A. Krosnick: Yes, we could do that. So if we said, "What job or office does this person hold?" then you could see them talking —

Unidentified Speaker: But isn't there a problem that it's like giving all of America a participation trophy? You've made it so easy to say, "Oh, yeah, that guy is the Supreme Court guy." That leads you to the conclusion that knowledge means nothing, when in fact a harder knowledge test, which may be who is John Roberts— whatever you want to do — actually does tap into some elite level of knowledge that meaningfully relates to added voting intentions and other attitudes, but that has to be something you've studied. It seems like you would have to do it three ways to say, you know, everyone who knows anything about this [inaudible] has no impact, but those who know the most behave this way, and so you have to do three tests.

Arthur Lupia: No, I think you're right. I mean, Markus Prior and I actually have, we have an article coming out in the next AJPS talking about different ways of trying to interpret political knowledge and how you have to measure it. In particular, what we did is we went on a Knowledge Network's poll and we ran a two-by-two experiment where we varied whether people had a very limited amount of time to answer a question, or whether they had an extended period of time.

Then we also varied whether or not we paid them for the correct answer — one dollar. It turns out that doing either of those things — giving them unlimited line or simply paying them — increased the rate of correct responses by 18 to 24 percent. So both of those mean something, right? I mean, the idea is that when you give them extra time it increases, I mean maybe they're looking it up, you know? That could be.

But if we're going into the online world and trying to measure political knowledge, it suggests that maybe we're measuring a different animal than we were measuring in face-to-face or phone.

The other interesting thing about that is paying people also had an effect, even if you give them no extra time. And the idea there is that in a normal survey, maybe people just don't try so hard. Hey, well I'll give you a dollar if you get it right. "Oh, well, it is Dick Cheney." Right. It suggests something, too.

So those types of experiments, I think, and building on the work that Scott did, you know, conceptually I think there is a sense of well, political knowledge just means one thing, and we all know what it is. And I think

what the evolving work is, well, we need to decide what we're measuring first, as you were saying, and then try to figure out the best way to do it.

Unidentified Speaker: Yes, and beyond that I would suggest that you're not measuring knowledge at all; you just measuring recall.

Jon A. Krosnick: Yes, absolutely. *[Agreement]* It's a very indirect measure of something other than. I mean, this is Skip's main point, and everything is riding on this, which is, answers to these questions are typically utterly irrelevant to what we're really trying to understand. They are proxies and not great proxies, at that.

But Scott, so could you just say what your view is? Cause I think that Rich's point is really fundamental here, and it's true about all of educational testing, right? You can make a hard test in math, or you can make an easy test in math. And if you make it a hard test, then hardly anybody is going to get it right. If you make it an easy test, lots of people are going to get it right.

In your book you had sort of like this bell — double thing where the bulk of people were moderately good and a few people were really good, and a few people were really awful. But if the questions had all been easy, then 100 percent of the people would have been excellent and nobody would have been awful.

Scott Keeter: Well, the question which is implicit or maybe explicit to what Rich said, is ultimately would your substantive conclusions about engagement levels across the demographic spectrum, be different if you allowed different channels of reception of information to, you know, be expressed back — rather than just the old-line verbal cues that are built into a lot of our questions?

But I don't think that this is an either-or, because especially if these are things that are going to go on the before/after a big face-to-face study. You already have imbedded in your core questionnaire plenty of very good knowledge measures of the old sort — and especially these comparisons of, "Are the Democrats to the left of the Republicans on this issue or that issue?" The left and right is irrelevant, you know, but which one falls here and which one falls there.

You've already got the ability to do it the old way. And then if you add a few of these then you will be able to say, "Well, this actually changes the way, you know, we think about the demographic characteristics correlate with things that are interesting in terms of knowledge and the things that knowledge predicts."

Jon A. Krosnick: So, actually, Rich, I think that one way to think about blending what Scott has just said and the educational literature is that you should never look at any one test question. What you need is a test with a set of questions that vary in difficulty across the continuum from really hard to really easy, and so you kind of don't have to choose one or the other — as long as you ask lots of them. You can then place people somewhere on the continuum of lots of knowledge to no knowledge. *[inaudible]*

Frank Newport: That sounds like an SAT score.

Jon A. Krosnick: Yes, exactly.

Unidentified Speaker: The other thing drawing on educational research is that people learn in different ways. And there are people who do not learn well from the written word; they learn visually. And if somebody sees a picture of Dick Cheney and we know that's Dick Cheney, they may have known it, but they couldn't get it from reading a question. They could get it from the picture.

And the same thing with the whole point of doing the video as well. Maybe people learn aurally, and that triggers the memory, the recollection. I don't even know the answer, but I think it would be interesting to sort of test some of these things out to see what the difference was.

John Curtice: Yes, but then if you're right, those relationships if you think that are conditioned on knowledge should be more strongly conditioned with your new measure than with your old measure. If your

old measure is better as a measure of knowledge, then it shouldn't be so good to identify what relationships are conditioned on which measure, you know? There is a good way of testing as to whether or not — simply because, you know, academics are creatures who are used to using written word. They are actually looking at the world in the way that they tend to think of it. If indeed visual cues and all cues are as valid as written cues or spoken cues in producing and evoking knowledge, then we should forecast the condition more, yes.

Jon A. Krosnick: From your work with Markus, do we have an answer to that? Is one of them more valid *[inaudible]*?

Arthur Lupia: I mean, it's a little bit of a different issue. It's not a question of more valid. I think that political knowledge is just too broad and too imprecise a term to measure what these things are giving you. And so what we're saying is that if you go through different measurement techniques, you're getting certain kinds of things. So, we talked a bit about what they are, but I think knowledge is too broad. I don't think that that's really what we're getting at.

I mean, to extend your comments, I think it's not just that we're used to the written word. I think that the questions we select and put on the surveys as somehow representative of knowledge are exactly the things that we wouldn't be caught dead not knowing. But when you ask what relevance they are to the randomly selected citizen, I think the answer to that is — the most positive thing to say is that it's hard to say.

One of the interesting things about the ANES is an argument that has been made, "Well, but it correlates to turnout." The questions that you have correlate to turnout, so that tells you something, doesn't it? Well, we actually went and looked back at how the knowledge questions were selected in the 1980s. One of the main ways that you reduced the set from 40 to 5 or 7 was which ones correlated with turnout.

John Curtice: But if I understand you correctly, that is an issue about what is it that we want to measure **—** ***(Yes.)*** As opposed to how we should measure it, right? ***(That's right.)*** I grant you that if we decide to do it two different ways, there's a presumption there that we know what it is we want to measure. You're trying to say there's a bigger issue which is what are we trying to measure in the first place?

Arthur Lupia: Yes, I'm writing a book on that now. The problem, the word that I would jump on in that sentence is "we." There is no *we*. People all think that they're talking about the same thing, but people actually want to measure very different things by knowledge.

Jon A. Krosnick: *[Portion inaudible]* So the question is with that reaction time stuff — the IAT, the misattribution paradigm — do you have any interest in trying to see whether that can predict people's vote choice better than their self-reports, or is this like way too hokey for you?

Unidentified Speaker: [inaudible] in the newspaper. Maybe on the Internet.

Jon A. Krosnick: The New York Times got a brain scan on it.

Unidentified Speaker: No, it would be easy to report that stuff. I just wonder if you could pull some of it off up, acknowledge something in that amount of time. I mean, Jon, you have trouble working your way through that one. [inaudible]

Jon A. Krosnick: Yes, there's a mini version of it. This is the long one.

Unidentified Speaker: I'm trying to think of the typical American in their living room after they've been enjoying a drink with dinner trying this—

Harold Clarke: *Jon, one of the things I was trying to do as you went— there seems to be a learning curve, and that's something that I noticed right away. I said, "Hey, I'm getting better at this."*

Jon A. Krosnick: Yes, everybody does and it's built in.

Unidentified Speaker: If pre-election polls were highly inaccurate of predicting vote choice, then I would say, "Yes, let's get into all of these ways at helping people figuring out whom to vote for." But since it's pretty accurate asking them whom they'll vote for, I don't see the need to go too far into it.

Unidentified Speaker: Were you going to use this with ads?

Jon A. Krosnick: Not this with ads, but we're open to the idea of putting ads on and asking people to evaluate them. Somebody said that that was of interest. Jeff, maybe you said it? No, but somebody said that evaluating ads was of interest.

Unidentified Speaker: [Inaudible] I don't know how valid it is on a national survey, but on a regional — ?

Jon A. Krosnick: Well, you would hope that nobody would say they saw them if they couldn't have, unless they were traveling to that place. Sorry — ?

Unidentified Speaker: [Inaudible] Maybe it would be useful like at a stage like now, in primary season, where you know, there are all these people running for president, to get a read of preliminary tests. [inaudible]

John Curtice: Jon, were you thinking of using — you know, you spoke this morning about doing stuff and getting social stereotypes by gauging people's reactions of being able to link them. Were you going to do that using this method? I mean, I suppose that the question going through my mind is could you not extend that technique to getting people's evaluations of candidates? In other words, how far they link candidates with descriptive words? *[inaudible]* I could see that, and particularly in the way that you train people. *[Portion unintelligible]* Again, you could compare it with the predictability of the traits as *[inaudible]*

Jon A. Krosnick: Absolutely.

Unidentified Speaker: I mean, if it were true that these kinds of techniques and the use of the visuals gave you a better reading of emotional reactions to candidates, [inaudible] then I could see the value of it in the context of the National Election Study.

I mean, a lot of the uses that we're discussing here are things that could be run on TESS or in any other context. The question would be is there something that you could capture in your panels about people's initial emotional reactions to candidates, you know, that could then turn out to be predictive of how they behaved later on in the campaign or the kinds of reactions that they have?

I don't know enough about the area of measuring emotions to know if this is better than, you know, something that's just more word-based. But it strikes me that it probably would be, because it seems like you'd get more of an emotional charge out of seeing Hillary Clinton than you do out of hearing her name read, or out of seeing Dick Cheney than just seeing his name.

Unidentified Speaker: Yes, and I would only add that it all seems a little based in unreality, you know? You might pull it back and test what Scott's kidding about — what Scott is talking about —but taking it down to earth and use ads. Because that's what people are truly exposed to rather than the photos [inaudible]. You know, Kenny Goldstein is really good on that subject that I'm talking about.

Unidentified Speaker: [Portion unintelligible] Showing little news clips similar to things that you would see on the news or that would be used in campaigns of, you know, the Twin Towers falling, you know, soldiers in Iraq. And just those little images and then asking candidate attributes produce an understandable reaction, and an interesting one — a politically relevant reaction.

So it isn't necessarily basically saying that we have a lot of things, and these images tap or push these buttons for these candidates. And that would allow you to, at least in a timely fashion, kind of get a jump ahead of the ads and say, here's what they're going to [inaudible] because this stuff works. Now, maybe your retrospective could say, "Well, we knew about that. We knew eight months ago that these images were used in print ads."

Arthur Lupia: This is very helpful. Why don't we move to the next topic? We're trying to do the lightning round thing, but it's working about as well as it did last night and the previous debates.

In this section we want to talk about questionnaire design, and in particular, some issues that we regularly face and get your points of view on it. So, do you have a presentation for this? All right, I can tell you about some of the things, but then Jon will make a presentation. We worry a lot about question wording, we do a lot of evaluations with various wordings of questions. We also do a lot with response categories and how those are described and their nature.

NES has a tradition of using scales, both within a question where you'll have people answer on a 7-point scale or a 100-point scale. And we do a lot of inquiry about how respondents see those, and whether they're all seeing it the same way. And if they're not seeing it the same way, what it means for the data. I think we're going to bring in front of you some of the ideas that we have, and some of the things that we're dealing with and try to get your feedback about how and whether to worry about these things.

Jon A. Krosnick: So, I just put together a little collection of questions from our surveys that parallel questions from your surveys to highlight some of the differences. I can imagine in some cases you will say, "Yeah, yeah, no big deal. That's basically the same question." And in some cases you might not say that, and maybe you would encourage us to move towards your way of doing it or whatever.

So here's our feeling thermometer question. We actually show the respondents the thing on the right-hand side. It goes from zero degrees to 100 degrees and it has labels on 9 of the points. And then there's this very long explanation that says, ratings between 50 and 100 meaning you feel favorable and warm; ratings between zero and 50 meaning you don't feel favorable towards the person, that you don't care too much for that person. You would rate the person at the 50-degree mark, if you don't feel particularly warm or cold towards the person.

So, obviously you don't do this over the phone. So, here's how at least two survey organizations represented here have done this over the phone on occasion. So there's this long version of it. "I'd like to get your feelings towards some people who have been in the news. I'll read the name of the person. I'd like you to rate the person on something we call a feeling thermometer, you can choose any number between zero and 100. The higher the number, the more and more favorable you feel towards that person; the lower the number, the colder or less favorable you'd rate this person. You'd rate this person 50 if you don't feel warm or cold."

Okay, versus here's a much quicker version: "I'd like you to rate how you feel about that person on a feeling thermometer using a scale from zero to 100. The higher the number, the warmer or more favorable you feel towards the person, the lower the number, the colder or less favorable you feel." Does anybody have any preferences among these? Do we need the picture? Do we not need the picture? Is the shorter just fine, or do you need a long description of it?

Unidentified Speaker: How about "strongly favorable, somewhat favorable, somewhat unfavorable and strongly unfavorable?"

Jon A. Krosnick: Yes, do you like that better than all this? Because?

Unidentified Speaker: [Inaudible] Well, I think people respond to the words better that to numbers.. There's a lot more agreement on what the words really mean— somebody is 63, somebody else is 71, and you really don't get anything.

Susan Pinkus: I agree.

Jon A. Krosnick: Okay, Susan agrees with that. Other points of view? Anybody who has asked this question like to defend it, or — ?

Frank Newport: Well, first of all, you're going to use it anyhow because it's historic, right? You're not going to abandon this, right?

Jon A. Krosnick: Actually, for the panel study, I'll confess to you, we've abandoned it. ***(Really?)*** Yes.

Unidentified Speaker: That's something long-term part of the NES, right?

Jon A. Krosnick: We will continue it in the face-to-face, but what were you going to say?

Unidentified Speaker: I was going to say you always end up collapsing anyhow. Maybe that's what Gary is saying. You give them a thermometer, but how do you analyze it? Well, I was just looking at the thermometer we had used recently, not on this, but on well-being. And it was everybody 60 and above or something, and you just end up collapsing it. Anyhow, now for multivariate or other statistical analyses, but then that's back to Gary's point, you may be making a big deal out of something. I don't know.

Jon A. Krosnick: Well, guess what? This won't surprise you even slightly. When you hand them the thermometer like this one on the right-hand side, ***(Oh they'll choose the numbers?)*** nobody says 72 or 57. Everybody picks one of the 9 points with words on them, so they're kind of saying what Gary said, which is they'll pick the points that are clear. Anthony, you want to say something?

Anthony Salvanto: [Inaudible] I'm sure you find in the data you've got some people who use the scale as if it's segmented in the 10 or in groups of 5 — others who say 72 when they're clearly using it differently.

Jon A. Krosnick: Yeah there's like nobody at 72; I'll guarantee you that.

Unidentified Speaker: If you're going to keep this, you should let the respondent use the slider on the screen. ***(Because?)*** *Because it allows them to put it where they want it as opposed to the ten points you've already laid out for them; because otherwise, they're going to look at that, and as you said, they're going to pick points.*

Jon A. Krosnick: Well, I mean, in the face-to-face survey they just say a number. And in the Internet survey they're going to type a number in a little a box. *(Use a slider.)* So you think a slider is preferable? *(Yes.)*

Unidentified Speaker: [inaudible] Is that going to foul—

Jon A. Krosnick: No, not really.

Unidentified Speaker: We're talking about the panel.

Jon A. Krosnick: Yeah *[inaudible].*

Unidentified Speaker: Well, looking at the wording in focus groups, too. People use the dial-a-meter so I mean, people do use the creations and it's readily available, too. We've seen some of that in [inaudible] debates from time to time.

Frank Newport: What's your question here? What's the best way to measure affect towards a political personality?

Jon A. Krosnick: Well, what I was pointing out is that in our version we present the picture, and people gravitate towards those nine labeled points. And in your *[inaudible],* your versions, most of you don't use this, but a few of you do, then you don't have a picture, and it's over the phone, and it can be described either at great length on the left-hand side here, or briefly on the right-hand side. And I was just curious about if you all would recommend if we had to do something like this — don't use the picture; do use the

picture; use the picture and know that it's really a 9-point scale; or don't use the picture and then you'll get the whole scale of 100, or use the slider so you'll get the continuity of it?

Frank Newport: You're assuming you're going to use the thermometer. It's just a question of whether to use the picture and how to describe it.

Jon A. Krosnick: Right, and so if anybody has advice or a preference.

Unidentified Speaker: I don't know whether just a 4-point scale is going to hopefully get you what you want. I've noticed that Clinton and Bush when you look at [inaudible], it's much more obvious that people are moving out to the end than it was with just the various [inaudible]. Academically, I shouldn't be saying this, because I'm more interested in what you said [background noise] [portion inaudible].

Unidentified Speaker: Well, Jon, you're not going to compare this to RDD, are you?

Jon A. Krosnick: No. The goal is just to optimize measure within a mode.

Frank Newport: Well, why not use the picture? If you've got the ability to, why not?

Jon A. Krosnick: Well, because it attracts people to those nine points—

Unidentified Speaker: Well, change the picture. Put zero and 100, but at least have the picture.

Jon A. Krosnick: Okay, would you like that better?

Unidentified Speaker: Have a picture of an actual thermometer with a little red mercury in it.

John Curtice: Imean, Jon, there is no point in having a zero to 100 scale if it's not being used as a zero to 100 scale. So if what you're saying is that that visual queue makes it less likely for you to use it, I mean, subsequently is the question of analysis, of whether or not doing the Internet and doing the statistical analysis you find that keeping it as a linear zero to 100 scale means that it's a better predictor or whether in fact you discover when you collapse it, it's better, in which case, you kind of know perhaps that you perhaps you shouldn't really make this a zero to 100 scale. If you know that the visual queue makes it less likely that people use it, then certainly you're biasing those eventual issues against the use of the zero to 100 scale. *[inaudible]* You could say to people, "Here, go and have a bit of fun." You can try it on somebody who's not a politician, you know? Give them a film star or pop star or something. Do it on that first idea. *[inaudible]*

I mean, I'm just thinking at the moment, I've done this very extensively, *[inaudible]* taking just zero to 10 scales, and I've ended up collapsing them on the grounds that they ended up being more useful in my *[inaudible] analysis collapsed*. *[Portion unintelligible]* The basic issue as to whether or not you're just introducing noise, but actually improving the quality of the analysis. It does also depend on whether you're going to use this as a dependent or independent variable. *[Portion unintelligible].*

Unidentified Speaker: On three of these examples you highlight that the midpoint —on two of these you highlight that 50 is the midpoint [inaudible]. In this version, the midpoint isn't in the middle on this scale. Unless it's some illusion — (Yes, it's too far down.) It looks like 50 is below the midpoint. If anything, the value of the scale would be to really highlight 50 as the middle

I mean, if you're going to use the scale, at least make sure that 50 percent is in the middle visually. And I guess, between the two other options I would go for the longer one, or the shorter one and add that 50 is the midpoint.

John Curtice: There is also the old *[inaudible]* issue *[unintelligible]*. It's not intuitively obvious — particularly, it's not intuitively obvious why you do it in a country which uses Fahrenheit as its principal measure *[inaudible]*. I mean, in my country where we now use primarily Centigrade, using a zero to 100

scale makes *[inaudible]*. But it a country that uses a zero to whatever, 212, scale, it's not obvious why you're using a zero to 100 scale.

Jon A. Krosnick: Yes, with the global warming you should have a zero to 112 now or zero to 122.

John Curtice: Yeah, it's the wrong scale.

Nick Allum: *[Inaudible]* My point was very similar to John's last, as a hypothesis, I suspect that if you allow the slider idea, which is technically possible, I'd strongly suspect you would introduce more error than you would anything else. And I think that ties in with what we know about response scales in general. *[Portion unintelligible]*

Jon A. Krosnick: There was an equivalent of that a long time ago, I think.

Susan Pinkus: I think the question is why would you reduce more errors if you had a sliding scale?

Nick Allum: Because it may be that people can't— that they simply aren't certain enough about their own feelings to be able to connect a very, very long response scale in any kind of meaningful way. I mean, that's the hypothesis.

Susan Pinkus: But if you're giving them a scale of zero to 100, it would be the same in their minds. I mean you'd probably get the same error.

Nick Allum: Right, and then you may find that 70 to 75 or 80, there are all smacking onto the same thing. *[Portion unintelligible]* You'd just introduce parameters around each.

David Sanders: And you're also assuming physical control of the mouse, which may not be, for instance, how I feel. I'm serious, but it's a probability.

Unidentified Speaker: That's a serious point. You're taking your attitude, and first of all, you're asking people to translate their attitude to a number, which people don't do well. And then with the slider you're adding the other alyer of actual manipulation. People respond to think and talk in words, not numbers and not sliders. [interjection inaudible]

Unidentified Speaker: I mean, one possibility is since you could take the slider and as they slide there's a number that's moving along to give them — I mean, if you don't like the number, you don't like the number. That's not going to help, but to get some calibration so that, Nick's point—[interjections inaudible]— *some way of getting multiple feedback.*

Jon A. Krosnick: David were you going to say something?

David Sanders: I'm just going to say something about this midpoint, of feeling neither warm nor cold. If I feel angry about a person, am I warm or cold, you know?

Jon A. Krosnick: It's supposed to be a metaphor. Warm is, probably the hotter you are, the more positive you are. But your point is that the more angry you are, the hotter you are, too. Why does that not count?

John Curtice: Well he likes skiing so therefore he prefers cold weather.

Jon A. Krosnick: Let me, just to tell you in response to Frank's question, this is what we're doing on the panel study. We're doing a 7-point scale with a branch, where we say, "Do you like them; do you dislike them; do you either like them or dislike them?" and then we branch the endpoints into three categories. And this is in *[inaudible]* what Nick said earlier about 7 to 9 kind of being the territory. So now, Gary you can talk about the midpoint.

Gary Langer: We've been over and over it.

Jon A. Krosnick: Yes, but I want to hear from others. You can start it, and then other people will say things. The reason you don't like the midpoint is?

Gary Langer: I think that it doesn't measure — when we insert the midpoint, we encourage people not to express opinions or attitudes that they actually hold. That if we want to measure the intensity and the strength of sentiment, we had better do that with a follow-up.

In other words, get them to express what attitudes they hold by not offering a midpoint, and then measure intensity in the follow-up. I understand if you put in a midpoint if your purpose is not to measure everyone's attitude, but to do the best possible modeling, then you're in effect by placing people more solidly in the extreme, then you jumpstart that. You don't have to use the intensity measure afterwards, and it fits academic models better. But for the kind of work that we do — [Technical difficulties]

Jon A. Krosnick: So anybody want to either endorse that or *[inaudible]?*

Susan Pinkus: I think it's almost like sometimes it depends on the question. And I think it's also like sometimes you use, "or haven't you heard enough about that to say?" On some questions if you put that answer in, you might get a higher response because people don't want to tell you. And other times, you know, you don't put it in. I think it depends on the question whether you want to put it in or not.

Jon A. Krosnick: How do you decide?

Susan Pinkus: That's a tough question! It depends on well, let's say in the earlier parts of the election when you're talking about a second tier candidate, and you're asking for an impression of somebody. You want to make sure that they have the opportunity to say rather than saying "not sure," because "not sure" is different from "haven't heard," that you put in that response and say, "Do you have a favorable or unfavorable impression of Tom Tancredo or haven't heard enough about him yet to say.

And I think that would be a legitimate way of using it, because he is not very well known. And so rather than having it in "not sure," I mean, eventually if you want you could combine the two. But I would say that you would want to see the difference and show like over time whether people who know about him, you know, more people know about him as we get closer to the election or not. So I think that in some respects you would want to put that response in.

Unidentified Speaker: Midpoints and volunteered opt-outs like this, I think for candidates that are well known and I think Susan's point is well taken for early in the campaign and for objects of evaluation that are clearly very obscure to a lot of people. But once you're talking about somebody like Rudy Giuliani, these midpoints and these options, it seems to me allows for psychological characteristics to be conflated with unwillingness to answer the question.

In other words, people who are not assertive are more apt to take the refuge in that, even though they may have an equally valid opinion about Giuliani. You're the psychologist and you probably know more about whether that's true generally, but that certainly is the way we look at it at Pew, and we almost never offer a midpoint or an opt-out like that.

Unidentified Speaker: Jon, there's another issue where there's a divergence between an academic poll that's going to be used particularly for heavy duty and multivariate analysis and a media poll, and that is the cumulative impact of asking the same question over and over by different media organizations. That, in a sense, produces a very newsworthy trend-line. For example, "approve/disapprove" or the bizarre question about "right track/wrong direction."

But nonetheless, you know, we have great trend-lines. Also, they're easy to interpret and they do, even if not precisely, indicate a sense of direction — we're either going going this way, that way or we're flat. Switching to a midpoint and I associate with the points that Gary made, you know, that people may in fact seek refuge

in the midpoint, but it also makes more difficult the reporting and the comparability of this question with what virtually every media organization here asks.

Jon A. Krosnick: We're not challenging what you're doing, but would you want us to shift over to what you're doing to increase this comparability?

Unidentified Speaker: [Inaudible] I would just modify this slightly and essentially lead people to say that they "like or dislike him," [inaudible] to say, "like or dislike him a little." It gives you more points, it will allow people to say, "Oh, I really don't have an opinion, you know?" They're kind of neutral, and stay where they are, or push them in a direction, and it's your call to lump them in with the [inaudible] responses that you get to the other two —

Jon A. Krosnick: So there is a paper coming out in PRQ in two issues I think on exactly this point by a guy named *[inaudible]* who you'll hear more from. He's leaving Stanford now, and he's going to be a professor somewhere else. It tested exactly what you were proposing. Because NES has been branching the midpoint for a long time on various questions, and it turns out it hurts; it doesn't actually help.

That when you take the people who put themselves at midpoint, and you say, "Do you lean in favor or do you lean against?" actually they're picking randomly. And so it actually reduces the validity of the measure. And the next time you ask those very same people just 15 minutes later, they'll go the other way. It's just a flip of the coin.

Frank Newport: For which question is that?

Jon A. Krosnick: A whole variety. I think there's probably like 40 questions in this paper across maybe four experiments, the different platforms — telephone, Internet, face-to-face. It's a pretty thorough analysis.

Frank Newport: Hmm, that's not our experience.

Jon A. Krosnick: Well, let me be clear that this is not for who you would vote for. It's not a dichotomous question like that where people say, "I don't know who I'd vote for" where you'd say, "Well, who do you lean towards?" In those cases, leaning increases accuracy.

But I'm talking about attitude — the favorable or unfavorable evaluations of a person or a party or a policy. In other words, if I say, "Are you going to vote for this one or that one?" there is no midpoint. You just have to pick one of those candidates. There it's best to press them to do it.

Unidentified Speaker: But what does that suggest for what most of us do? Where do those people go now when we just ask them if they like him or dislike him? Presumably, it's a flip. ***(Exactly.)*** *Some proportion are just flipping coins. If you ask them 15 minutes later, it moved from "like" to "dislike."*

Jon A. Krosnick: That's why this is an important paper!

John Curtice: But it means your aggregate *[inaudible]* measure will be fine, because as long as those who don't have a real opinion just go 50/50 to either side *[inaudible]*. Jon, *[inaudible]* have you done any work on whether or not the probability of choosing a midpoint is a function of the number of points offered? In other words, I mean, you're going through an initial 3-point offer and then you're branching. Have you ever compared the use of the midpoint of that structure with, for example, offering people five points in the first place?

Jon A. Krosnick: Yes, if you offer them all at once, the more points you offer the fewer people end at up at the midpoint. But if you do it like this, then it doesn't —

John Cutrice: Right, more people end up in the middle. *[inaudible]* So then the crucial step is if you provide them with the non-branching structure, does that result in greater or less predictability.

Jon A. Krosnick: Well, going out to seven is better. And then once you go past seven, then it complicates things.

John Curtice: Yes, sure. *[inaudible]* Does branching versus non-branching increase or reducepredictability? ***(Increase.)*** So branching is better. ***(Yes.)*** So in other words, it's better to encourage people to be in the middle.

Jon A. Krosnick: Exactly, that's what this image is.

Unidentified Speaker: [inaudible] You asked Bush approval in 2004. "Do you approve or disapprove of his handling his job as president? Would you say that the 2004 election was fair or unfair?" So, — [inaudible]

Jon A. Krosnick: It's filled with dichotomous questions. This is one of the wackier things about this survey. It has evolved over 50 years. And you've got dichotomous questions, 3-point scales, 4-point scales, 5-point scales, 7-point scales, 9-point scales, 101-point scales —the survey clearly doesn't have a preference for scale format.

Unidentified Speaker: The thing to do — you and I have discussed this — but if you want to test your wacky notion that people randomly self-assign if they can't have a midpoint, then test it, you know?

Jon A. Krosnick: That's what *[unintelligible]* paper does.

Unidentified Speaker: No, do it here. Do it on questions like—

Jon A. Krosnick: His data are from the NES *[inaudible]*

Unidentified Speaker: Bush approval? ***(Oh not Bush approval.)*** *Do it on questions in the land that we work in.*

Jon A. Krosnick: Oh no, actually, I'm sorry, Bush approval is in his paper, actually Sorry. John, let me just go back and correct one thing that you said. You said that the gap won't be changed, but that isn't really true. The paper has a distribution of like 10, 30 and 60 — in the middle. So if you split 30 and 30, the two sides seem closer together, right? It's now 40 and 70 instead and 10 and 30, so there are three times as many people.

John Curtice: It depends on how you report it—

Susan Pinkus: I was just going to ask you, do you use the question "like or dislike?" I think "dislike" is a word people don't like to say they dislike somebody.

Jon A. Krosnick: So what would you say instead?

Susan Pinkus: I mean, "rather favorable" or "unfavorable" is a softer way of asking whether you like or dislike someone. And I would think most people — I don't know what your survey shows, but I think many people would not like to say they dislike somebody. It's a very harsh word.

Jon A. Krosnick: So it turns out that there are lots of people who say, they're happy to acknowledge they really, extremely dislike George W. Bush, for example. So there is a literature on this suggesting that the more extreme your endpoint labels are — the words are — the more valid the scale. So if you put "like" and "dislike" or "favorable" or "unfavorable," the scale doesn't work as well as if you say "extremely favorable" and "extremely unfavorable."

[Tape 3: 01:00:00]

Susan Pinkus: Right, and so I'm saying to open up the scale. What's interesting and what our report is like is the "very favorable" and the "very unfavorable" scales for George Bush, I mean it's so huge, the "very

unfavorable." So, I mean, it really shows the difference. But I guess for me, the word "dislike" ***(You wouldn't say it.)*** I wouldn't use it, I don't think.

Jon A. Krosnick: So if you think of somebody that you dislike — if I gave you that scale you would say "neutral" or "like," rather than saying "dislike?"

Susan Pinkus: It would be dependent on who it was, but —

Jon A. Krosnick: You can think of somebody.

Unidentified Speaker: That's a different question than what she's saying. She's not saying if I dislike somebody, I would be hesitant to say I dislike them — she's saying I may not think of them in terms of dislike. [inaudible]

Jon A. Krosnick: You may not dislike anyone, right? *[inaudible]*

Susan Pinkus: Yes, there's a very high correlation in our polling between favorability ratings and job approval ratings let's say for President Bush, so we don't really get a lot additional out of favorability than we do out of approval. So, I was looking at your question wondering, it's a much more personal rating. "Favorable/unfavorable" incorporates both their job and their personality. "Like or dislike" is very personal. And so I wondered, did you do that specifically to try to make it different than how they perform in office. Are you really trying to make it who they are as a person?

Jon A. Krosnick: Yes, we are trying to have it be about the person so it can apply to somebody who has not been in office, right? You're never going to say, "Do you approve or disapprove of Barack Obama's job as a senator?" People don't know what that is. We want to ask for a judgment of him.

Okay, so in the few moments we have left, I want to show you one of my favorite NES questions. I don't know if you know this 7-point scale. Some people believe that we should spend less — oh, sorry, let me show you one more picture. *[inaudible]* Some people believe that we should spend much less money for defense. Suppose these people are at one end of the scale, at point number one. Others feel that defense spending should be greatly increased. Suppose that these people are at the other end, at point number seven. Of course, some other people have opinions in between —at 2, 3, 4, 5 and 6. Where would you place yourself, or haven't you thought much about this? Thoughts, please.

Unidentified Speaker: Can you eliminate it from the poll forever?

Jon A. Krosnick: Go right ahead. Tell us why.

Unidentified Speaker: I don't think most people have a metric in their head that goes from one to seven. What's it mapped to? What's it mean?

Jon A. Krosnick: Okay, others agree, disagree? Neutral? Neither agree nor disagree? That was a joke.

John Curtice: So you're not providing us with a visual queue, are you?

Jon A. Krosnick: Well, there actually is a visual queue.

John Curtice: Yes, *[inaudible],* do label the endpoints?

Jon A. Krosnick: Yes, there are labels on the ends, and there's a little line— *[inaudible]*

John Curtice: Yes, because that makes quite a difference.

Jon A. Krosnick: Yes, sorry I was looking for that but it's not in this file. So assume there's a little *[inaudible]*. Rich why are you smiling?

Rich Morin: I really haven't thought much about this. If you presented this scale to me, this thing that no conversation has ever offered to me before. This is the first time I've heard about this, or I would expect a huge, a large — it seems a much simpler way, a quicker way.

Jon A. Krosnick: Go right ahead, *[inaudible].*

Unidentified Speaker: Well, I mean just "favor" or "oppose."

Jon A. Krosnick: Well, here, let me show you some examples of your versions of this question. "Do you think that we should increase our defense spending, keep it the same or cut it back?" Does that strike you as improvement? *(Yes, great question.)* "I'm going to read you a list of present federal government programs. For each I'd like you to tell me whether you feel it should be expanded, cut back or kept about the same." *(It's fine.)* It seems fine? Any preference between those two?

Unidentified Speaker: The first one. [Several in Agreement]

Jon A. Krosnick: So this is sort of like Rich's, "do you favor or oppose increasing defense spending by at least $20 billion by the year *[inaudible]" ?*

Unidentified Speaker: If that's the proposal on the table, that's the best one. If it's not, then the others. It's just a theoretical question. [inaudible]

Unidentified Speaker: The trouble with all of these, of course, is relative. How much is defense spending? And if you don't know that, then what does your answer to any of these questions really mean — scale or no scale? [Portion garbled]

John Curtice: Can I raise an objection to these questions? My concern about these — *[unintelligible]*. In other words, people may find it more difficult to say something should be cut back or to oppose something, whereas the advantage of the ANES structure in which you are legitimate in both points of view in the way in which the question is introduced. I mean I'm not sure whether or not the acquiescence bias is as serious for this kind of structure or not *[inaudible]* The nice advantage always of having both end points put forward in a legitimately similar way, because you're trying to eliminate, I think —

Jon A. Krosnick: So does the first one of these seem not to do that, increase — ?

John Curtice: Well, it's a question of load. It's whether or not you think people are going to find it difficult to use the words "cutting back," or not. I mean, maybe it's okay, but it doesn't introduce the options in a way that clearly says, one of these views is just as legitimate as the other, you know?

Unidentified Speaker: [inaudible] Keep it about the same.

Jon A. Krosnick: Evans, what are you thinking there?

Unidentified Speaker: [inaudible] and I were just having a little side discussion. Basically, this is a question about whether or not you have a policy preference for defense spending. And asking about $20 billion, unless there's a well known — (Value judgment.) Well, well known proposal that is being debated and covered in the news media, so people have some clue of what you're talking about — adding a dollar number to it is completely hopeless.

Unidentified Speaker: But that's what that question is; it's a proposal.

Unidentified Speaker: Then it's just "should defense spending be a priority or not a priority?"

Jon A. Krosnick: How about this one: "Imagine the President and Congress decided to cut defense spending by 10 percent, not 11 percent, — and directed this money to improving education, strengthening Social Security and paying down the national debt instead? Would you support this movement?" *[Laughter]*

Jon A. Krosnick: Okay, here's another one. "Thinking about all of the programs the federal government spends money on, do you think defense spending should be given the highest priority, or are there other programs which deserve higher priority?" Or, "I'm going to read you a list of federal programs, and for each one please tell me how important you think the program is. Is it very important, somewhat important, not very important or not at all important? What about defense spending — ?"

Unidentified Speaker: Well, Jon, but the point is that we all know what all these questions mean in a way that when you have these numbers and you're a seven and there's a six — I don't know what those mean. [inaudible] I don't know how to report it, but I also don't know how to answer it. With these, you know, there's good questions; there's terrible questions, and there's in-between questions here, but we know what they mean. Words are a good thing.

Jon A. Krosnick: Okay, good, all right. Skip is going to say we're over time. Let me just for the fun of it show you this one now. Approval and then we'll stop. "Do you approve or disapprove of the way that George W. Bush is handling his job as president?" Okay, and then "Do you approve strongly or not strongly?" So everybody think that your wording is the same as that? It's not. So, let me just show you what I've found. So there's our wording again. Here are some media polls.

"Do you approve strongly or somewhat" instead of we're saying "not strongly." "Do you approve very strongly or not so strongly?" There's another one. "Would you say that you strongly approve, or just somewhat approve?"

Mark was talking earlier about trend lines across survey organizations. You can do it with the first part of the survey, because you all asked that the same way — but you don't all ask the follow-up the same way. And so if you want to draw trends, it may be worth doing. I mean, would you say these are so trivial that there is no point in bringing out these differences, or would you hesitate to combine them? *[inaudible]*

Unidentified Speaker: I mean, instead of asking us whether we think it is, we should be looking at the data and deciding.

Jon A. Krosnick: Okay, and if the polls aren't going on at exactly the same time, then you'll see it bouncing around. It's a little hard *[inaudible]*. Anyway, thank you very much. You may not realize how helpful this feedback was, but you have reinforced a lot of our personal hesitations about these items. We tried to get it from you in a way that didn't push you in one direction or another.

I hope that you felt you weren't pushed in one direction or another, but you have on your own voiced a lot of the concerns that we have about these items, and then lead us to worry about them and reconsider whether to stick with them. And so we're moving in new directions with the panel study in response to already a lot of the feedback you've provided, so thank you for that.

Arthur Lupia: So, would you like to take a break? Okay, maybe, what do you think, 10 minutes, is that all right? All right, 10 minutes we'll come back.