

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

UNIVERSITY OF HOUSTON
CENTER FOR PUBLIC POLICY
SURVEY METHODOLOGY: NEW DEVELOPMENTS
HAROLD CLARKE AND MARIANNE STEWART
April 30, 2008

BE IT REMEMBERED that the aforementioned
proceedings were heard on the 30th day of April, 2008,
beginning at 10:09 a.m., at the University of Houston,
4800 Calhoun, Heyne Building, Room 135, Houston, Texas
77004, reported by Dorothy A. Rull, a Certified Shorthand
Reporter in and for the State of Texas, as follows,
to-wit:

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

A P P E A R A N C E S

HOST:
 Jim Granato, Ph.D.

SPEAKERS:
 Harold D. Clarke, Ph.D.
 Marianne C. Stewart, Ph.D.

CENTER FOR PUBLIC POLICY:
 Mike Angel

GENERAL PUBLIC

* * * * *

1 P R O C E E D I N G S

2 MR. GRANATO: Okay. I would like to
3 bring this talk, workshop to a beginning here. As we go
4 through the process of devising a plan for the Houston
5 Region Panel Study, one of the issues and challenges that
6 the prior workshop discussed was the issue of mode. And
7 I could think of no better individuals to discuss the
8 evolution of mode and the future challenges dealing with
9 mode, that is, how we collect our survey data, than
10 Marianne Stewart and Harold Clarke, who have not only a
11 national reputation, but an international reputation in
12 survey methodology.

13 One of the things that they do is they
14 are PIs for the British National Election Study, which
15 next to the American National Election Study, is probably
16 one of the most prestigious in the world. Certainly at
17 NSF when I was there, this was considered one of the top
18 studies in the world.

19 So the structure of the discussion is
20 going to be two hours long. They're going to present
21 evolution of the process, some contemporary issues that
22 they're dealing with in their own research, and then
23 there will be a question-and-answer session.

24 Without further delay, I would like to
25 introduce Harold Clarke and Marianne Stewart.

1 MR. CLARKE: Marianne is going to begin.

2 MS. STEWART: Well, thank you for joining
3 us today. Today gives us an opportunity to do at least
4 two things, one is to be part of a formative stage for
5 the Houston study and, second, and relatedly to talk with
6 you or speak to you about our experiences doing
7 large-scale surveys in a variety of places, including
8 Canada, the United States, and the United Kingdom.

9 Our presentation or talk with you today
10 will have three or four parts. First, I want to give you
11 an overview of surveys. I'll try and do that in as
12 concise and as helpful way as possible. Second, we'll
13 talk with you about where we are now. Harold will
14 particularly talk about the use of Internet and
15 face-to-face survey in a large heterogeneous population,
16 the U.K. Third, we'll talk about into the future.

17 Now, there are a lot of things to think
18 about with regard to surveys and what they might look
19 like as they unfold, evolve and then evolution in a
20 biology way, if you will, over the next decade or so.

21 There are two points that we want to
22 address with you today. One has to do with what the
23 future capabilities of devices such as Internet will
24 allow all of us to do. A key point here is that the
25 Internet mode, much more than any other modes that

1 surveys have been used with or for, really is a very,
2 very democratizing mode. All of a sudden, it allows
3 people to do research who previously have been deprived
4 of doing any kind of survey work at all. The second
5 point I want to discuss with you focuses on technology
6 and demography. So I'll come back on those in just a
7 moment.

8 The overview part of the presentation
9 today, the first part, kind of where have we been, what's
10 been the past, there are three or four bits that I want
11 to mention to you. One is just a very brief chronology
12 of survey work. The second has to do with what are the
13 goals of most surveys, at least, goals of most social
14 surveys or political surveys. The third point has to do
15 with what are some of the key questions in survey design.
16 And then the fourth point has to do with mode and mode
17 comparisons. And so each of these four points in the
18 first part cumulates, if you will.

19 The first bit, the overview or the
20 historical chronology, most survey texts will tell you
21 it's really very difficult to determine when the first
22 surveys were done. We know, based on the written record,
23 that during the time along the period of the Roman
24 Empire, there were orders from Rome to have census done.
25 We also know that in the aftermath of conquest of the

1 U.K., Britain, in 1066, the Dome's Day record was
2 commissioned, collected, and still to this day
3 constitutes a very interesting historical archive.

4 We skip through time to the late 19th
5 Century, and this is when we really first begin to see
6 social surveys being done. And these social surveys in
7 the late 19th Century focused on two questions. First of
8 all, they focused on social conditions; and secondly,
9 they focus very specifically on conditions of poverty in
10 urban centers in the United States and the United
11 Kingdom.

12 We move forward a bit further. In the
13 mid 1930s, we now see George Gallup introducing or
14 initiating his political polling. 1946 brings together
15 academic and market researchers to form AAPOR, American
16 Association of Public Opinion Research. It has its
17 counterparts in a variety of countries. Their key point
18 is to ensure the integrity of survey work.

19 1950s and '60s, particularly on the heels
20 of some dramatic failures in political polling or
21 political forecasting, especially coming off the 1948
22 election, we begin to see much more rigorous
23 methodologies being used for personal or face-to-face
24 interviewing. In the '50s and '60s, face-to-face is the
25 dominant mode. In the '70s and '80s, it's telephone

1 interviewing. '80s, '90s, we move more into the
2 mail-back mode. Today, Internet is not the dominant
3 mode, but it is the mode that's capturing a lot of
4 people's attention because of its potential.

5 Today, then, there are surveys sponsored
6 by government agencies, private corporations,
7 universities and colleges, private consulting companies,
8 advertising agencies. The key point is that surveys are
9 prevalent and most of us know that.

10 The second bit, if you will, of this
11 overview section, what are the goals of most surveys,
12 most social surveys? Essentially, there are four goals.
13 The first is to try to get a sense of a distribution of
14 attitudes, beliefs, and behaviors among a population of
15 interest. The second has to do with trying to discern or
16 identify changes or trends in those attitudes, beliefs,
17 and behaviors. The third involves identification of
18 group differences, again, in attitudes, beliefs and
19 behaviors. And a fourth -- and the trickiest goal, if
20 you will -- is to try to be able to locate or again
21 identify causal patterns in order to make interesting and
22 hopefully important causal inferences about connections
23 in these attitudes, beliefs and behaviors.

24 To achieve those goals, survey design
25 becomes critical. And this is the third bit, if you

1 will, of our introduction. There are lots of things to
2 think about when we design surveys, what populations
3 should be studied. That's a question we're going to come
4 back on later in the presentation in the third part, if
5 you will, on technology and demography. Another question
6 is who should be interviewed. And a third question is
7 how should the data be collected.

8 This third question leads us directly
9 into the last bit of the introduction in the overview and
10 that has to do with mode of data collection. There are
11 several modes as I have already mentioned; face to face,
12 telephone, mail-back or self-completion, and Internet.
13 And each of those -- next slide, Merrill. Thank you.

14 Each of those, when you're thinking about
15 which of them to use, are critically important, which
16 combination of them to use depending on the research
17 question you have can be compared in terms of several
18 dimensions. One dimension -- and the dimension you hear
19 most frequently -- has to do with what's the cost of the
20 interview. We're not going to go through all of this --
21 all of the cells of the grid in complete detail.

22 But in sum, face-to-face interviewing is
23 extremely expensive. Current estimates for a 2009/10
24 British Election Study is approximately 325 pounds per
25 interview. That's 700 -- roughly at current exchange,

1 \$700 per interview, okay. If you're going into ethnic
2 minority communities where you have to do face-to-face
3 interviewing and you have a lot of cultural conditions to
4 think about, the estimated cost per interview, at least,
5 doubles and possibly triples. So it is very expensive.
6 Telephone interviewing used to be considered inexpensive.
7 That's no longer true. Mail-back is cheap. Internet is
8 still relatively inexpensive.

9 A second dimension to think about is what
10 the response rate is. This is a key point on all of
11 these modes. Generally speaking, face to face is
12 considered to have the best or largest response rate.
13 Harold mentioned some of the key face-to-face surveys
14 done in the U.S., also, in the U.K. are about 62 percent.

15 Telephone in the mid '90s, it was argued,
16 achieved response rates of about 70 percent. That's no
17 longer true. Our best estimate is probably about
18 20 percent on telephone.

19 MR. KLINEBERG: 20 percent?

20 MS. STEWART: 20 percent on telephone.

21 MR. CLARKE: It depends. It varies
22 widely, and it's hard to get the figures as well. You
23 know, they use different denominators, different ways of
24 measuring it.

25 UNIDENTIFIED SPEAKER: How you measure

1 it?

2 MR. CLARKE: Our friend and colleague,
3 David Sanders estimates, in the U.K., for major agencies,
4 reputable places you go to to do telephone surveys, the
5 true response rates are probably closer to 5 percent than
6 20 percent, so...

7 MR. KLINEBERG: Oh, geez.

8 MS. STEWART: Mail-back response rates
9 may be about 10 percent. Internet, at least based on our
10 experience, you can get 50 percent -- achieved response
11 rate of 50 percent or higher, okay.

12 A third dimension we'll talk about is
13 length of interview. The key cost of face to face is
14 getting the interviewer to the door. Once the
15 interviewer gets to the door and in the house, it doesn't
16 matter -- it matters to the respondent. It may not
17 matter to the interviewer -- whether she's there for an
18 hour or two hours or three hours. You don't send
19 interviewers to people's houses or places of employment
20 for five minutes, okay. That's just not cost effective.
21 So you can do relatively lengthy surveys face to face.

22 Telephone, it's about 20 minutes on the
23 phone before people hang up. And mail-back, they used to
24 be considered almost infinite in their length, but now
25 the best estimate is probably about 40 or 50 questions

1 depending on how you define question.

2 Internet, you can probably capture them
3 for about 100 to 120 questions for a major survey. But
4 there are all sorts of conditions and cautions to think
5 about. So the point of story is that face to face does
6 tend to be longer. Mail-back, phone are briefer.
7 Internet, somewhere in between.

8 A couple of other dimensions and then
9 we'll turn this over to Harold. One dimension has to do
10 with what's called "candor of the interview" and this has
11 an enormous implication for data quality. The candor of
12 the interview, just among other things, means how freely
13 does the respondent feel in terms of his or her ability
14 to respond to questions. We know that there are a number
15 of interviewer bias effects introduced in face to face
16 simply because you're sitting there and you're talking
17 with another human being across from you; you don't know
18 who that person is and you're not as likely to say some
19 things that you might, you know, if you're in "a private
20 circumstance" such as on the Internet. You get more
21 candor on the Internet. You get more candor face to
22 face. You get less when there's another person who is
23 conducting your interview.

24 The final bit I'll mention has to do with
25 data collection time, and here a large face-to-face

1 interview can be in the field for anywhere from two
2 months and longer, which means the real world has the
3 opportunity to intervene and change people's attitudes.
4 The Internet typically is in the field for about two
5 days. Data come back very quickly.

6 Okay. Harold, do you want to talk a bit
7 more about...

8 MR. CLARKE: Marianne has, you can see, a
9 bit of an allergy and maybe a cold, but I'll pinch-hit
10 here. I will carry on with this particular point.

11 It's very a important point theoretically
12 and that's the time with which it takes for you to gather
13 the data that you're interested in. Marianne mentioned
14 that traditionally face-to-face surveys, again, many of
15 these things we're never worried about, but theoretically
16 we should.

17 Your traditional ANES or CES or BES
18 post-election survey will be in the field for anywhere
19 from six weeks to three months typically. And most of
20 your catch will come in the first month, but still a
21 discernibly large portion comes in later. And you can
22 satisfy yourself very easily that the world has
23 intervened. If you're asking people, for example, about
24 what were the most important issues facing the country
25 when they were making up their mind how to vote, you'll

1 have -- numbers that I know from the British 2005 study,
2 a lot of them were mentioning things that were occurring,
3 you know, six weeks, two months later and saying that
4 this is what determines their vote back in May and then
5 something happened in July.

6 So if we're thinking in terms -- one of
7 the big things that we're interested in, of course, is
8 the dynamics of political behavior, social attitudes,
9 economic evaluations, and what have you. We assume that
10 T and T minus 1 can be really anything we want them to
11 be. But of course, theoretically, that's crazy.
12 Psychologists are going to tell us right away that your
13 assumption that it takes a month for people to update
14 their opinions about something, that's wild. It's
15 bizarre. It doesn't make any sense at all. And so there
16 is an irreducible problem with traditional, in-person
17 surveys in this regard; and it's been there from the
18 beginning and it will always be with us.

19 I would emphasize, also, the point
20 Marianne made about the democratization of research.
21 We're talking about within the community of scholars, of
22 course. Because any of you who have been -- Jim, of
23 course, was a political science director at the NSF.
24 Marianne was, and other people have been on panels and so
25 on. And you will see right away that the chunk of money

1 that's taken by the major survey projects, like the ANES
2 and political science in the United States, BES in
3 Britain and so on, takes a huge chunk of that budget.

4 And typically nobody else can come in and
5 say, "Gee, I really want to do some study. I want to
6 study this election. I need some survey data." They'll
7 say, "Well, go use the ANES data. They'll be ready in a
8 year. You can use those."

9 You can say, "But they don't ask the
10 right questions and they don't do the stuff that we want
11 to do." And they say, "Well, tough." And that's really
12 the way it's been for 50 years. And as a result, an
13 awful lot of very good research has never been done.

14 Now, that changes with other modes.
15 Telephone helps, to be sure, and not as much as it used
16 to.

17 Mail-back questionnaires can help. Of
18 course, they're incredibly, incredibly cheap. They have
19 problems as well, of course.

20 Internet is incredibly good in this
21 regard. You get vast amounts of data. Lots of people
22 can be involved. Cost considerations are minimal. And
23 the speed with which data are gathered is extremely fast.

24 What we're going to talk about this
25 morning, in particular, is an experiment that we did in

1 2005 concerning the quality really of data that we can
2 gather for a particular type of survey. We're interested
3 in voting behavior in elections. And one of the things
4 that we want to know, given the considerations I just
5 met, is can we get high quality data via the Internet?
6 It has all these advantages in terms of cost, time,
7 et cetera, et cetera, its possibilities for
8 experimentation, which I'll talk about a little bit
9 later. It has all these really neat things.

10 But the data, are they just garbage
11 and/or are they good data? Are they data comparable in
12 quality, at least, to what we get from traditional sort
13 of gold standard studies, such as the ANES or BES
14 face-to-face survey in the United States.

15 In Canada, Canadians went very easily
16 from face to face to telephone in the late 1980s under
17 the leadership of Dick Johnston and Andre Blais, moved --
18 simply for cost reasons. Can't afford it. Canada is a
19 big country, small population, spread out, can't afford
20 that anymore. Well, let's do telephone. So they just
21 did it, and it was done quickly.

22 They produced some, I think, very high
23 quality interesting studies. They recognized -- Dick
24 recognized immediately the possibilities, the leveraging
25 the opportunities of the telephone surveys and produced

1 these rolling cross-sections, a design that he worked
2 theoretically with Henry Brady in the survey operation at
3 Berkeley. Some very interesting ideas. So they really
4 represent, for RDD, that brand of survey, I call it the
5 CES gold standard. And so what we're doing with the
6 British case is really trying to see whether we can come
7 up with an Internet standard really which approximates
8 these.

9 One of the big -- you know, there are a
10 couple of big things. You think about doing these
11 studies. One, of course, is the idea that Internet
12 surveys, you can't get a list. Can't get a sample.
13 Can't do probability sampling in a traditional way.

14 Whereas for the in-person and RDD
15 surveys, of course, we can do probability samples; right?
16 That's one of their great virtues. Leslie Kish really
17 codified all this stuff for us long, long ago and
18 provided really sort of the theoretical basis, the
19 underpinning of why do people really believe this stuff.
20 "Well, because of this," okay. If you can believe
21 this -- if you're willing to believe the traditional
22 Naman Pearson kind of statistical inference machinery, if
23 you buy that and you know how these examples are drawn,
24 then you can say, "Okay. Yeah. We can proceed. This
25 fits hand in glove. Away we go."

1 The problem, of course, is that anybody
2 can get a list. I can get a list easy, you know, in most
3 cases for a lot of the surveys. Not all of them, but a
4 lot of them we can get a list. In Canada, it was very
5 easy for us, I'd say, early on to get a list. In
6 Britain, it's easy to get a list. Not quite as easy in
7 the United States.

8 We get the list, but that doesn't mean
9 that people answer -- they're willing to be respondents.
10 I can knock on your door, I can send you a letter and
11 invite you, but you may ignore me. You know, you can
12 simply opt out.

13 Now, when Angus Campbell and his
14 colleagues began doing the American National Election
15 Studies back in the 1950s -- actually, the first one I
16 think was done in 1948 -- response rates were really high
17 by contemporary standards. And indeed I can remember
18 when I started doing the Canadian National Election
19 Studies back in the 1970s, we had extensive
20 conversations, of course, with our survey firm. We never
21 talked about response rates. We just assumed they could
22 go out and get the data. We drew the sample -- my
23 colleague, Larry LeDuke and I actually drew the sample,
24 gave them, "Look, here they are. You're going to go to
25 these constituents. These are going to go into these

1 wards. You're going to go to these places, and here are
2 the names we want." And we'd just assumed they'd bring
3 them back, okay.

4 In the 1950s, the published data from the
5 ANES, you know, says that they were drawing 85 percent
6 response rate. The response rate for the 2004 ANES was
7 62 percent. The British figures are almost identical.
8 It's interesting. British figures are almost identical
9 with one exception. In 2001, when we first did the
10 British Election Study, our survey firm there at that
11 time was NOP in the first study. They went down as low
12 as 52 percent, 51-point-something percent. That was
13 judged unsatisfactory both by the ESRC, which is like the
14 British NSF, and by us the PIs. And so that was a big
15 deal for us when we interviewed survey firms for the 2005
16 study, and we drove rates back up to 62 percent. But
17 still that's a long way away from the full sample, let
18 alone the 85 percent, you know, that we were drawing back
19 in the 1950s.

20 And you have to be incredibly heroic, I
21 think, to assume that people who are not answering are
22 just -- that that's a normal missing. And those units of
23 nonresponse, that the idea is those people are somehow
24 just a random subset of everybody else and it's just in
25 the error term that they're not there is, I think, a very

1 heroic assumption. You wouldn't want to make that. That
2 just doesn't seem simple. Like I say, if it's just too
3 simple to be true, it's probably not right; and I'm sure
4 it's not.

5 So unit nonresponse, as we fancy that
6 people won't answer our questions, is now very large in
7 both in person and RDD. Massive in RDD in some cases, in
8 the British case, for example.

9 Internet surveys on the other hand, of
10 course, are nonprobability samples, with the exception of
11 knowledge networks. We'll talk a little bit about how
12 these different firms -- you know, how do they get a list
13 of people to talk to. And the general way that it's
14 done, of course, is to -- in one way or another to
15 advertise, you know, "Join our panel," of Internet
16 respondents, whether it be for YouGov, which is the firm
17 that we work with in the U.K., or Harris Interactive or
18 Polimetrix, which is now owned by YouGov by the way.
19 They have bought a controlling share in Polimetrix. Or
20 in the case of Knowledge Networks, which was founded by
21 Doug Rivers and Norman Nie about 10 years ago now, what
22 they did, they said, "Oh, well, what we'll do is we'll
23 use an initial telephone contact with people" and say,
24 "Hey, come and join our Internet panel. We will give you
25 a computer or an Internet device. And if you will agree,

1 you can have this thing for a year or however long they
2 specified. All you have to do whenever that red light
3 comes on is to answer our surveys." And some of the
4 surveys had to do with political. Some had to do with
5 toothpaste, et cetera. You know, they had commercial
6 clients and so on. And that was their model.

7 Now, that model is an interesting one,
8 but it suffers. Of course, if you can't get people to
9 respond to your telephone to begin with, then this is not
10 going to work as well as you would -- as you would like.
11 And, also, if I believe Doug Rivers -- of course, he's
12 party free now to some extent, having moved over to
13 Polimetrix. I don't know what kind of noncompete
14 agreement he had. Not a very strong one evidently. But
15 in any event, he tells us that there are lots of problems
16 in terms of people dropping out of the Knowledge
17 Network's panel very early. I don't know if that's true
18 or not.

19 It's hard to get good information. One
20 of the points about trying to get research on how to do
21 this stuff is that everybody you talk to has got a
22 horse -- or a dog in this fight and a horse in this race,
23 and so it's hard to get really good information
24 sometimes.

25 But certainly the big rap on the Internet

1 surveys is they're nonprobability samples. People have
2 to opt in. The big rap sampling now from the RDD and
3 in person is they opt out at any moment.

4 MR. KLINEBERG: So it's hardly an
5 improvement.

6 MR. CLARKE: And the question is, is
7 opting out better than us going out? You can think about
8 ways you might want to do this experimentally and so
9 forth, but this is really sort of what the debate is
10 about now. Clearly, it seems to me it sort of maintains
11 the hypothesis here, to start with, is that all modes
12 have selection biases, okay, and they're nontrivial.

13 Okay. What we did in the 2005 BES was to
14 try to design a study that would allow us, from the point
15 of view of political science, okay -- and that's what,
16 you know, people in the different social sciences have
17 different interests -- as political scientists, we say,
18 "Gee, they've got all this great stuff we can do with the
19 Internet if we can believe it, if we believe the
20 results." It has all the advantages that I talked about
21 in terms of cost effectiveness, speed, size of sample,
22 experimentation, et cetera, et cetera. But can we
23 believe this stuff?

24 Some people are very traditional -- our
25 friend Jon Krosnick at Stanford and others say, "No. You

1 can't believe this stuff." Well, yeah, these guys are
2 serious, so you better pay attention to what they have to
3 say. You don't want to gather data which are garbage and
4 are no good. You want good data.

5 And so what we did was to design a study
6 where we have two parallel streams of data gathering,
7 okay. One is the traditional British Election Study face
8 to face. It's a panel, because we do -- it's like the
9 American National Election Study with a pre/post design.
10 Except, unlike the ANES, we top off the post-election
11 wave to make sure we have as representative as possible
12 post-election survey.

13 So we go into the field just before the
14 campaign begins. And of course, in Britain, that's not
15 terrible easy sometimes because you're typically at the
16 discretion of a -- at the majority level, it's going to
17 be at the discretion of the Prime Minister when he
18 decides to ask the Queen to dissolve Parliament. But we
19 could usually make pretty good guesses as to when this is
20 going to occur, and we've been lucky.

21 But anyway right before the election
22 begins, the campaign begins, we go into the field --
23 recognizing how long it takes to do in-person surveys, as
24 I mentioned before, we actually -- since we thought the
25 election would be called right around April Fool's Day,

1 interestingly enough, and it was. It was delayed for a
2 couple of days because of the death of the Pope, as it
3 turned out in 2005, but we started in mid February. So
4 we'll have most of the stuff done -- of these interviews
5 done. We didn't want to start too early because we
6 wanted to have a measure -- you know, like I talked to
7 you about time before. We want to compress it as much as
8 possible, so we started in about the end of the second
9 week of February. Then we have preelection. Then
10 precampaign. Then immediate post election, the day after
11 the election, we start in. And again it takes a couple
12 of months for this to be completed. And then for
13 experimental purposes, we asked all these people whether
14 or not they had an Internet address, and then we came
15 back to them a year later.

16 UNIDENTIFIED SPEAKER: Just a clarifying
17 question. I know in England there's a tradition of quota
18 sampling, and I'm assuming that wasn't what was done in
19 this case.

20 MR. CLARKE: Oh, no. No. What we did --
21 the British Election Studies were modelled right from the
22 get-go on the American model. In fact, on Stokes, went
23 over and spent a term at Nuffield College working with
24 David Butler. They have always been probability samples,
25 sort of gold-standard sample.

1 There are a number of firms in Britain
2 that we don't really hear very much about who sort of do
3 quasi-governmental kind of operations, working for the
4 national statistical office and so forth. They have
5 always used probability. I guess, as a political
6 scientist, at least, we don't hear about them. If you're
7 a socialist or a demographer, you probably know these
8 studies quite well. This is to say this is what we're
9 talking about.

10 In Britain, like the United States as
11 well, there's a tradition -- a long tradition of
12 face-to-face quota sampling with commercial pollsters,
13 firms such as ICM, MORI -- in particular, Robert Wooster
14 at MORI -- Gallup, of course, Bob Wybrow at Gallup for
15 many, many years did quota sampling. Those guys only
16 moved off quota sampling in the early '90s after the
17 debacle of the polling firm's inability to predict the
18 outcome of the 1992 general election.

19 And at that time, they did two things.
20 Some of them -- virtually all of them really went to RDD
21 is what they did, telephone surveys. A few decided they
22 would only do a survey now and then and went to try to do
23 face-to-face probability study or some mixture thereof.
24 Try to get sort of -- the very last poll they would do
25 before an election would be sort of the Cadillac or

1 Mercedes kind of things face-to-face probability and
2 otherwise do RDD.

3 For us, this is RDD. This is strictly
4 ANES, the BES is like in terms of how it proceeds.

5 Paralleling that, though -- and this is
6 the experiment -- is down here. It is the 2005 British
7 Election Study Internet campaign, and it's actually a
8 rolling campaign panel survey. Before the election
9 begins, before the campaign begins, rather, we
10 interviewed 8,000 people almost. And in this case, we
11 could do it just on the weekend before the election
12 begins because, again, the great speed of the Internet.
13 We do not smear the interviews out over six or eight
14 weeks, but rather over approximately -- virtually, it's
15 all done in 48 hours, okay.

16 And then we took those 8,000 people and
17 randomly interviewed about 200 of them every day during
18 the period of the campaign. This is the Brady Johnson
19 rolling campaign panel design. Then we interviewed
20 everybody, even if they didn't participate here, if they
21 had participated there, we tried to interview them again.
22 "Interview," quote/unquote, to survey them again in the
23 post-election survey, which again was completed. You
24 know, we took the interviews for four or five days or
25 surveys for four or five days, but virtually all were

1 done in about 48 hours.

2 Then notice this. A year later, we come
3 back to these guys again and we're again back to
4 everybody who had participated here. Notice the
5 retention rate, which is really interesting. We have
6 6,186 of the original 8,000 respondents. All right. So
7 you get very high panel retention, which we're very
8 pleased with because we're very interested in the
9 dynamics of particular attitudes and beliefs and
10 behavior, particularly the dynamics of party support,
11 party demonstration.

12 MS. STEWART: Merrill, there's a
13 question.

14 MR. KLINEBERG: You know, you may have
15 said it. I didn't understand it. Where did you get the
16 e-mail addresses for that as well?

17 MR. CLARKE: Let me talk a little about,
18 this is the design. Now, see, the idea -- the idea, of
19 course, for us as political scientists, is the body of
20 data that we gather via the Internet -- we use exactly
21 the same questions, okay. We try to write as close as we
22 could to match these question, question order, question
23 wording, scales, looking at identical to the show cards
24 you would see on the screen as opposed to the show card
25 that would be administered face to face, to really match

1 these things. Then we said, okay, now we have a set of
2 models of political behavior. And we've just published
3 this in a book with Oxford called, POLITICAL CHOICE IN
4 BRITAIN. And what we want to do is go back there, okay.
5 This is based on the 2001 and earlier surveys. Let's
6 take the model, the set of models -- and we have several
7 different rival models, competing models, theoretically
8 in form competing models of voting behavior and let's do
9 this experiment.

10 Let's suppose that all we had in 2005 was
11 the Internet data, the Internet data set, or all you had
12 was the face-to-face survey. What happens in terms of
13 the differences you would draw? Do you make different --
14 because as political scientists, this is typically what
15 we're after. We're not after point estimates as to which
16 party is ahead in the horse race or something like this.
17 We don't use the data for that.

18 And certainly, we don't use it as a
19 record of what the voting shares were. The government is
20 going to tell us that. We've got that. We have got
21 umpteen polls giving us the horse race leading us to the
22 election. We don't need that.

23 What we need is to be able to model
24 things. We're not as interested in the means as we are
25 in the poll variances, right. This is what's driving

1 what we're doing.

2 MS. STEWART: Harold, to get back on that
3 for just a moment, these e-mail addresses were requested
4 from these respondents.

5 MR. KLINEBERG: So it's the face-to-face
6 interviews?

7 MS. STEWART: Okay.

8 MR. KLINEBERG: Do you then ask, "Do you
9 have an e-mail?"

10 MS. STEWART: Yes. That's these people.
11 These people are part of the YouGov Internet pool. So
12 YouGov has a variety of recruitment practices, some of
13 which are proprietary -- actually proprietary to any
14 Internet company, how do you get your respondents. Well,
15 we're not going to tell you all that. We'll tell you
16 some of the methods we used, so -- but these methods
17 culminate in this pool, and that's how they get the
18 e-mail addresses.

19 MR. CLARKE: Now, this one here, this out
20 here was to say, "Hey, one of the interesting things
21 is" -- is -- to use the sort of Knowledge Network's
22 approach, in a sense. You say, hey, let's start off with
23 a probability sample, but can we get high quality data by
24 going back to those guys via the Internet? So we're
25 going to mix the mode, and that's what this has to do

1 with this thing here.

2 But for you guys, you say, okay, how do
3 these guys get their respondents? How did you get these
4 6,000 or 7,000, almost 8,000 people? As I say, there are
5 a variety of ways.

6 But what YouGov does, the methodology
7 they use -- and we can talk about the other firms as
8 well. But in YouGov's case, what they do is they
9 construct a big panel. They advertise. If you get up,
10 you see these banner headlines on various web pages and
11 so forth. You know, "Join the YouGov panel." And they
12 may have a little sort of teaser, "Do you think that
13 Beckham is the world's greatest football player? Vote
14 now." So you click on that and then you give your
15 answer, and you get invited to join the YouGov panel.

16 UNIDENTIFIED SPEAKER: Ahh.

17 MR. CLARKE: And so they collect -- they
18 build these panels. And YouGov terms -- I don't know how
19 many hundred thousand people they have, a lot of them
20 anyway. Polimetrix has, I don't know, a million five or
21 something now. YouGov not quite as big, but a lot.

22 And what YouGov does is to divide these
23 up into a 48-cell demographic grid, okay. And they --
24 you come along and you say to YouGov, "Gee, I would like
25 to do a survey on pet food or whatever. We're launching

1 a new line, and we'd like to get some information. You
2 know, does it look better if we had a Dalmatian on the
3 can or a poodle," you know, or something like this. And
4 so they say, "Okay. That's fine." And they will draw
5 from this 48-cell grid. They will go ahead and draw
6 proportionately a sample, then, you know, that looks like
7 a sample from the panel that will be looking like what
8 the census tells us the British electorate looks like.

9 MR. KLINEBERG: It's a quota sample.

10 MR. CLARKE: That's right.

11 So what's this thing really like? When
12 people ask me like in two words, "What is it?" I say
13 it's a quota sample.

14 Now, it also, though, has a waiting
15 scheme built in because they achieve -- these guys here,
16 this 7,793 represent 52 percent of the people that they
17 contact, okay.

18 And so what they do, of course, is they
19 impose weights after the interviews are over. They look
20 and do post-stratification weighting designs. And they
21 use standard -- or standard demographic characteristics
22 that you would think of in terms of variables like age,
23 gender, region, education and so forth. But they also
24 use the British case newspaper readership. And newspaper
25 readership, it turns out in Britain -- at least,

1 historical. Newspapers are in trouble in Britain.
2 That's another story. But historically, newspapers in
3 Britain are overtly partisan, okay, and people know it.
4 And so if you want to have a political variable, which
5 actually tells you something -- and it's interesting. In
6 a number of analyses we've done adding newspaper
7 readership into the analysis actually has additional
8 predictive power in terms of a number of political
9 attitudes. And YouGov guys know this. And, of course,
10 they have designed -- they actually do a study every
11 year, a big study with face-to-face interviewing to
12 gather data on media consumption patterns. And they use
13 these -- particularly newspaper readership -- as a
14 weighting variable with their samples, okay. And so
15 that's how YouGov does it. So this is what we have --
16 this is what we have here. We'll talk a little later
17 about other companies.

18 MR. KLINEBERG: How do you spell YouGov?

19 MR. CLARKE: Y-o-u-G-o-v.

20 MS. STEWART: Time for the next slide.

21 MR. CLARKE: Time for the next slide.

22 We'll come back. In fact, I have a pretext right now if
23 anybody wants to take it. I have got one we can do.
24 We're going into the field.

25 So you can see this thing here. We're

1 not done with these guys. We are going in the field with
2 them again now. It's not actually two years out, it's
3 like two and a half years out or three years out. We're
4 going back into the field with them this weekend, so if
5 you want to become a YouGov panelist --

6 MS. STEWART: Quick comment on YouGov
7 panels. It's also true for most Internet panels. As
8 respondents stay in for about 50 survey administrations
9 and then they go out because otherwise, you have got
10 respondent fatigue and attrition problems.

11 UNIDENTIFIED SPEAKER: How is YouGov
12 compensating respondents?

13 MS. STEWART: They set up an account on
14 their system. And so every time somebody answers, I
15 think that it's a pound or two pounds now.

16 MR. CLARKE: 50P.

17 MS. STEWART: So they have to build up
18 the account to 50 pounds and then a check or a deposit
19 into their bank account or wherever.

20 MR. CLARKE: Yeah. We actually wanted to
21 give them more because we really wanted to build that up.
22 And they said, "No. Don't do. That will screw them up.
23 They will get conditioned to expect two or three pounds."
24 We give them 50 pound, and we compromise on a pound for
25 our guys.

1 MR. KLINEBERG: So YouGov gets these
2 numbers by inviting people to vote on funny things;
3 right? So they get on the --

4 MR. CLARKE: That's right. They found
5 them like sports, you know, Pop Idle.

6 MR. KLINEBERG: They followed up with a
7 questionnaire to them asking, getting those demographic
8 characteristics?

9 MR. CLARKE: What happens is, when you
10 join the YouGov panel -- say I want you to join. At that
11 point, they collect a series of standard demographics on
12 the person, okay. And so when they have this big panel
13 and this 48-cell grid, they can put you in your cell.
14 They say, "Okay. Here he is. This guy is 50 years old.
15 He lives in Colchester. He has a BA Ox on in," you know,
16 whatever and so that's -- that's how they do it. And so
17 they actually have you, you know, demographically pretty
18 well sorted out, and then they've got your newspaper
19 readership as well. They say, "Well, he studied at
20 Oxford. Do you read a daily newspaper?" You say, "Yes."
21 "Oh, which one?" "Oh, I read The Telegraph." Oh, okay,
22 that's -- that -- or "I read The Daily Mail" or I read
23 "The Scotsman" or whatever, you know.

24 MR. KLINEBERG: Someone has voted on
25 Beckham --

1 MR. CLARKE: That's right.

2 MR. KLINEBERG: And it gets funneled
3 back, "Would you like to join this panel?"

4 MR. CLARKE: That's right. They have
5 a -- yeah. They use a variety of recruitment devices.

6 They also buy lists from other companies.
7 There are companies actually that do nothing but gather
8 e-mail addresses. That's what they do for a living. And
9 then a company like YouGov comes along, and they buy
10 sample from these guys, potential sample from them.

11 UNIDENTIFIED SPEAKER: One more sample
12 question.

13 MR. CLARKE: Sure.

14 UNIDENTIFIED SPEAKER: I mean, in effect,
15 you invited -- YouGov invited 15,000 and got roughly
16 50 percent?

17 MR. CLARKE: That's right. 52 percent.

18 UNIDENTIFIED SPEAKER: Why the 15,000? I
19 mean, you're comparing responses from 3,500 for your
20 info, for your RDD, and 7500 or whatever. And normally
21 if I was going to do a comparison, I'd want 3500 to
22 3500 [sic]. I mean, why would you choose the 15?

23 MR. CLARKE: Well, we have got a variety
24 of concerns, okay. This experiment -- and we'll talk --
25 we weight these things later on. For certain purposes,

1 we do things differently.

2 But the idea is this, the main -- one of
3 the main things we wanted to do with the Internet survey
4 in addition to the experiment, the mode comparison
5 experiment which I described a couple minutes ago, was of
6 course to do the Brady Johnson rolling campaign survey.
7 Now to do that, we would ideally like to have a sample of
8 1,000 people, independent draw every day, okay.

9 But even YouGov, you know, we didn't have
10 the money to do this and, also, they said, "That will
11 exhaust our panel. And, by the way, we're also doing
12 surveys for the BBC and et cetera, et cetera. We're not
13 willing to exhaust our panel on you."

14 So we said, what we would like to have
15 achieved end of about 8,000, okay. You tell us how many
16 people you got to contact to get that because then we're
17 going to divide those up into 30 daily replicates and
18 whatever it was to get an achieved end with an attrition
19 rate of about 200 a day, because then we can construct
20 these moving rolling -- you know, these rolling panel
21 surveys with our panel data, which we're going to put on
22 our website and which we'll go down and talk to the BBC
23 about, which we have a contract with and so on and so
24 forth. So that was the other thing.

25 Now, when you come to do the mode

1 comparison, you say, "Well, gee, I'd like to have exactly
2 the same number of people or what have you." Then, you
3 know, that's -- they say, "Well, we can do that. We can
4 easily achieve that. That's no problem."

5 But if you got in, you can always weight
6 it down. If you don't have in, weighting it up is
7 comforting, but you're going to get smaller staired
8 errors, but you know they're not -- this is not real,
9 okay. So I'd rather weight down than weight up. And the
10 other thing is it's cheap, as I say, so there's no reason
11 not to go with big ends.

12 One of the problems -- now, here we come
13 to the mode comparison. Ideally, we would not worry very
14 much about marginals at all. We're only interested in
15 co-variances ultimately. But nevertheless, it's
16 interesting to take a look and see whether or not the
17 face-to-face survey or the Internet survey, for that
18 matter, or any other kind of survey actually matches the
19 actual shares of the vote in the election. It's going to
20 start there. There are so few things -- in social
21 science, there are so few things, at least in political
22 science, where we've got an objective sort of indicator
23 out there, all right?

24 And one of them, of course, is vote
25 shares after an election. We've got a pretty good idea.

1 Sometimes there may be some debates, as there have been
2 in recent American elections and so on. But
3 nevertheless, we usually have pretty good -- and Britain
4 is really good about this. The British system is so
5 great. We know who won. We know all the shares. It's
6 fantastic. People get together in a room like this, and
7 they count them up and then they make a report at
8 10 o'clock at night. Right, Guy, you have been there
9 when they do this?

10 UNIDENTIFIED SPEAKER: Yeah. Sometimes a
11 little later.

12 MR. CLARKE: Yeah, sometimes a little
13 later. But, you know, it's so easy. It really looks
14 like democracy. It's really amazing.

15 And so here is the result of just
16 comparing the shares, okay. And you can see the Internet
17 bar is the black bar. The gray bar is the actual vote
18 share. The in-person survey was done by NatCen, which is
19 National Center For Social Research, which does a great
20 job historically of this stuff, one of the so-called
21 Mercedes kinds of firms which we use for the
22 face-to-face. And you can see these things are really
23 close, okay.

24 And these are -- these are the data, the
25 weighted data, using the standard YouGov weight and the

1 standard NatCen weight -- with data face to face, of
2 course, with weights attached to them as well.

3 Here's a variable we can't compare, but
4 it's a central variable in theories of electoral choice,
5 party identification. And I would be very leery of any
6 survey -- if there's a big discrepancy between the
7 in-person and the Internet surveys on party ID, I would
8 be worried right from the get-go.

9 And so we see again this is for the pre-
10 and post-election surveys, the precampaign, that is, and
11 post-election Internet survey. Pre/post face-to-face
12 surveys. You can see those shares are really close. And
13 the first two bars in each of the little groups of bars
14 are the preelection comparison and then the last two are
15 the post-election comparison. Strikingly, strikingly
16 similar.

17 Okay. Now, here's a variable we all get
18 wrong, okay. There's a couple that look like we're doing
19 a pretty good job on. Here's one that's really -- here's
20 one that's sort of scandalous in political science in
21 terms of measurement.

22 MS. STEWART: Harold, please.

23 MR. CLARKE: No scandals in political
24 science?

25 MS. STEWART: No.

1 MR. CLARKE: No scandals. Lots of
2 politics, not much science.

3 MS. STEWART: It just gets worse.

4 MR. CLARKE: Okay. Here is what we got.
5 Actual turnout from Britain in 2005 -- actually updated
6 that figure a little bit. Final figure was 61.4, I
7 believe, but that's close enough. Internet, way off,
8 82.9. In-person, though, 71.7. 10 points roughly over
9 the official.

10 This is the standard question. As you
11 may know, over the years, there's been all sorts of
12 debate on how to word this. I'm going to show you the
13 next slide. It doesn't matter how you word it, okay. It
14 doesn't work.

15 This is really cool. Here's a whole
16 bunch -- these numbers here are overreports. These are
17 the overreports in a whole bunch of surveys. The blue
18 ones are -- blue bars are face to face. Yellow bars are
19 RDD. And the red bar up here at the top is the 2005 BES
20 Internet. And you can see, as I say, here is -- these
21 numbers are scandalously far off. They always have been.
22 They still are. There is some debate as to whether it
23 matters or not. We'll come to that later, but you can
24 see they are way off.

25 Look at the 2005 or 2002 ANES, for

1 example, telephone survey, off by nearly 40 points.
2 Everybody in the sample basically said they voted, okay.
3 Similarly, with Dick Johnston's Annenberg study in
4 2002 -- actually, Dick didn't direct that one, but he's
5 back now -- and they have got about 90 percent of the
6 people who say they voted.

7 MR. KLINEBERG: We get this on our
8 surveys. 80 percent say they voted.

9 MR. CLARKE: Yeah. That's right. So
10 everybody gets this way wrong, okay.

11 MS. STEWART: So to be clear, this is the
12 British Election Study, the Canadian Election Study, the
13 American National Election Study --

14 MR. CLARKE: And the NAES study, the
15 Annenberg study.

16 MS. STEWART: -- and then the National
17 Annenberg.

18 MR. CLARKE: I just pulled these out and
19 said, let's look at a few recent ones here, along with --
20 this is a track record for the BES over 11 surveys, off
21 by about 10 percent. We do vote validation, by the way,
22 just as a footnote to this as the NAES has done several
23 times and will do again.

24 If you do that in Britain, you will find
25 you will knock that number down by about 50 percent.

1 50 percent of what you have there is accounted by people
2 misreporting, lying, in other words, as to whether they
3 voted. They're practically all false positives, right,
4 false negatives.

5 MS. STEWART: Which means they said they
6 voted, but they didn't.

7 MR. CLARKE: That's right. That's right.
8 And the rest of it, then, you say, "Gee, then why are we
9 still getting this wrong?" Well, that's giving you some
10 insight into the nature of the sample, okay?

11 UNIDENTIFIED SPEAKER: You validated on
12 both the Internet and the face to face?

13 MR. CLARKE: Yeah. And one of the things
14 we haven't analyzed yet, but we were able to -- and we
15 can with the help of the survey firm, of course,
16 YouGov -- is that we do have -- I'm sitting on a data set
17 which I have not looked at yet. We validated, vote
18 validated, a sample of the Internet respondents. We went
19 to the 128 constituencies that we used as our primary
20 sampling units in the face to face to try to get a
21 comparison. The problem -- we just ran out of money in
22 terms of doing vote validation. Vote validation
23 basically, in the British case, is you have to get
24 graduate students and send them down to London and have
25 them beg with -- actually now with a company called

1 Pickford's which is a private firm but collects
2 government documents for the government -- and say,
3 "Please, you know, give us the following -- you know,
4 here we are from Colchester South or whatever
5 constituency for this particular ward, this particular
6 polling state. Give us that sack -- give us that
7 information." And sometimes -- you know, mostly they do,
8 but it's very labor intensive.

9 MS. STEWART: Sack of paper bells.

10 MR. CLARKE: Well, it's a report. It's
11 actually a report. It's actually done at the
12 constituency level, but you got to sort through it. We
13 just didn't have enough money to do 7,000 of these guys,
14 so we picked 128 -- say, 128 constituencies as we used
15 for the face-to-face survey. But we can. The point is,
16 yes, we can vote validate. And that will be really
17 interesting to see how that works out.

18 UNIDENTIFIED SPEAKER: Do you have a
19 theory explaining the 10 percent difference between the
20 face to face and the Internet?

21 MR. CLARKE: Right now, Allan McCutcheon
22 who some of you may know who is the head of the Gallup
23 Research Center at the University of Nebraska, and one of
24 his colleagues are getting ready to submit a paper where
25 they are actually studying this across a whole bunch of

1 countries, including -- I gave him the British data to
2 work with very early on, and so they have got some models
3 of that, some interesting stuff. That will be -- I think
4 they're going to send that POQ.

5 MS. STEWART: What's his key conjecture?

6 MR. CLARKE: Well, he's got a variety. I
7 haven't actually read the paper, but he's got some -- you
8 know, some of that stuff has to do with who lies and like
9 why and et cetera, et cetera. But I don't know yet.

10 Okay. So what we want, though,
11 basically -- that stuff is sort of mood music, okay.
12 Because as political scientists, what really matters to
13 us is being able to make inferences about parameters in
14 rival model, theoretically directed models of electoral
15 choice. And so we had just, as I said, finished a book
16 called POLITICAL CHOICE IN BRITAIN and we have got a
17 variety of models there.

18 Well, let's just grab those models and
19 let's estimate their parameters with the 2005 data,
20 face-to-face and Internet, and let's impose parameter
21 quality constraints and we do this in a variety of ways.
22 We just did it with -- pooled our data and data
23 interaction effects, very simple methodology.

24 And in our labor vote model, for example,
25 these are standard binomial, multinomial logit models

1 and so forth. And some were regression models for leader
2 images and what have you.

3 Key variables here in what we call the
4 valance politics model of electoral choice, party
5 identification, party best and most important issue,
6 party leader images. Party ID, by the way,
7 conceptualized very much along the lines of theory and
8 Akin and others, not as a social psychological conception
9 out of Michigan.

10 MR. KLINEBERG: What does that second one
11 mean?

12 MR. CLARKE: Party best and most
13 important issue, one of the things that is really
14 striking in Canada, the United States, Britain, and
15 probably in other countries as well is that issues people
16 are concerned with tend to be what we call "valance
17 issues," not positional issues. The press spends all
18 sorts of time worrying about people's stands on same-sex
19 marriage, abortion, issues like this, pro/con
20 positioning, left/right ideologies and so on.

21 But when you ask voters about what's the
22 most important issue, they don't talk about that. They
23 talk about healthcare. They talk about the economy.
24 They talk about high quality education for their
25 children. They talk about security, personal security,

1 crime, terrorism, et cetera, et cetera. These are the
2 things -- and these are all issues upon which everybody
3 has got the same opinion. Everybody wants a healthy
4 economy; right? There may be a few ghouls on Wall Street
5 who are trying to short and so they may not want this.
6 But basically this is not what they -- people want the
7 same thing. They want personal security. They want high
8 quality education for their children. They want high
9 quality healthcare. These are the issues that people
10 talk about.

11 The mix of those issues has changed over
12 time, but as Don Stokes pointed out in his famous '63
13 article, his critique of "Spatial Models and Party
14 Competition," valence issues dominate. You say, okay,
15 when you ask people what are the issues that you're
16 concerned about, they say, "Which party is best able to
17 handle this issue?" All right. And this is really where
18 the political debate typically ends. It's not what to
19 do, but who can do it best, okay. Who can do it best.

20 And so that's -- that's that variable and
21 that's very sort of -- it's a core variable in our
22 valence politics models. But, of course, we test
23 other -- other variables as well; Downsian party issue
24 proximities; special attention to economic evaluations,
25 which are part of the valence politics scheme; and in

1 2005, in Britain, of course, opinions about the Iraq War.
2 Just as in the United States in 2006, in the British
3 election in 2005, there were very widespread conjectures
4 that the -- that people disaffection with the Iraq War
5 would have important negative consequences for the Labour
6 Party in that election, and so that's an interesting
7 variable. And then tactical or what we call strategic
8 voting as well, which has been a popular topic in British
9 political science since the late 1990s.

10 Originally -- of course, the idea being
11 you take your second choice in order to defeat a party
12 that you really don't like. This kind of idea that
13 you're really going to, if you will, maximize your
14 utilities by taking into account party competition at the
15 constituency level rather than just making a sincere
16 vote.

17 Then some controls for demographics,
18 particularly social class in the British context which we
19 have beaten up pretty badly in our research, but
20 nevertheless it's always one you want to take account of
21 particularly, I think, in a mode comparison where there's
22 a lot of discussion about coverage, of course, with the
23 Internet, something we haven't talked about too much yet.
24 But in the British case, somewhere over 70 percent now of
25 people have regular -- had easy Internet access. So,

1 nevertheless, that's not 100 percent and that's
2 demographically skewed, so you want to take a look at the
3 demographics as well.

4 The story is really easy. The story from
5 this -- from this analysis really astounded us. We
6 didn't know what we'd find. We didn't really know what
7 we'd find. But what we found was, both for the turnout
8 model, which I haven't talked about here, but for these
9 party choice models of which I'm just illustrating with a
10 very simple logent analysis of voting labor, voting for
11 another party here is that the inferences you draw are
12 virtually indistinguishable by mode; that mode doesn't
13 matter. It doesn't make any difference.

14 And indeed even at the detail level, as
15 we'll see, you just sort of summarize here -- I did have
16 a slide with all the little coefficients and all the
17 standard errors up there, but there's about 50 of those I
18 can show you. This paper is in Political Analysis by the
19 way. It's published last summer in Political Analysis if
20 you're interested in it, in the Journal of Political
21 Analysis.

22 Anyway, the idea is that there's
23 astounding similarities across modes. So it's not the
24 case that you get a very different view of what matters
25 for voting in Britain and what doesn't matter if I give

1 you the Internet data as opposed to the face-to-face
2 data. You say, gee, the valence politics model dominates
3 no matter how you do it. Social class doesn't matter.
4 All those inferences right down to the level of ranking
5 the variables by their -- the variables and the
6 associated models by their explanatory power turns out to
7 be exactly the same. And the model selection criteria
8 testify exactly the same way.

9 MR. KLINEBERG: What was the AIC?

10 MR. CLARKE: The AIC and the BIC are what
11 we call model selection criteria. Imagine you have rival
12 models, typically rival non-nested models where you can't
13 get from one to the other by simply imposing some sort of
14 parameter constraints on X to get Y or so on. It's not
15 just a question of looking at differences.

16 MR. KLINEBERG: There were some big
17 differences?

18 MR. CLARKE: Well, the big difference --
19 the way to compare these -- let me go back. The way to
20 compare them is across the issue. There sure are some
21 big differences. There are big differences because
22 sample sizes -- the comparisons we want -- the
23 comparisons we want are really not the absolute size of
24 these things, but you want to go -- better models have
25 smaller AICs and smaller BICs, okay. So you have got to

1 rank order.

2 Here is our traditional, in-person
3 surveys, okay. And you've got this tournament of models.
4 And you say, okay, were Butler and Stokes right that
5 everything is class, all else is embellishment and
6 detail? If you have this in-person data, no way. Class
7 explains nothing, okay, and it never explained nearly as
8 much as they claim, but that's another story.

9 So you rank order these competing models.
10 There's class. There's all the demographics. Economic
11 voting. Apologies to Mike Lewis-Beck and Helmut Norpoth.
12 Issue proximity, down, okay. Most important issues,
13 scope, like the valence -- just like the quota valence
14 politics model. Party ID under either a Michigan or a
15 Rochester interpretation. Leader images. And then a
16 composite model, which is the one that actually has
17 significant predictors. You know, if you put them all
18 together, this is the one that has got everything in it
19 that's significant. There's a rank order there.

20 Compare that rank order with this rank
21 order, okay. And you can do it in terms of the pseudo R
22 squared. You can do it in terms of the model selection
23 criteria which -- of course, one thing all they really do
24 is they impose a penalty on you for the richness of the
25 parameterization because, as we all know, if I want to

1 explain everything, I just give you the data back, okay,
2 there is Joe and Susie and Sam and so on. I can explain
3 it, but that's really not terribly interesting,
4 especially when you get older and your memory is not as
5 good. We want parsimonious models. That's all I can
6 carry around anymore.

7 So they're saying, "Hey, yeah. You can
8 beat me with your models, but you have got 10 parameters
9 and I only have two" or something like that. So I am
10 going to discount your explanation in terms of the
11 richness of your parameterization. The AIC and the BIC
12 both do this. If you have a sample size any or greater,
13 the BIC will give you a greater penalty for putting in
14 extra parameters, but we practically always have that.
15 Theoretically, it's argued recently the AIC is actually
16 somewhat better. It doesn't really matter. That's
17 another talk.

18 The key things for us is that this rank
19 order here looks like this rank order there, amazingly
20 so, amazingly so. They say, okay. So if I'm having a
21 debate about what matters for voting in Britain and I
22 give you the one data set and I have another one, we're
23 going to reach the same conclusion. To me, that's a key
24 point for political scientists because this is what we do
25 basically. If we want just marginals, we're going to get

1 them from umpteen different firms now. Everybody is
2 doing horse-race stuff, okay?

3 Here is some more. Very simple, just how
4 good a job did we do in terms of predicting the vote.
5 Now, here is one here. This shows exactly the same
6 things here again, zoop, zoop, same rank order, very
7 similar percentages as well, astoundingly similar.

8 Let me show you the next one. This is
9 one that Bob Erikson suggested to me. When we sent the
10 paper to Political Analysis, we got to revise and
11 resubmit. He says, "Well, Harold, you know, one of the
12 neat things you might do would be to cross-predict."
13 Okay. So you take your model, estimated using the
14 Internet data. You've got your parameters to your model.
15 Take those parameters, put them into your face-to-face
16 data and just try to predict the vote share doing that
17 with that and then vice versa.

18 I said, "Oh, boy, there goes that paper.
19 I'll never be able to do this, you know." So I remember
20 on a Sunday morning I set this thing up, and I went like
21 this (indicating). You know, I hit the computer and then
22 looked back and there's the result.

23 MR. KLINEBERG: Wow.

24 MR. CLARKE: All right. Look at that.

25 So I'm using the Internet parameters -- the parameters

1 generated using Internet data to predict Labour voting
2 with the face-to-face data and vice versa. You get
3 exactly the same result. So on the basis of this and
4 parallel analyses for turnout, plus some stuff on
5 marginals.

6 Now, we did stuff on just like do people
7 like Tony -- one of the things that would be disturbing
8 would be if in our face-to-face survey everybody is like,
9 "Tony Blair, he's a dog. We hate him. We'll give him
10 zero out of 10 on the scale," and then you've got 7 out
11 of 10 as a mean like with the Internet people. You know,
12 that kind of stuff would bug you, you know, or the rank
13 orders like do they like Charles Kennedy better in the
14 Internet than Tony Blair, and you get vice versa when you
15 do face to face, that kind of thing.

16 So we did marginal stuff, too. And there
17 were very, very few significant, significant differences.
18 There were a few. There are a few. But in no two
19 things -- they don't explain hardly knowing the
20 variables. Let me explain. Like less than 1 percent --
21 knowing the mode, you explain less than 1 percent of the
22 variance.

23 Second thing in the variable in
24 question -- excuse me. And the second thing was in no
25 case were the stylized facts changed, okay. In no case

1 would you say, "Oh, gee, in my survey people really liked
2 Tony Blair." I go, "Well, not in mine. They didn't like
3 him in mine." I did Internet and they hated him, and you
4 did face to face and they loved him. It's not like that.

5 There are small differences in some of
6 the variables. They tend to be, I think, scale --
7 there's still lots of things we need to know about how
8 people respond to scales in the different modes and so
9 on. But it looks like all we're doing in general, at
10 most, is shifting the constant a little bit in terms of
11 how we would model this. Of course, that's the inference
12 that's coming very much from the model comparison study,
13 right. That we made shifts of these constants, but we're
14 not moving, we're not moving the covariant structures
15 around and that's the key thing. So that is where we are
16 in terms of what we have done with mode comparison.

17 Now, we have got a paper that we are
18 doing at the -- just to give you some idea of the power
19 the Internet in Britain. If you ever get a National
20 Election Study award, you can be sure that very soon
21 thereafter Hermann Schmitt and Ian McAllister from the --
22 what's CSEC -- the Comparative Study of Electoral Systems
23 will send you an e-mail asking you, begging you, telling
24 you that you must --

25 MS. STEWART: Explaining that.

1 MR. CLARKE: -- improve the full CSEC
2 election study module on your study. Well, this is like
3 a 12-page thing, you know.

4 And so with our surveys, we typically
5 have a self-completion mail-back questionnaire that goes
6 with the face to face, but we also have 40 or 50 members
7 in the user community in Britain who want to put
8 questions on. And we typically use this device for
9 expanding the data that we can gather. And it's the same
10 people, of course, as the major face to face. So they've
11 got all those data as well.

12 So if we do the CSES, put it on the
13 mail-back, that's the end of the mail-back, okay, because
14 you can't ask for a 30-page mail-back. It ain't on. I
15 mean, you can give it to people, but it's not going to
16 come back, okay.

17 And so we said, "What are we going to do
18 about this?" Well, in 2001, we just told McAllister and
19 Scmitt to take a hike. "No, we're not going to do it.
20 Sorry guys."

21 In 2005, we were tempted to do the same
22 thing for the same reasons, however, we said, "No. We
23 can do better than that. We're going to give you a
24 complete Internet survey all your own. Most election,
25 only your questions. Nobody in the world gives them what

1 they want, by the way, but we will. We're going to give
2 you your questions, your entire module, start to finish
3 free. That's 3600 respondents. And then we put a few
4 questions at the end -- at the end, which allowed us then
5 to compare the CSES measure of partisanship with the
6 traditional British Election Study measure and so on. So
7 we did this.

8 So we gave them those data. And then the
9 CSES in their wisdom -- their board met and they said,
10 "Oh, no. We couldn't accept the nonprobability sample."
11 And I said, "Whoa. Wait a minute here. I know how you
12 guys get your data in various countries." And so this is
13 not -- this is not a wise decision.

14 So what they actually did, though, John
15 Curtis and Steve Fisher from -- you know, Steve is at
16 Oxford, and John is at Strathclyde, long affiliated with
17 though Nuffield -- they got some money from the ESRC and
18 did a self-completion, not a mail-back because John does
19 the British social attitude survey. They went into the
20 field in July, though. The election is in May. And they
21 gathered a CSES module by self-completion paper and
22 pencil questionnaire. That means we can compare.

23 So we have another mode comparison we're
24 going to do at the APSA meeting where we compare our
25 Internet CSES with theirs. Now, it's not a fair

1 comparison in the sense that they didn't get into the
2 field until July. Of course, they would have got in
3 earlier if they could, but that's not in the BSA study.

4 UNIDENTIFIED SPEAKER: You still did the
5 CSEC module on the Internet?

6 MR. CLARKE: Yeah. We did it on the
7 Internet. But it's a totally separate survey,
8 freestanding, done exactly the way. You know, in the
9 wildest dreams of Ian McAllister, this is how he would do
10 it, okay. And so we have a paper now which compares
11 those two and again -- and against the BES face to face,
12 okay.

13 And because that's one of the big
14 questions with the SCES because they try to get questions
15 which can be compared across like France, Canada, the
16 United States, Britain, Botswana, everywhere, you know.
17 And invariably, they don't ask the question -- or very
18 often they wouldn't be the same question that you would
19 find in the ANES or the BES survey. And so you could
20 have people saying, "Well, British scholars" -- no, that
21 doesn't matter. CSEC by doing the CSS British data --
22 oh, yeah, it's a totally different model. And so we want
23 for them to compare back to BES wherever possible face to
24 face.

25 The results, of course, look a lot like

1 what I've just showed you. Just like we said. The same
2 like -- it's the same thing.

3 MR. KLINEBERG: Probabilities don't
4 matter --

5 MR. CLARKE: Not for modeling electoral
6 choices, it doesn't matter.

7 MR. KLINEBERG: -- response rates don't
8 matter.

9 MR. CLARKE: This is the inference
10 because this is what we need to know, "Can we model
11 what -- as political scientists, can we do this whole
12 range of activities and do it well with the Internet
13 surveys?" And we say, "Yes. In Britain, at least, you
14 can." And that's one of the points.

15 I gave this talk at Washington State
16 University last year, a version of this talk. And some
17 of you may know Don Dillman, who is a very eminent survey
18 researcher, has done a lot of work on mode comparisons,
19 response rates and so forth. And he was sitting there; I
20 could tell that he was uncomfortable with what I was
21 saying. So finally he asked, "Well, I can't believe
22 this." Well, the data are there. Go ahead and do it.
23 It's a pretty simple analysis. I don't think we made any
24 errors here, but it's all public.

25 And I said, "But the bigger point is that

1 what works in milieu may not work elsewhere," okay. What
2 works in Britain where we have very high, you know,
3 response -- or very high coverage and good response rates
4 through a firm, high quality Internet firm, may not work
5 as well elsewhere. And the other thing is, you know, it
6 may not work even in the United States. It's not just
7 between, you know, a developing country and a developed
8 country. It may just not work as well as in the United
9 States as it works in the case of -- in the British case.

10 I said, "Our fundamental point is this is
11 empirical." It's an empirical thing. Everybody has got
12 vast nonresponse problems now. Do not confuse Kish and
13 Campbell. In others, it's a little aphorism. Do not
14 confuse Kish and Campbell. Kish developed the
15 statistical theoretical basis for doing the ideal face
16 variety anyway you want to do it actually. Mode doesn't
17 matter for Kish.

18 But that is not what we achieve in
19 practice with any mode. We're not anywhere close. We
20 were a -- you could sort of dream about it back in the
21 '50s when you were getting 85 percent, but not now. And
22 so it's really an empirical question. If we're willing
23 to buy the face-to-face results -- and I think there's
24 some reasons sometimes not to buy them -- and we get
25 exactly the same results with our Internet surveys -- and

1 I'm encouraged that we have got a real good future here
2 for these kinds of projects in a particular milieu, okay.
3 And I think you've got to test it. You've got to test
4 it.

5 Now, we have done studies in Canada and
6 the United States as well. I've done a 2006 Canadian
7 National Election Study and also by the Internet with
8 YouGov and a 2006 American congressional election study
9 by the Internet. And, again, the same idea is, do I get
10 the same kind of -- do these data look like they should
11 look like, based on all the other kinds of stuff we have
12 got out there. The answer in both cases is, yes, right
13 on the money. It's the same kinds of inferences. They
14 look just the same. You wouldn't know what the mode was
15 if I didn't tell you, okay. And so that is encouraging.

16 The American -- both of these are
17 reported in a book that we have got called MAKING
18 POLITICAL CHOICES, CANADA AND THE UNITED STATES, coming
19 out with the University of Toronto Press later this year.
20 But it's -- you know, the findings are so standard in a
21 way that you -- you'd never even think about it. Once we
22 got going working with it, we never thought about mode
23 again. Because we've got -- we've been analyzing ANES
24 and CES data, you know, Al Kornberg and I and Tom Scott
25 are working on this between the three of us for a long

1 time. So we didn't think about mode very much. At the
2 beginning, we did, just doing sort of tasks and looking
3 at this and that. We said, "Nah. This looks like
4 election study data. Let's go."

5 Is there anybody here who is a member of
6 the project, the Congressional Election Study project?
7 You know, the one that was run by Doug Rivers and Lynn
8 Vavreck and those people? You know, I think it's
9 actually called the CCES. They have all these acronyms
10 now, but the Cooperative Congressional Election Study
11 project. I haven't seen results from that yet. I've
12 seen some stuff that Lynn has done for a conference we
13 did last year, but I haven't actually seen the voting
14 behavioral models or anything. So I'm looking forward to
15 that.

16 We're doing another study right now
17 with -- with the Polimetrix side of YouGov Polimetrix
18 called the CCAP project. Some of you may be involved
19 with that. That is, again -- it is a cooperative
20 campaign analysis project. This is a six-wave national
21 Internet survey done with the Polimetrix methodology.
22 I'll talk a little more about methodologies as well.

23 But it started in December, and then
24 teams could buy in. You could buy in for one or more
25 subsequent waves. We bought in for all the waves we

1 could buy in for. Very cheap, \$33,000. We're going to
2 have five. We've got January, March, August, October and
3 November waves. And so we've got some stuff. We're
4 really interested in the dynamics of partisanship in the
5 United States, a variety of things about testing the
6 generality of the valence politics model in the face of
7 new issue agenda, which are clearly here with us now.

8 And so we bought into this for 33 --
9 imagine getting -- for \$33,000 getting five-wave national
10 panel with 40,000 respondents or whatever the end is
11 going to be. It's so big you don't even worry about it.
12 It's wild.

13 MR. KLINEBERG: This is Internet?

14 MR. CLARKE: Yeah. This is Internet.
15 It's done Polimetrix -- by YouGov/Polimetrix as they are
16 known now, and that was a really easy decision. We did
17 this as part of our British Election Study comparison.

18 Marianne.

19 MS. STEWART: Next slide. We're into
20 that topic.

21 MR. CLARKE: Mode doesn't matter.
22 Internet survey is the future. Cost effective. Insanely
23 huge N's. Super fast. Cool experiments.

24 Do you want me to talk about the
25 experiments?

1 MS. STEWART: Uh-huh. For 10 minutes,
2 yeah.

3 MR. CLARKE: Yeah. A couple of things.
4 One of the things you're going to be interested in as
5 social scientists in the Internet is the ease with which
6 you can do really sophisticated experiments involving
7 audio and visual stimuli. In the 2005 British Election
8 Study, we started to experiment with this kind of stuff
9 doing something called a feedback to respondent.
10 Feedback to respondent experiment was designed to test
11 one of the fundamental assumptions of Downsian spatial
12 competition theory, namely, the exogeneity of -- you
13 know, taste exogeneity that, as you know, the idea is
14 that voters have preferences in a linear multidimensional
15 issue space. Very simple model. They assess where
16 candidates and parties are. They look at those loss
17 functions and they pick the one that will maximize their
18 utilities. It's a very, very simple model.

19 And we said, "Well, gee, you know, that's
20 one straight -- but there's like 50 years of social
21 psychology that says that's not true; that people do --
22 and indeed we studied political campaigns and a lot of
23 the stuff we studied has to do with people shifting
24 around and do they shift, do they move in result if they
25 get stimuli. So we said, "Okay. Let's go ahead, take a

1 traditional -- two of the traditional dimensions in
2 British sort of Downsian modeling." One is a tax cut
3 thing. You know, cut taxes -- do you want to cut taxes
4 even if it means we're going to reduce services or are
5 you willing to pay more taxes to get more public
6 services, which is the sort of -- that's the left/right
7 one, if you will, although they actually ask left/right
8 using those terms as well in the BES surveys.

9 The second one is punish criminals versus
10 rights of the accused dimension, all right. That's sort
11 of authoritarian/libertarian dimension. We said, okay,
12 we will ask people where they are in these dimensions.
13 We'll ask them where they put the major parties, and
14 we'll just do this in a very standard way. Then at the
15 end of the survey, we're going to come back and we're
16 going to show them a two-dimensional space on their
17 computer and it's going to have them located in this
18 space. "Here you are." Okay. And then we are going to
19 do a variety of treatments for -- you know, we've got
20 this big N, so we can do a whole bunch of different
21 treatments. And the treatments involve presentation of
22 positions of parties and/or leaders and sometimes leaders
23 are identified as, you know, liberal leader or liberal
24 Democratic leader Charles Kennedy or Labour leader Tony
25 Blair, et cetera, and sometimes they're not.

1 And the idea was, do people move in
2 reaction to those stimuli relative to your control group
3 that didn't have any stimuli, we just had the individual
4 out in the space? And the finding, perhaps not
5 surprisingly, at least to me, was that people moved.
6 They did move, okay, significantly. And moreover what
7 matters were partisan cues, not leader cues for this. So
8 it wasn't just that -- you know, there wasn't just a
9 general sort of like, hey, there's another object and we
10 move toward it, which would have been a troubling sort of
11 counterhypothesis. No. It's partisan stuff. It's the
12 partisan stuff that moves them and it is an attraction,
13 not a repulsion thing. You know, there is a possibility
14 of sort of a spider sat down beside him and you move away
15 kind of thing. That's not it. Consistently, people move
16 toward -- toward cues.

17 MR. KLINEBERG: You identify with the
18 Labour Party and you hear a message from a neighbor here,
19 it moves you in that direction?

20 MR. CLARKE: That's right. That's right.
21 So this is just an illustration. But the idea is we
22 gather information, we present it back to the respondent;
23 that provides a basis for the experiment involving this
24 visual cue. Of course, you could do all sorts of other
25 things. You can show them pictures of Tony Blair. You

1 can show them picture of the buses blowing up in -- at
2 the time of 7/7 with the terrorist attacks in London a
3 couple years ago. You can do all sorts of stuff. And so
4 we're experimenting more with that right now.

5 We're going to do a bunch more
6 experiments. We're interested in immigration, for
7 example, and attitudes toward immigrants. Threats --
8 there are various kind of threats. We're doing some work
9 with Jim Gibson, who some of you may know, who worked on
10 tolerance in this country and elsewhere, South Africa and
11 elsewhere. We're going to do an entire British minority
12 election study, which hopefully, if they give us the
13 money, which will involve a bunch of experiments having
14 to do with attitudes toward minority groups, particularly
15 immigrants.

16 So it's all sort of stuff you can do.
17 You can do some of this -- you can do virtually none of
18 this by telephone, okay, or very limited, hardly any of
19 it by telephone. You can do some experiments, but you
20 can't really do very much, okay, historically. Now, once
21 we get like -- with phone records changing, of course,
22 with cell, you know, you get these phones that show
23 things, you can do them face to face by turning the
24 computer around so to speak, okay, but that's really
25 cumbersome. That's really cumbersome. Talk about

1 something, "Oh, here's my computer. Take a look at
2 this." I think that's not going to work very well.

3 You can do it. It's really easy on the
4 Internet. People expect to see pictures. It's just
5 seamless. It's easy, okay. And again, it's cheap.

6 The British findings travel well. That's
7 what I mentioned before, how far is it from Wivenhoe Park
8 to Ann Arbor to Montreal to wherever. There's
9 encouraging findings from what we've done so far.

10 And then we talked a little bit about
11 where we might go in the future; more mode comparison,
12 more experiments. We're going to have a truly huge
13 Internet campaign survey this time. Absolutely humongous
14 compared to what we've done before. At least, 400 people
15 every day, ideally 500. We actually want to have the
16 dream of like the independent study every day during the
17 campaign. That's what we want. We really want to do a
18 little time series analysis or really get even to where
19 we do multilevel modeling in confidence, and it's a story
20 and so we can get it and so we are going to do it.

21 Monthly continuous monitoring survey.
22 Okay. I want to make sure I mention this, okay. Some of
23 you may -- how many of you know of Lupita-Mutz, the TESS
24 project? Are you familiar with TESS? How many of you
25 have actually done stuff with TESS? Have you?

1 UNIDENTIFIED SPEAKER: Yeah.

2 MR. CLARKE: Are they using Internet now
3 exclusively or partially?

4 UNIDENTIFIED SPEAKER: Not when we did
5 it. It was telephone. I mean, the "T" is for telephone.
6 Time sharing. It's time sharing.

7 MR. CLARKE: It's time sharing
8 experiments in the social sciences. I need to talk to
9 Skip about whether they're strict. I know they're partly
10 Internet now.

11 The deal is in Britain with the election
12 study -- again, there's only one team that gets an
13 election study, even with the advantages that we're
14 talking about here this morning with the Internet mode.
15 What we want to do is -- and we are going to do is to
16 develop a TESS-like project called the Continuous
17 Monitoring Survey in Britain. Every month starting in
18 July, we'll be interviewing at least 1,000 people and
19 we'll have a core questionnaire and then a periphery, if
20 you will, where anyone, including people in this room of
21 course, can submit proposals to us which will be reviewed
22 by our editorial, quote/unquote, our advisory board for
23 the study, like an editorial board for a journal. And if
24 the proposal is accepted, we will then go ahead and
25 schedule you for doing your experiment or just your

1 regular question battery. It doesn't have to be an
2 experiment. We will do experiments as well, you know,
3 assuming they're feasible; that we can do it technically.
4 Some things you can't do, are too expensive or can't do
5 at all. But that's going to be part of this as well.

6 You could see like how this expands again
7 the frame, you know, the opportunity. The opportunities
8 become massive for people to be involved. And user
9 community actually do work now on a continuing basis, not
10 every four or five years, because we know a lot of that
11 stuff that's really interesting happens in between
12 elections, all right. That's why you want to do
13 continuous monitoring.

14 And also experiments you may want to do
15 in public opinion, doing them during an election campaign
16 is actually the worst time you could do them. It's
17 absolutely atypical. There are all sorts of stimuli
18 going on that are clouding the picture, so you don't want
19 to do them then. And historically you have no
20 opportunity really to do them any other time or very
21 limited opportunities. You might be working -- maybe
22 this university has got like a survey research center so
23 you could do something here in Houston or in Texas; but
24 to mount a national survey, historically, is impossible.
25 And this stuff is all free, by the way. This is totally

1 free. It's totally free, so you don't have to do
2 anything but send us your proposal.

3 UNIDENTIFIED SPEAKER: This is starting
4 in July 2009?

5 MR. CLARKE: It starts in July of this
6 year. Now, we may actually start the test part of it a
7 little bit later because we have to get our website up
8 and get things organized, but certainly not more than a
9 month or two.

10 MR. KLINEBERG: So you got this from
11 YouGov and it's a --

12 MR. CLARKE: No. This is like the grant
13 for the British Election Study for 2009/2010. It has a
14 face-to-face component. It has the Internet component,
15 which is the campaign survey, plus the continuous
16 monitoring survey.

17 MR. KLINEBERG: So this is Britain?

18 MR. CLARKE: This is in Britain. But
19 there's lots of things that you say -- you know, sort of
20 convenience where you are. Britain is great for what I
21 want to do because I can study this. If it is that way,
22 then we're interested in your proposal.

23 The idea is to work with people. The
24 proposals have to be topically relevant. I mean, there's
25 a range of things that you might want to study that are

1 really outside of our framework, like juvenile
2 delinquency and things. We're not going to go there,
3 okay. Those are good topics. You have got to have
4 another problem.

5 You know, we want stuff basically that
6 have to do with political topics, you know, opinions,
7 attitudes and behaviors about political topics. So
8 topical relevance, all right. Theoretical motivation, of
9 course. You have to tell us like what, how do you -- you
10 know, what are we testing here? What's this about, you
11 know? You didn't just find this written in a telephone
12 booth or something like this. You know, this is actually
13 testing, you know, a proposition that's interesting, is
14 an issue here, a scientific issue.

15 And then third is to say administrative
16 feasibility. There will be limits to what we do. You
17 know, we're providing this gratis. And we'll say to
18 YouGov, "Yeah. These guys want to experiment with
19 pictures of this, that and that," and that's fine. But
20 if they had to do a lot of software development, then it
21 went beyond. We'll let you know right away. The idea is
22 to work with people, too. If we like the idea in general
23 but it's got some hairs on it or what have you and we got
24 to work with it a little bit, we will give you an R&R
25 basically.

1 Yes?

2 UNIDENTIFIED SPEAKER: Well, this is a
3 general question, I guess. I believe the Internet can be
4 used for these kind of, you know, surveys, et cetera,
5 et cetera. Obviously, the scope is very limited, right.
6 As you mentioned, people in Albania, it's not going to
7 work there.

8 MR. CLARKE: Not yet.

9 UNIDENTIFIED SPEAKER: Probably in 30
10 years or so. But my question is, I understand how a
11 probability sample is made. I understand that. However,
12 what is a little bit nebulous is how this company
13 designed the sample. Why this e-mail and not this
14 e-mail, on the one hand?

15 And on the other hand, what about
16 reliability and how -- I mean, probably you got lucky
17 here in 2005 and that's why you have -- everything
18 matches so nicely, right? And as you said at the
19 beginning, if it's too simple, probably it's not right.

20 So, I mean, how can I be confident that
21 if I replicate whatever this Polimetrix or YouGov did,
22 I'm going to get the same thing, since you are using the
23 same people that you had at the beginning? So I mean
24 what's your take on that?

25 MS. STEWART: Do you want to speak to

1 that, about data quality? And I'll talk a little bit
2 about coverage.

3 MR. CLARKE: Okay. Yeah.

4 UNIDENTIFIED SPEAKER: And also
5 minorities, right. I mean, that's --

6 MR. CLARKE: You've got a number of
7 questions there that we can deal with and are interesting
8 to talk about.

9 We have done a lot of -- we have actually
10 been running continuous monitoring surveys in Britain
11 using telephone and more recently Internet mode since
12 1992 actually. I talked Bob Wybrow into letting us do
13 this stuff for free with Gallup at one time before they
14 got bought out by Gallup US. So we've been doing this
15 for a long time.

16 And so we've got these sort of standard
17 models of voting behavior, and we run them cross-mode,
18 okay. We have done it a million times. We have done it
19 more recently with the YouGov -- related YouGov project
20 that we run monthly surveys with, and what's really
21 striking is the extent to which the same models work
22 again and again and again and again and again with
23 different sample, okay, with different sample.

24 And not just regression-type models, but
25 some of the stuff we have done, confirmatory factor

1 analysis models we have done of the structure of public
2 service evaluations and things like this, a variety of
3 topics. So once you leave the world of Leslie Kish --
4 and if I can convince you this morning of one thing, I
5 hope you will believe no matter what mode you use, we're
6 not in that world, okay. Now, it would be great to get
7 back there, but we're not there.

8 MR. KLINEBERG: Nothing will get us back
9 there; right?

10 MR. CLARKE: No. I don't think anything
11 we can think about will get us back there.

12 So that's it. Then, I mean, one option
13 is just to give the game up and say, "We can't do this
14 stuff."

15 The other view, though, is to do
16 empirical work, comparing what we think is the very best
17 surveys done with the gold standard methodologies against
18 what we're doing with, for example, the Internet. And we
19 have done a lot of these comparison studies, and they're
20 very encouraging.

21 MR. KLINEBERG: Now, they are consistent
22 and lots of other studies have been done similarly.

23 MR. CLARKE: Yeah. Of course, also, when
24 you do a traditional survey -- remember, you say, you
25 have repeated random samples -- you have an infinite

1 number of repeated random samples. We will have a
2 population parameter, you know, X bar, which is equal to
3 U or something like this, right? Hey, we don't have. We
4 have got one. We've got one, okay. We always have one.

5 UNIDENTIFIED SPEAKER: And this is only
6 for voting -- in particular for voting issues; right? I
7 mean, all of this is only for voting and that's it?

8 MR. CLARKE: Yeah. This is what I'm
9 saying, one of the things is what works in one milieu --
10 this is Don Dillman's point. He made it largely with
11 regard to using different devices to maximize response
12 rates, what he called his tailored survey method. But I
13 think it applies like a fortiori to the questions about
14 mode and what is really useful for doing the kinds of
15 work we do as political scientists in Great Britain or
16 Canada or the United States may not be the way you want
17 to design surveys to study minority populations living in
18 South Dallas or, for that matter, in the British case.

19 We wanted to do --

20 MR. KLINEBERG: Hispanic immigrants in
21 this city are much the same.

22 MR. CLARKE: Of course. And in Britain.
23 So if we do this Benz project, we get funded for that, we
24 are actually going to do face-to-face interviewing.
25 We'll do some work with the Internet to start learning

1 about what we can do for -- and maybe by telephone as
2 well to see about mode comparisons. But the fundamental
3 methodology will be face-to-face interviewing. Now, it's
4 not going to be a probability sample, okay. It's going
5 to be some kind of snowball kind of sampling thing, I
6 believe, to the extent what we have learned so far.

7 But the costs per interview are so high
8 that we could not have accommodated in the British -- it
9 would have been the tail that like ate the dog. It would
10 be like that. The cost for interviews are absolutely
11 enormous to do this work. So we have to put in for a
12 separate, what they call, response mode bid. We will go
13 to both the ESRC, I'm sure to the NSF here, maybe Pew.
14 You know, we'll try to get some more funds to do this.

15 This thing will be almost as expensive as
16 the entire British Election Study, just to study the
17 minorities, and then we won't even have national
18 coverage. It will be a study probably -- and it'll be
19 London, for sure, and it may well be Birmingham and what
20 have you. We're not sure yet, but that's what you're
21 really facing. And so it's very expensive. It's very
22 expensive, but you have -- if you want to have some
23 reasonable data set, you've got to go a different way.
24 So that's my general message here.

25 MS. STEWART: Your comment points to

1 several things, one of which is coverage that Harold is
2 touching on and you had just mentioned, too. Right now,
3 there are unreachable -- Internet unreachable populations
4 or groups. It could be because they live in poor
5 countries, which the Internet infrastructure is not
6 available or Internet infrastructure is available but
7 it's highly government censored.

8 MR. CLARKE: Yeah. China.

9 MS. STEWART: China, United Arab
10 Emirates, et cetera. It could be because people are
11 members of poor groups that, by definition, they can't
12 afford even the cheapest computer and connectivity.

13 So that's all true, but the other thing
14 that's true is face-to-face is, as we keep saying,
15 prohibitively expensive. Telephone is increasingly less
16 reliable. There's a whole stratum of population you
17 can't reach through telephone surveys. Young people
18 don't have land lines. They often have cell phones.
19 There is no national cell phone directory or sampling
20 frame, so you can't get that. Well, they're one of the
21 most critical groups for a lot of reasons, okay. If
22 you're studying -- we know that voter turnout rates
23 amongst the 18- to 25-year old segment of the population
24 has really dropped. You can't get to them through face
25 to face, right.

1 MR. KLINEBERG: I thought you could get
2 cell phone.

3 MR. CLARKE: Maybe you can now.

4 MR. KLINEBERG: That's just changing.

5 MS. STEWART: Is it?

6 UNIDENTIFIED SPEAKER: Gallup has
7 introduced the cell phone component to its surveys, but
8 I'm not sure what frame they're using.

9 MS. STEWART: It will be interesting to
10 find out.

11 UNIDENTIFIED SPEAKER: Either that or
12 they made a breakthrough.

13 MS. STEWART: It may be that -- I talked
14 to some of my criminology colleagues, Internet surveys
15 for the kinds of things that we're interested in, which
16 have to do with crime recidivism among certain prison
17 populations don't make sense.

18 So the point of story is there some big
19 research question there. There's some big topic of
20 interest. But it may just be that for slice of time,
21 whatever that is, multi modes are going to get you the
22 models that you need to estimate in order to get the
23 answers that you want to try to get, which used to be a
24 bad thing. Oh, no, you can't mix up the modes too much.
25 Well, you know, you're going to have to.

1 MR. CLARKE: Yeah. Our research
2 suggests, at least for the political stuff in Britain, as
3 I say, it doesn't matter. You can. You can, if you want
4 to do the kinds of stuff that political scientists
5 typically do.

6 UNIDENTIFIED SPEAKER: That was actually
7 one of my questions that I was going to ask you, is what
8 have you-all done for the cell phone component in the
9 samples in the U.K. where we actually were able to obtain
10 a sample here -- we're actually survey sampling. You can
11 actually get them now in an RDD format and put them in,
12 you know, add your percentages that you have tried. So
13 we were actually reasonably successful with that.

14 MR. CLARKE: That's great.

15 UNIDENTIFIED SPEAKER: But I was
16 wondering if you-all were using that.

17 MR. CLARKE: We have not, but YouGov is
18 very interested in that kind of stuff. If you talk to
19 Joe Twyman, who is our service director or some of their
20 other people, they are very much in tune with this idea.
21 There are more cell phones in Britain than there are men,
22 women, children, dogs, cats, birds, et cetera.

23 UNIDENTIFIED SPEAKER: Here it's 15 to
24 18 percent that people don't have the actual land lines
25 at home.

1 MR. CLARKE: Yeah. There's nobody under
2 the age of 30 in Britain with a land line, things like
3 this. And so you want to be interested in this
4 technology as a way of doing this stuff, and they are.

5 MS. STEWART: So ideally if you're going
6 multi-mode, you're going to end up with some core set of
7 questions you can ask people, but clearly some of the
8 visuals that are interesting in order to get manual
9 line -- causal mechanisms on Internet you can't do with
10 telephone.

11 UNIDENTIFIED SPEAKER: I suspect, though,
12 with the Internet you are picking up some of the younger
13 folks from the Internet as opposed to the old folks. So
14 you are kind of balancing out the cell phone as opposed
15 to the, I would suspect that, right?

16 MS. STEWART: Yeah. You can use the mode
17 for multimodes as a compensating sample.

18 UNIDENTIFIED SPEAKER: Right.

19 MR. CLARKE: It's interesting. Just look
20 at the demographics, one of the things we did -- and
21 actually we published this as an appendix to the
22 Political Analysis paper. Just look at the unweighted
23 demographics for the face to face and the Internet
24 surveys. You say, oh, gee, what would be think -- where
25 do you think the skew would mainly be with the Internet?

1 Well, people typically say age, young people. It's
2 education. It's education. In Britain, at least, it's
3 education. That was the one big one that jumped up.

4 MR. KLINEBERG: Age plays a role as well.

5 MR. CLARKE: Age plays a role, but age is
6 correlated with education. But it's not -- it's still
7 that correlation is going away because Britain, like the
8 United States, now has a mass of higher education. And
9 they have achieved this, in part, by just taking all the
10 community colleges and universities, which was one way of
11 quickly expanding it. We really expanded higher
12 education.

13 MS. STEWART: Well, Harold, we're doing
14 that in Texas.

15 MR. CLARKE: I know. They do it
16 everywhere. But the point is, it's education and it
17 really jumps out at you. Just scan that table --

18 MR. KLINEBERG: That's a pretty powerful
19 bias.

20 MR. CLARKE: But it's irrelevant. The
21 interesting thing is it's irrelevant to these models.
22 There's always literature on political sophistication, of
23 course. You say, well, gee, that would suspect there
24 wouldn't be -- you know, that we would find differences.
25 And so we have pursued this actually in our book

1 be a curvilinear thing. That people get some
2 information -- the easy information is about leaders. So
3 leader effects go up and then they go down. Then the
4 third one -- we call that the smart money hypothesis.

5 That is, the third hypothesis we call the
6 really smart money hypothesis, which are the guys who
7 have read Lupita and McCubbins and so forth, and they
8 know -- they don't care on diversity and so on. There
9 aren't very many people like this probably. But then
10 they suggest -- I also suggest a quadratic effect, but
11 it's shaped the other way. So that people who are really
12 smart, they just go right to the leaders because they
13 know -- they live in a world of sort of radical
14 uncertainty. We have got to get a cue. We don't know
15 what's going to happen. But this guy, whoever he is,
16 will be here -- or whoever she is, is the person we want
17 in charge. It turns out that the smart money hypothesis
18 works.

19 However, having done all this work and
20 spent months developing these models, the purchase you
21 get in terms of explanation using your AICs and BICs and
22 so on, suggests, yeah, you can go there, but there isn't
23 a lot of reason to do so. You really don't get a
24 stronger model statistically by doing this. You may say,
25 well, yeah, there's some nice insights there and I can

1 talk to my friends at Starbucks about this or whatever,
2 you know. This is fun. I can impress my friends with my
3 mixed logent models. But statistically it doesn't look
4 like there's a lot of reason to go there, okay.

5 MR. KLINEBERG: So response rate does not
6 matter as much as we thought? Self-selection does not
7 create as many bias as we thought?

8 MR. CLARKE: It doesn't matter. No.

9 The other thing about Britain is that the
10 models that work now worked 50 years ago. There's some
11 really interesting stuff there. The stuff that worked --
12 the valence politics model is not only good now. It was
13 good when Butler and Stokes -- when Don Stokes first went
14 to Nuffield in 1963. It works.

15 Party identification. One of the great
16 things about the British study is they got all sorts of
17 panels, okay. And we're able to study the dynamics of
18 party ID. Party ID in Britain has big-time dynamics all
19 the way back to the '60s, okay. It's not that the world
20 used to be the way that Angus Campbell described it in
21 1960. It never was that way in Britain. And it wasn't
22 that -- and, also, McCutcheon and I have done a paper
23 coming hopefully from the NLQ where it shows it was not
24 like that in the States either back in the '50s.

25 The panel data we have suggests that what

1 we call these generalized user stair models, which are a
2 generalization of the converse models which we estimate,
3 those things work just as well with the 1956 to '60s now
4 as they do with the 1992 to 1996 panel. And they suggest
5 the same thing.

6 At the latent variable level, you know,
7 there's lots of movers out there and there always have
8 been, okay. And so it's not just the things -- like one
9 of the things is back in the great days of the gold
10 standard, we got 85 percent response rates and all this
11 stuff was done ever so carefully and so forth and
12 everybody was honest and there's no social desirability
13 biases or anything. Everybody spoke their mind like good
14 midwestern people they were. It was a different world.
15 It doesn't look like it. What strikes me is how similar
16 things are over time.

17 MR. KLINEBERG: So it really wasn't that
18 different.

19 UNIDENTIFIED SPEAKER: A more modest
20 question about, you two have done a lot of work on these
21 models of voting and turnouts. I mean, you know a lot
22 about those things.

23 But what about if I wanted to go to
24 YouGov and do a study on vote intentions, I feel pretty
25 good about doing that because I can read your paper and

1 stuff. What if I want to go there and get data on
2 predicting placement on 11-point scales or something or
3 on political ethics, you know, some things are squishier
4 and we don't know as much about, are you buying into sort
5 of --

6 MR. CLARKE: Well, we've got all that
7 stuff.

8 UNIDENTIFIED SPEAKER: You could do a
9 unique comparison with 2005 data on that?

10 MR. CLARKE: Sure. Sure. I think that's
11 really -- what you want to do -- and you can do a lot
12 with the 2005 data. I welcome you to download the data
13 and go to it -- is to start doing an inventory. Let's
14 start branching out, looking at a variety of different --
15 like interesting -- I think we have worked on party
16 identification in the models in the Internet and the
17 dynamics look a lot like the face to face, but there are
18 other things. There's stuff like on advocacy. There's
19 stuff on political knowledge, measuring political
20 knowledge.

21 We mentioned the education bias. So you
22 say, gee, wow, we have a lot more knowledgeable people.
23 It's the structure. It's not just the means on these
24 variables. But what's the structure like? Does, for
25 example, the internal and external advocacy, do you get

1 the same kind of structure? Do you get two factors? Do
2 you get one?

3 I think there's a really rich sort of set
4 of comparisons that we can do to expand our knowledge.
5 We have lots of things we want to know yet. There's
6 stuff we can do in the lab, too, like psychology kinds of
7 lab experiments in terms of here is how the Internet
8 presents an 11-point scale or, at least, this is how
9 YouGov does it. Here is how Polimetrix does it. And
10 they actually -- I found out by actually now working on
11 the CCAP project with Lynn Vavreck and these people, they
12 do it differently. And I said, "Lynn, I don't think
13 you're right. I think this is wrong. I think this is
14 going to give you some very different data." She said,
15 "Well, I don't have money for a programmer to fix it,
16 Harold. If you want to try to fix it, you can do it." I
17 said, oops.

18 Anyway, there's lots of stuff that we may
19 want to do in lab setting, like just writing lab. Say,
20 hey, go to your friend in the psychology department and
21 say, "Look, let's do some experimentations and see
22 whether it makes a difference within an Internet mode.
23 Would it make a difference if we did it this way or this
24 way or we had the pointer up here is here as opposed to
25 over here and so on." There's a bunch of that stuff that

1 we need to know.

2 Krosnick has done some of this stuff in
3 his work when he was at Ohio State, but there's a lot
4 more to do it yet to learn that kind of stuff. But
5 you're right, there's lots of stuff we can do and you can
6 do it with the 2005.

7 UNIDENTIFIED SPEAKER: You can use that
8 sort of a Rosetta stone?

9 MR. CLARKE: That's right. Get started.

10 UNIDENTIFIED SPEAKER: Just go to the
11 future and just to Internet. But you can feel confident
12 that's -- because I talk to people, you know, in
13 political science and I say, "I've got somebody to do
14 YouGov this summer." They are like, "Oh, geez, I
15 wouldn't" -- people are skeptical. So if you want to
16 write something with this stuff, you have to worry is it
17 worth your time, you know, doing this stuff. But I think
18 I'm okay if I can sort of go back to the study and sort
19 of --

20 MR. CLARKE: Psychoanalysis paper.

21 UNIDENTIFIED SPEAKER: A psychoanalysis
22 paper, that would be good, so...

23 MR. CLARKE: Yeah. And you can, like you
24 said, validate it. This is worth validating. I'm going
25 to develop this project, but to begin with I need to know

1 whether I'm getting equivalent measures. Well, gosh, the
2 2005 BES allows me to look at the efficacy batteries or
3 whatever, and so that gives me confidence and then away
4 you go or, no, it suggests caution, I need to do more
5 work, I need to learn more. Yeah, I think that's exactly
6 the way to do it. Right. Sure.

7 UNIDENTIFIED SPEAKER: Right.

8 MR. CLARKE: Guy?

9 UNIDENTIFIED SPEAKER: I was wondering,
10 sort of around those lines of areas that people have
11 skepticism, I guess, sort of Hawthorne effect where these
12 people are participating in all these surveys --

13 MR. CLARKE: Panel conditioning signs.

14 UNIDENTIFIED SPEAKER: -- on wine or dog
15 food and then a political one, but does that somehow make
16 them different. And I know that's a tricky one to try to
17 measure.

18 MR. KLINEBERG: Certainly different
19 meaning eager to do surveys? That's not a usual trait.

20 MR. CLARKE: Yeah. I mean, there are
21 some people who are professional survey takers as you
22 know. They can make enough money to buy a six pack every
23 week or whatever they want to do.

24 But, no. I think that's right. You
25 know, in general, because we do multi-wave panels in

1 various modes because we're interested in dynamics of
2 attitudes and behavior and so on, panel conditioning is
3 always a question and sort of the main hypothesis is that
4 people try to be consistent in their behavior and so you
5 minimize dynamics really in terms of measurement. But
6 that's a good point.

7 Now, whether that would be greater in the
8 Internet or not -- see, most of the stuff -- if you're a
9 YouGov panelist -- I don't know if you are or not. But
10 if you are -- it's not meant to become one. And what you
11 find is that most of the surveys you get, the vast
12 majority are about soap flakes kinds of surveys, you
13 know, and dog food and stuff like this. They're not
14 about political topics.

15 And you might just -- once in a while,
16 you might turn out to be a person who gets the YouGov
17 telegraph monthly survey and then gets the BES coming in
18 the next day, but that would be very, very rare. Because
19 they try to control that, you know, because they can.
20 When they go out of their cell grid, they're not going to
21 pick you if they just picked you to do the tele survey.
22 They'll pick somebody just like you in your grid. They
23 do that, you know, because they're selecting anyway. So
24 they can choose randomly, but they can still, "oh, not
25 Whit. We just asked him yesterday."

1 MR. KLINEBERG: So they do that, they tag
2 them and say --

3 MR. CLARKE: Well, they look to make sure
4 they're not oversampling these guys. I mean, this I
5 haven't actually done, but this is what they tell me. So
6 that they won't pick you -- if I pick you for the
7 political survey yesterday and your name just happens to
8 come up for the BES the next day, they'll avoid you, you
9 know. Maybe they will miss once in a while, but they try
10 to avoid it.

11 MR. KLINEBERG: Your sense is that this
12 is what the future is going to be?

13 MR. CLARKE: I think it's going to be --
14 it's got to be something other than what it is. There
15 are reasons theoretically, even if face to face was as
16 cheap as it was in 1956, that we have got to do more. It
17 is so slow. Its ability to capture dynamics, you know,
18 like all these models that we have where we just blithely
19 put T, T minus 1, T minus 2 on these repeated measures.
20 And they're like one panel is every two or three years,
21 we're submitting an election panel. And the other guys
22 are interviewing like, you know, a month apart or
23 something. This is not the same. And the psychologists
24 would say, "This is crazy. You can't do this. This is
25 not the same."

1 And so there are reasons we want to do
2 other kinds of work and we want to get outside the lab.
3 And the college sophomores -- you know, bless their
4 hearts, occasionally at least -- you know, they're
5 limited. One thing we know, one thing we know is that if
6 we want to generalize, that it's going to be really hard.
7 You know, the external validity off that is always
8 problematic; and we're in much better shape if we have a
9 national sample that we're running our experiments on,
10 other things being equal. And if these guys are getting
11 it right again and again and again and again, then, you
12 know, this is probably a pretty good way to go.

13 MR. KLINEBERG: Well, let me ask you a
14 personal question. I mean, we have been doing this
15 survey by telephone every year for 27 years. We want to
16 continue this in a comparability. Response rates were in
17 the high 70s in the beginning. They're in the low 40s or
18 high 30s today. The responses seem to be as good as they
19 have ever been. You know, when you compare, for example,
20 ratings of job opportunities with official unemployment
21 rates, the parallelism of those two curves is as great
22 today as it was 30 years or 40 years ago. Shouldn't we
23 just continue what we're doing?

24 MR. CLARKE: Yeah. I was going to say,
25 what's your question? It seems to me --

1 MR. KLINEBERG: Should we combine?

2 MR. CLARKE: Yeah. It seems to me
3 that -- you know, there's always this question about
4 these guys, who are these guys who aren't there? If
5 there aren't many of them, you could sort of pass it off.
6 Maybe you shouldn't -- maybe you shouldn't pass it off
7 because it's easy to imagine a situation where it makes a
8 difference, even if you're getting 85 percent. But,
9 yeah, I don't know.

10 I worry when response rates get that --
11 well, I may worry about the survey enterprise in general.
12 I mean, I don't like the idea that the Internet has got
13 52 percent. I would like a much higher rate. I don't
14 like the idea that our face to face is just 60 percent.

15 MR. KLINEBERG: The Internet is not a
16 probability sample, so that seems to me to be a serious
17 challenge.

18 MR. CLARKE: Well, it is. But imagine --

19 MS. STEWART: It is now, but it won't be
20 10 years from now.

21 MR. KLINEBERG: That's right.

22 MR. CLARKE: No, no. But imagine this.
23 Imagine this. Imagine you've got a list, okay. You've
24 got a list. And you know 70 percent of the people, which
25 is about what you're telling me, 60 to 70 percent, you

1 don't interview, okay. And you have got to make this
2 heroic assumption that the 30 percent that are there are
3 just a random subset of the guys --

4 MR. KLINEBERG: They seem to be.

5 MR. CLARKE: They seem to be. Well,
6 that's great if they are. I mean, it's really good news.

7 MR. KLINEBERG: It's remarkable.

8 MR. CLARKE: I mean, we can have good
9 news in social sciences. It's quite possible.

10 There are two things that we don't like
11 to talk about very much. One is that things can go right
12 as well as go wrong and also that we actually might learn
13 something, we might know something. Is it really
14 possible that we actually know why people vote? You
15 know, it's a radical thought, but it's quite possible.

16 I mean, it's a really serious sort of
17 question on both sides, because the problem is with --
18 you know, you've got this nonprobability sample because
19 people are opting in. But with the traditional modes,
20 they're opting out in vast numbers, you know. And so you
21 say, gee, are those guys who are still left, the soldiers
22 still standing, are they representative of the larger
23 group that you originally started off with when you went
24 over the top? I don't know.

25 MR. KLINEBERG: The demographics from the

1 census is as good today as they did --

2 MR. CLARKE: Yeah. Well, you may be in a
3 very happy situation then. You know, you may well be in
4 a great situation.

5 One of the things I want to do -- I'm
6 going to try to do, you know, I hope we can do with the
7 2009 study is get a little more information about the
8 face to face, like we talk in traditional modes, about
9 the guys who do not participate, okay? Because right
10 now, they don't give us any information about
11 nonrespondents, but they know a fair amount about these
12 guys. We know a lot.

13 We can do this, quote, neighborhood stuff
14 that Ron Johnson, this famous political geographer in
15 Britain. We can learn a lot about the characteristics of
16 the nonrespondents, okay, right down to the level of the
17 streets they're on and so on and so forth. We could --
18 we could actually make some inference -- you know, do
19 some comparisons between the guys who answer and those
20 who don't. And I think that's worth doing because it
21 starts to move towards the -- well, we don't know about
22 their attitudes to be sure and that may be what
23 differentiates them, but we can certainly start to make
24 some moves on this. I think we should. It's well worth
25 doing because it's a really big question.

1 Forget about the Internet. Just deal
2 with your question. Traditional modes, massive
3 nonresponsive, is this a problem? And you say, gee, it's
4 really encouraging. The comparisons I've done and so on,
5 it looks just as good now as it did 30 years ago. That's
6 great.

7 But I'm not sure if that's true again for
8 the political surveys. So political surveys are, by
9 their nature, sort of sensitive beasts. And it may well
10 be that the people who don't -- I'll give you an example,
11 real fast. Let's go back. Quebec, sovereignty
12 referendum, 1995, you've got a very close division of the
13 vote, extremely close. You have got three population
14 groups. You have got the Francophones, you know, the
15 French-Canadian community, traditional community. You
16 have got the English speakers, the Anglophone community.

17 And then you have got a third force. You
18 have got people who are immigrants, the so-called
19 Allophones. Allophones differentially don't answer the
20 survey, okay. A lot of them come from countries where
21 you're asked about whether you support the country or
22 whether they're going to leave the country or you're
23 going to get shot if you give the wrong answer to the
24 question and so they're not answering, but they're
25 voting. There's good evidence that these people are

1 going to vote, and they're going to vote massively
2 pro Canada. They came to Canada because they like
3 Canada. They do not want to move to Quebec and see
4 Quebec leave. That's not what they want. They came to
5 be part of Canada. But the question is how many are
6 going to vote and what the division is going to be.

7 Public opinion pollsters made the
8 division 2 to 1 pro Canada and readjusted their polls
9 that way. See the panels, you know, back of the
10 envelope, whatever.

11 We did it using Gary -- one of these
12 algorithms, like these Ruben kind of algorithms that Gary
13 King popularized in political science.

14 And we developed this really big multiple
15 imputation model that came up like 69 percent or
16 something like his. It's hardly different. So we spent
17 all this time, and it came out just about the same, you
18 know. But we were worried. The idea is you're worried.
19 There are reasons why, with some of these political
20 surveys, where you're worried about this stuff.

21 And again, in this case, we wanted a
22 point estimate. We really wanted to know what the
23 division of the vote was going to be. Typically, as I
24 say, we're not -- I don't care about that, but we did in
25 this case. And you care about some of these things in

1 your surveys. You really want to know what the
2 distribution of opinion is, and then those things that I
3 think this becomes -- you know, I don't know much about
4 that right now. I don't know much about that except
5 empirically, votes look the same, turnout is bad, but
6 it's bad for everything, so...

7 MR. GRANATO: We have time for one more
8 question.

9 MR. KLINEBERG: To follow up on Guy's
10 general theme of concerns about Internet survey, you had
11 mentioned the '92 where the British pollsters got it
12 wrong. And one of the reasons I've heard to account for
13 this is that they -- they use quota sampling, and the
14 census had been done like eight years ago and so the
15 weights that they developed based upon the census weren't
16 in some sense really reflective of the demographic -- the
17 current demographics in '92. And in many ways, what's
18 going on is similar with the Internet surveys. There are
19 48 cells, and then they are drawing proportionately.
20 That's the quota sampling.

21 So then the question is proportional
22 relative to census, about how long ago was the last
23 census and how comfortable do we feel following the last
24 census when it wasn't done prior to 2005? When was the
25 last time it would have been done? By 2009 then. You

1 know, is this a concern? Can we go do data census? Are
2 we stuck with sometime eight-year-old data?

3 MR. GRANATO: The census gives you
4 estimates, right?

5 MR. CLARKE: Yeah. Let me -- you know,
6 they work -- the census in Britain, like the census in
7 the United States, produces updated estimates on key
8 variables. They also run these enormous, which I use in
9 dating myself, called Cadillac kinds of face-to-face
10 surveys with the ONS, the Office of National Statistics,
11 and so on. So they use a variety of data, not just the
12 10-year-old census. These guys are -- they're not
13 Einstein, but they're not Dobbies, you know, and so they
14 try to get updated information.

15 Some of you may be familiar with the '92
16 case in Britain was the one where they all predicted that
17 Labour would win and Labour got all excited and had this
18 huge triumphant rally in Sheffield at the football
19 stadium the week before and everything. And then, of
20 course, the Conservatives won with a very, very small
21 almost sort of a non -- just a barely workable majority
22 in Parliament.

23 And in the wake of that, the public
24 opinion polling -- what do they call the market research
25 society? Yeah -- held an inquiry based in postmortem,

1 like what happened, you know? Who committed the crime?

2 And there a variety of hypotheses, and
3 our friend, Ivor Crewe, wrote this great article called
4 "A Nation of Liars." One of the hypotheses was that
5 people lied; like they didn't want to admit they're
6 Tories. Like you're nasty if you're a Tory and you're
7 mean-spirited and so on. And so one hypothesis is quota
8 samples are bad for the reasons, in part --

9 UNIDENTIFIED SPEAKER: It's not that they
10 were bad, but they were outdated in some sense.

11 MR. CLARKE: Yeah. In this particular
12 context, they gave misleading estimates. We stopped
13 polling too soon. People lie, but it's really about it's
14 not -- it's really not a nice thing to say you're a
15 conservative. So they had a variety of hypotheses.

16 But the effect of this was, as I said, to
17 get them off quota sampling. That really moved them to
18 RDD and then with the occasional, quote/unquote, Cadillac
19 probability sample, which they would run like MORI would
20 run just before an election. And that's how they did it,
21 then.

22 And then starting in around 2000, just
23 before the 2001 election, YouGov started doing this stuff
24 by the Internet. And their first success was with
25 something called the Pop Idol quiz, where they predicted

1 that some guy would win like 52 to 48 or something like
2 this, and everybody else said, "No, no. Beckham is going
3 to win it by a mile," or whoever it was, you know. And,
4 of course, they got it dead on. And people said, "Oh, my
5 god, they got it dead on. This is really something."

6 And it turns out that a lot of people are
7 actually interested in Pop Idol. There's a lot of money
8 to be made by associating yourself with this and running
9 commercials on this and so on. So that got them started.

10 And then they predicted a whole series of
11 elections correctly; the 2001 general election, the
12 London -- not this one that's going on now, but the
13 previous one. The European election, they got right on.
14 A whole bunch of them. So they established this success.

15 They also -- in 2001, they gave us a
16 gratis, just free, "Here we'll run you. Give us your
17 survey and we'll run it for you" and that turned out
18 actually -- we had a paper we published a few years ago,
19 which I didn't pay much attention to at the time. Oh,
20 yeah, it worked out really well. But they have got this
21 string of successes, which is how they have built their
22 reputation. I mean, maybe they're just waiting for a
23 Literary Digest moment. I don't know.

24 UNIDENTIFIED SPEAKER: That's the
25 concern, right, that at some level if they're sampling

1 from these cells and it's based, to the extent to which
2 their sampling is based upon other data, right, in some
3 sense, it's like data. The question is how
4 comfortable -- you know, they got it right last time.
5 Maybe that's --

6 MR. CLARKE: That's always true, as I
7 said before. The colleagues only do one survey --

8 UNIDENTIFIED SPEAKER: It's not true with
9 the RDD. It's always true with the political sample.

10 MR. CLARKE: We know that in the long run
11 if we do these things, we are going to get to the
12 population, but we always only do one. And so we have
13 always got this possibility of a rogue poll. We can
14 always have it wrong.

15 And one of the nice things, of course, is
16 we do a lot -- we do a lot of commercial polls now. So
17 we do a whole bunch. So we have sorted out the rogue.
18 Like, in the last race election campaign, MORI had an
19 obvious rogue. I mean, they were way off on this one and
20 it was quickly identified because everybody else wasn't
21 there, okay. There wasn't the big move to Labour at this
22 point.

23 MR. GRANATO: What about a place like
24 New Hampshire?

25 MR. CLARKE: They can get it wrong for a

1 variety of reasons. You can get it wrong for a variety
2 of reasons. But I'm saying basically our problem with
3 the election study is we're always only going to have one
4 election study, okay. That's it. So if there's a rogue
5 and we know we can draw -- by chance alone, we can draw
6 rogues. We know that. And that's it, folks. There goes
7 your \$8 million and that's what you got. And so that can
8 happen with the very best of probability samples.
9 That's -- that's always with us. That doesn't go away.
10 The thing that we know, though, is that we know it's not
11 going to happen very often.

12 MR. KLINEBERG: With election studies,
13 you have got an absolutely clear objective truth out
14 there that you can compare it to.

15 MR. CLARKE: Simon Jackman has got a
16 great sort of methodology now for pooling the polls.
17 Some of you may have seen his paper yesterday in Election
18 2004. So you can really put that information together
19 really very nicely, as long as you know sample size,
20 so...

21 MR. GRANATO: There will be a reception
22 at 3 o'clock over at the Center For Public Policy. They
23 have an informal discussion about these issues and also
24 about the American Journal of Political Science,
25 questions about that as well. Marianne is the editor of

1 that. So it will be at 3 o'clock today. And thank you
2 very much, Harold and Marianne.

3 (Applause.)

4 (Proceedings concluded.)

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

1 STATE OF TEXAS :

2 COUNTY OF HARRIS :

3

4 I, Dorothy A. Rull, a Certified Shorthand
5 Reporter in and for the State of Texas, do hereby certify
6 that the preceding proceedings were reported by me at the
7 time and date stated on page 1 hereof;

8 I further certify that said proceedings
9 were reported by me in machine shorthand, later reduced
10 to typewritten form under my personal supervision, and
11 the above and foregoing 103 pages constitute a true and
12 correct transcript thereof.

13 GIVEN UNDER MY HAND AND SEAL OF OFFICE on this
14 16th day of May, 2008.

15

16

17

18

19 _____
Dorothy A. Rull
20 C.S.R. No. 3978, Expires 12/31/08
Notary Public, State of Texas
21 Commission Expires: 05/27/2008

22

23 Allied Advanced Reporting, Inc.
Texas CRCB Firm Registration No. 252
1647 Colquitt
24 Houston, Texas 77006
Ph 713.524.6777 Toll Free 1.800.223.9409
25 Fax 713.524.6888
aari@alliedadvancedreporting.com