

TO: The Board of Overseers
FROM: Giovanna Morchio, Maria Sanchez, Santa Traugott
RE: Mode Differences. DK Responses in the 1984 Post-Election Study

In this our first report of a series exploring differences in Mode of interview, we will concentrate on **DK** responses. The 1984 Post Election Survey presents a rare opportunity to compare telephone (T) and face-to-face (FTF) modes of interviewing using the same sample frame, in this case, multi-stage area probability.

Differences in T and FTF results can have many sources.

One has to do with the situation itself. In the T situation there is no personal contact. The interviewer is removed from the respondent. The voice is the only means of communication for both parties. All body language clues are lost. The respondents can keep a much higher degree of anonymity. The interviewer can not see their expressions, their clothes, their houses or neighborhoods. Moreover the respondent has always the possibility of terminating the interview by simply putting down the receiver. In the FTF situation there is personal interaction. The interviewer has been admitted to the respondent's house. The interaction has to conform more to a "guest script." As a matter of fact, many interviewers are invited to have coffee or tea, sometime even a meal. There is a higher degree of small talk and in many cases confidences. The respondent has revealed much more of him/herself, and the interviewer has to reciprocate by not straying too far from a guest behavior.

Another possible source of mode differences can be found in the interviewing staff. The Telephone facility offers a centralized setting. Training is highly standard for each new group of hired interviewers. They are more closely monitored and supervised. Training and supervisory practices in the field are looser. Each supervisor hires, trains and supervises the interviewers in her area. There is no opportunity for "on line" monitoring and immediate feedback. In the 1984 Election study, virtually all of the field interviewers were newly-hired for the pre-election study either to replace terminated interviewers in self-representing areas, or in order to staff the SRC 1980 National Sample primary areas where SRC had never interviewed before. In contrast, the telephone interviewers were winding down from a very busy year in which the demands for interviewing had run extremely high. So the two staffs of interviewers differed significantly with regard to interviewing experience, training, and intensity of supervision.

Yet other sources of mode differences are slight changes in question wordings, the use or non-use of visual prompts and context (the telephone interview is shorter). Other memos will discuss those.

The 1984 NES Study

All 1984 NES respondents were initially interviewed face-to-face during September-October of the election year. The study design called for a second (post-election) interview with all pre-election study participants but, for a number of reasons, the decision was made to reinterview half of the respondents in the field and the other half from the Telephone Interviewing Facility at ISR. Accordingly, near the end of the pre-election study, the sample was divided into random halves. The face-to-face (FTF) half-sample was reinterviewed in November-December by SRC interviewers who had worked on the pre-election study. The telephone (TEL) half-sample was reinterviewed during the months of November and December by SRC telephone interviewers.

Some of the pre-election respondents scheduled for a telephone reinterview were interviewed in person at the time of the post-election study. Those handled this way included persons who had refused to disclose their telephone numbers when first interviewed, those living in households without telephones, and all persons identified by the pre-election interviewer as unlikely to tolerate a telephone reinterview for a variety of reasons (too old, poor hearing, ill-health, etc.). (The pre-election coversheet included a question addressed to interviewers to help us identify these special cases.) Non-telephone cases and all cases flagged by interviewers as candidates for a FTF reinterview are excluded from the analysis reported below, without regard to the half-sample into which the cases were subsequently randomized. This procedure results in purified half-samples (FTF/TEL) which, statistically speaking, have comparable demographic characteristics with regard to age, education, income, sex, race, marital status, employment status, region of the country, and population density of the area of residence.

Put another way, the differences that emerged while comparing the two half-samples cannot be attributed to biases in the demographic composition of either sample. Nor are differences explained by differential treatment of respondents on the part of interviewers who conducted the pre-election study. At the time this study was taking place,

interviewers had no knowledge of the half-sample assignments for individual study participants. Finally, differences in the data collection periods between the FTF and the TEL samples cannot account for the results obtained since the field periods were identical for both studies.

ANALYSES

A systematic difference was detected early on in comparing post-election results obtained from the two modes. The FTF sample had a **higher** overall incidence of **DK answers** over a large number of variables. Our analyses will deal with three content areas characterized both by yielding a large percentage of DK responses and strikingly large differences in their distribution by mode: Reagan and Mondale Traits, Performance Evaluation of candidates or institutions, and Political Information Items.

TRAITS

The Reagan/Mondale traits battery asks the respondent if a series of sixteen adjectives fit their impression of the two presidential candidates. The battery was asked both in the Pre and Post interviews, albeit with different scales. The scales were: EXTREMELY WELL, QUITE WELL, NOT TOO WELL, NOT AT ALL WELL in the Pre; and A GREAT DEAL, SOMEWHAT, CAN'T DECIDE, A LITTLE, NOT AT ALL in the Post. The CAN'T DECIDE alternative in the Post, was not read to the respondent, and it was actually probed as a DK. We will be lumping together the CAN'T DECIDE and DK categories for the purpose of this analysis.

We have chosen this battery because it shows a higher than average percentage of DKs; the number of questions involved gives more stability to our estimates; by their nature they require an evaluation of two candidates of different visibility, and they offer the possibility of two measures in time.

Table 1

Distribution Of DK-Can't Decide Responses For Reagan And Mondale
By Mode

	Post-Election		Pre-Election	
	Reagan	Mondale	Reagan	Mondale
	T/ FTF %DIF	T/FTF %DIF	T/FTF %DIF	T/FTF %DIF
Hard-working	1.3	4.6	0	.8
Decent	1.7	4.5	-.1	-.2
Compassionate	2.5	8.8	.7	1.9
Respect	1.1	6.2	-.1	0
Intelligent	1.4	4.5	-.2	.9
Moral	4.7	9.3	-1.1	3
Kind	4.4	9	-1	1.3
Inspiring	3.2	6.8	1.2	.4
Knowledgeable	2.4	4.3	0	1.1
Good Example	2.3	8.3	-.5	-.1
Cares	4.3	9.1	-.3	2.5
Leadership	2.6	8.3	.2	1.5
Understanding	4.4	7.2	-.2	2.2
Fair	3.6	9.1	-.4	2.4
In touch	5.4	7.9	0	2.1
Religious	14.3	24.2	-3.1	2
Average Differ	3.7	8.3	.37	1.36

Table 2

Pre-Post Election DK, Can't Decide Averages To Traits Questions
For Reagan And Mondale By Mode

	<u>Post-Election</u>			<u>Pre-Election</u>		
	T	FTF	DIF	T	FTF	DIF
Reagan Av DK	.67%	4.4%	3.72%	2.59%	2.90%	.37%
Mondale Av DK	.97	9.29	8.25	8.25	9.55	1.3
Av Cand Differ	.30	4.76		6.1	6.65	

In order to discount sample differences we run a parallel analysis of traits items in the Pre, using the same T and FTF sub-samples that we used in the Post. Table 1 shows the total DK % for Reagan and Mondale in the Pre, and the separate figure for each candidate in the T and FTF mode in the Post. Table 2 shows the average DK % in each sub-sample in both surveys. The differences between sub-samples in the Pre are negligible: .37% for Reagan and 1.3% for Mondale. The same differences become substantial in the Post 3.72% and 8.25% pointing to a clear mode effect.

The higher visibility of four years as president, translated in respondents having an easier time judging Reagan than Mondale. Both the Pre, and Post FTF batteries show a lower average of DK responses for the president. It is puzzling though that the DK average for Reagan went up in the Post, while the average for Mondale stayed basically the same, ruling out a scale effect. A more disturbing piece of information is the almost negligible difference in DKs average among the candidates in the T. A degree of difference, like the one found in the FTF situation, seems more in tune with reality. The presidency provided Reagan with constant visibility, showed him performing on a day to day basis the duties of the office; presumably making it easier for a respondent to rate him in a number of areas. The T situation seems to be producing an indiscriminate reduction of DKs to trait questions, at odd with common sense

expectations represented by the FTF results.

Note that the higher levels of DKs, and the largest mode differences are on the religion and morality areas, where persons are normally thought not to express opinions. One wonders if the very social situation of FTF interviewing operates to reinforce this reluctance.

APPROVAL OF POLITICAL FIGURES AND INSTITUTIONS

In this section we will look at four explicit approval items:

Do you approve or disapprove of the way Ronald Reagan is handling his job as President?

Do you approve or disapprove of the way Ronald Reagan is handling the balancing of the budget?

In general, do you approve or disapprove of the way the U.S. Congress has been handling its job?

In general, do you approve or disapprove of the way Representative (NAME #33 or 34) has been handling (his/her) job?

We will also include the "like" items, since they imply an evaluation:

Was there anything in particular that you liked about (NAME #31, 33 or 35) the Democratic candidate for the U.S. House of Representatives?

Was there anything in particular that you liked about (NAME #32, 34, or 36), the Republican candidate for the U.S. House of Representatives?

Table 3

Approval Items, %DKs Distribution By Mode

	T	FTF	DIF	DECREASE RATE
Reagan as president	.4%	2.6%	2.2%	5.5
Reagan budget	2.4	7.6	5.2	2.16
Congress	5.2	13.3	8.1	1.55
Incumbent	9.0	23.5	14.5	1.61
Like Dem Candidate	16	26.2	10.2	.63
Like Rep Candidate	18.3	28.1	9.8	.53

Decrease Rate was calculated by subtracting the T% from the FTF% and dividing by the T%.

Beside the familiar pattern, higher DK% in FTF half sample, we can see a tendency toward increased use of DK as the subject being evaluated is less familiar. The evaluation of Reagan as president generates the lower level of DKs, while the evaluation of the House of Representatives candidates generates the highest. But when we look at the decrease rate of DKs going from FTF to T, the inverse is true. The more familiar, the subject being evaluated, the steeper the decrease rate. Since the six items under consideration have a dichotomous answer: APPROVE, DISAPPROVE for the "Approval" items; YES, NO, for the "Likes" items. We looked to see if the DK decrease translated in a gain for the "positive" or the "negative" alternatives. The results are mixed. In the evaluation of Reagan as president, and his handling of the budget, the DK decrease results in an almost parallel increase of the "disapprove" alternative. The reverse is true for the approval of Congress and the Incumbent. While for the Like of the democratic and republican candidates the NO alternative shows the higher gain. We can not discern a positive or negative bias as a result of DKs conversion.

Political Information

Political information was measured by four factual items:

J2. Do you happen to know which party had the most members in the House of Representatives in Washington before the election (this/last) month?

J3. As a result of the election (this/last) month, which party will now have the most members in the House of Representatives?

J4. Do you happen to know which party had the most members in the U.S. Senate before the election (this/last) month? (IF NECESSARY: Which one?)

J5. As a result of the election (this/last) month, which party will now have the most members in the U.S. Senate?

Table 4

Political Information Items By Mode

	<u>Republican</u>		<u>Democrats</u>		<u>DK</u>		<u>Na</u>	
	T	FTF	T	FTF	T	FTF	T	FTF
House majity pty before election	17%	9%	63%	56%	20%	35%	.7%	.1%
House majity pty after election	26	19	59	51	14	29	.7	.4
Senate majty pty before election	38	29	35	24	27	47	1.1	.2
Senate majty pty after election	47	35	34	23	18	41	1	.2

The puzzling initial conclusion appeared to be that, for some unknown reason, the TEL respondents were more informed about politics than their counterparts reinterviewed in the field. However, this conclusion was not supported by other survey results and, in particular, by the answers to another post-election factual item in which respondents were asked to name the House candidates running in their districts and to provide each politician's partisan affiliation. For this question, about the same proportion of respondents (45% TEL, 46% FTF) indicated they could give the information requested, but FTF respondents had a higher overall number of correct mentions than TEL respondents.

The pervasive nature of the unbalanced use of DK in the T and FTF samples, and the magnitude of the differences in the Evaluation, Traits and Political Information areas, required an explanation.

Having ruled out earlier other extraneous sources of variation having to do with half-sample demographic composition, differential treatment of pre-election respondents and noncomparable field periods, we looked at the interviewers themselves as a possible explanation. As stated in our introduction, the interviewing staff were different in terms of their experience, training and supervision. It is a well-known fact that less-experienced interviewers have a tendency to accept DK answers without probing.

Goves and Kahn, in their major work on differences between personal and telephone interview, found that telephone interviewing resulted in more DKs, probably as a result of less training in probing. At the ISR, probably as a consequence of his work, a major effort was made to train telephone interviewers in probing.

So, not only were telephone interviewers more experienced, there was a pronounced effort to reduce DKs by probing. The centralized setting of the Telephone Facility operates to reinforced the emphasis on probing, since supervisors review a higher percentage of questionnaires and monitor some of the actual interviewing.

We found some support for the probing explanation in a cursory examination of about 100 questionnaires. We looked at the information items, and we found that "every" T DKs had been probed, while the FTF questionnaires showed very little evidence of probing. Nevertheless, to look at T as a "probe situation" and FTF as a "non probe" is far from been a perfect experiment. Obviously some level of probing was going on in the FTF situation as well.

DK Probing - Does it elicit mere guesses or valid information?

Attitudinal and opinion survey questions do not usually have right or wrong answers. These types of questions offer less desirable testing ground to explore the effects of DK probing than factual information items. Fortunately the NES political information quartet displaying large differences in the proportion of DKs across samples consists of questions with 'true' or 'false' answers. These questions are especially useful to analyze the quality of the information elicited from respondents after probing DK answers. (See table 3)

The four items can be combined to produce a new variable indexing the number of correct mentions each respondent gave out of a possible total of four. Since each question presents a dichotomous choice (T or F), one can argue that by chance alone the combination of 2T and 2F items would be the most frequent occurrence when someone answers by merely guessing for each of the four questions. To be precise, the probability of this event is .375. Next in line as outcomes predicted by chance alone are the combinations '3T-1F' and '1T-3F'. Each of these outcomes has an expected probability of 0.25. The most infrequent outcomes obtained by guessing are represented by the combinations '4T-0F' and '0T-4F', with an expected probability of .063 each.

If DK probing elicited real information among TEL respondents, then a couple of things would follow. First, the joint distribution represented by the index would show a higher proportion of TEL respondents than FTF respondents in cell outcomes corresponding to combinations on the high end of the political information index (all correct mentions; 3 out of 4 items correct). Conversely, proportionately fewer TEL than FTF respondents would be represented in the low information end of the index (all items incorrect).

On the other hand, if DK probing elicited mere guesses, the sample index proportions for the TEL and FTF samples would be fairly comparable for most cells except two: the one corresponding to the most favored outcome predicted by chance (2 correct mentions out of 4 possible), and the cell corresponding to zero correct mentions represented by 4 DKs. Further, these two cells should compensate each other across the two samples. The TEL index distribution should show a higher proportion of

respondents than the FTF distribution in the outcome category "2 out of 4" where most of the telephone "guessers" would land, and the difference in these proportions should be mirrored in a comparable disproportionate bulging in the FTF proportion for the cell corresponding to the 4 DK answers which field interviewers did not probe systematically. Beyond this, we would predict proportionately more TEL than FTF respondents in the next most likely outcome categories, the cells corresponding to the patterns '3T-1F' and '1T-3F'. Finally, small differences in the sample proportions would be predicted for the least likely outcomes, '4T-0F' and '0T-1F'.

The results presented below support the hypothesis that DK probing elicits "mere guesses", at least as far as these factual items are concerned. The empirical results displayed in the table validate the predictions derived from theoretical considerations. The relative probabilities of various index outcomes in the FTF and TEL samples match results predicted under conditions involving answer guessing on the part of uninformed TEL respondents.

Information Index by Mode

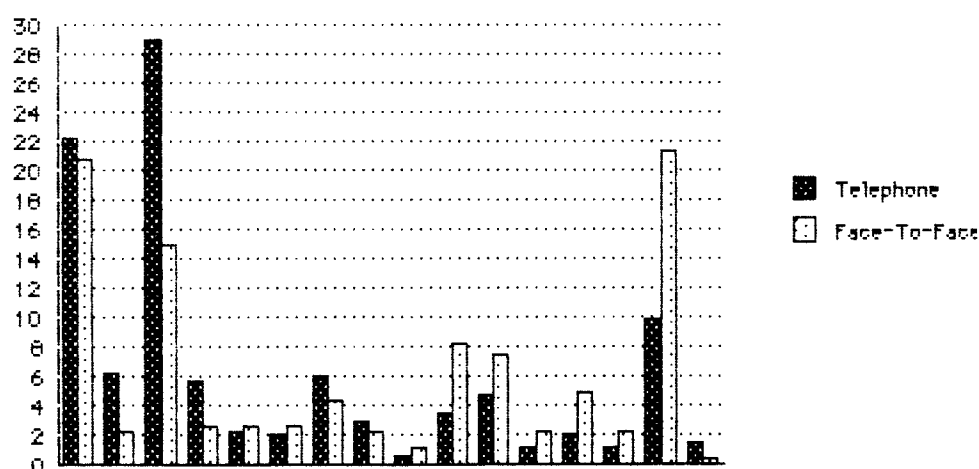
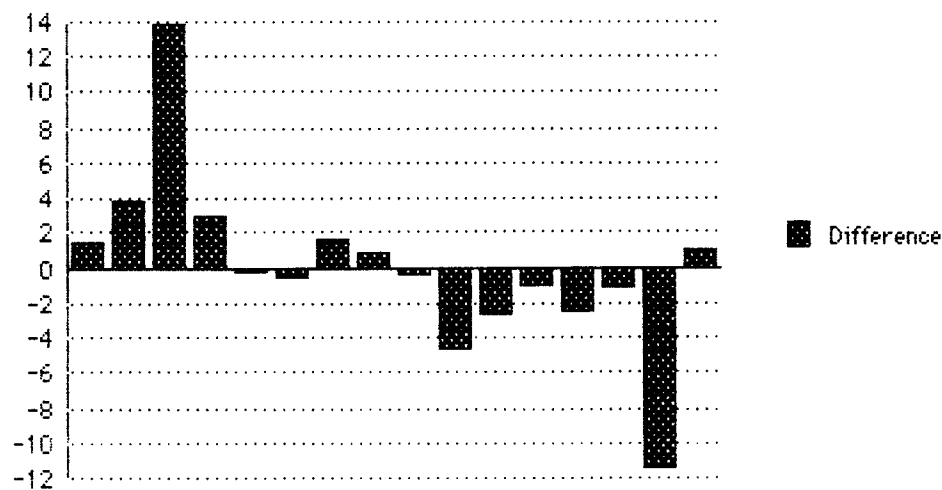


Table 5

INDEX CATEGORIES BASED ON 4 ITEMS.

CODE	NUMBER CORRECT	NUMBER INCORR	NUMBER DKS	T N	T %	FTF N	FTF %
1	4	0	0	178	22.2	170	20.8
2	3	1	0	49	6.1	19	2.3
3	2	2	0	232	28.9	123	15
4	1	3	0	44	5.5	21	2.6
5	0	4	0	18	2.2	21	2.6
6	3	0	1	15	1.9	21	2.6
7	2	1	1	48	6	36	4.4
8	1	2	1	24	3	18	2.2
9	0	3	1	4	.5	8	1
10	2	0	2	28	3.5	67	8.2
11	1	1	2	38	4.7	61	7.5
12	0	2	2	9	1.1	18	2.2
13	1	0	3	17	2.1	39	4.8
14	0	1	3	9	1.1	19	2.3
15	0	0	4	78	9.7	174	21.3
99	NOT ASCERTAINED			11	1.4	3	.4

Difference in Information Index across Modes



It seems clear that probing elicits "guessing" behavior. The next question would be: Is "guessing" equally distributed in our sample, or some groups are more sensitive to probing than others? If the second is the case, the situation would be more serious. We would be not only introducing a large level of noise, but the noise would not be randomly distributed, distorting our distributions. We will explore the effect of probing by systematically examining the two mirror Information Index categories most likely to be affected by guessing: four DKs and two correct-two incorrect responses to the information items. We will look at their distribution in the T and FTF sample by sex, race, education, age, income, political interest, and vote validation. If the effect of probing is constant across the board, we would expect to find for each demographic group, the DK category (four DKs) decreasing from FTF to T by approximately the same rate as the "guess" category would be increasing.

If the DK decrease rate turns out to be significantly higher than the "guess" increase rate, it would mean that the data elicited by probing is of better quality than we would have expected by mere guessing. These would probably be the kind of DKs that really mean "I am not sure." Conversely, if the DKs decrease rate is significantly lower than the increase rate for the "guess" category, it would mean that the level of increase in guessing is higher than expected as product of probing DK responses.

Sex

Table 6

DKs And Guess Distribution By Sex And Mode

	DKs			Guess		
	T	FTF	Dec Rate	T	FTF	Inc Rate
Men	6%	15%	1.5	25%	14%	.78
Women	13%	27%	1.07	33	16	1.06

Decrease Rate was calculated by subtracting the T% from the FTF% and dividing by the T%

Increase Rate was calculated by subtracting the FTF% from the T% and dividing by the FTF%

Women are more likely than men to give DK responses in both modes. Their DKs decrease rate is almost identical to their "guess" increase rate over the phone. It looks like any gain in substantive response was produced by guessing. Men start out with a much lower DKs percentage. The DKs decrease rate for them is much lower than their "guess" increase rate. So the quality of the information produced by probing is better than expected by guessing.

Race

Table 7

DKs And Guess Distribution By Race And Mode

	DKS			Guess		
	T	FTF	Dec Rate	T	FTF	Inc Rate
Whites	9%	21%	1.33	30%	15%	1
Blacks	20	28	.43	22	11	1

Again, blacks are much more likely to give DK responses than whites. For them the DKs rate of decrease is lower than the comparable "guess" increase rate. The increase in guessing is higher than expected by probing.

The small number of black respondents, 66 in the T, and 73 in FTF situation, make this estimate rather unstable. For whites, the DKs decrease rate is much higher than for blacks, and also higher than the "guess" increase rate. As for males, the quality of data elicited by probing is better than expected by guessing by a factor of .33.

Education

Table 8

DKs And Guess Distribution By Education And Mode

	DKS			Guess		
	T	FTF	Dec Rate	T	FTF	Inc Rate
Less High School D	23%	33%	.43	17%	12%	.41
High School Dipl	10	27	1.7	34	14	1.42
Some College	5	12	1.4	30	18	.66

The level of DK responses decreases as the level of education increases in both modes. As always the T shows a much lower DK 's level. The rate of decrease from FF to T is different for each education group. It ranges from 1.7 for high school diploma, no further academic training, to 1.4 for some college and above, to .43 for less than high school diploma. The "guess" increase rate for the two lower education groups is almost identical to the DKs decrease rate. The substantive gain for these groups seems due entirely to guessing. College people show an entirely different pattern; their DKs level is the lower of any group, their decrease rate is the steepest and the "guess" increase rate is much lower. The data recovered by probing are much better, by a factor of 1.04, than expected by guessing. This pattern seems to substantiate the notion of two kinds of DKs at work: "no knowledge", and "I know, but I am not sure." The DKs given by the lower education groups seem to belong to the first kind, those given by the college respondent to the second.

Age

Table 9

DKs And Guess Distribution By Age And Mode

	DKs			Guess		
	T	FTF	Dec Rate	T	FTF	Inc Rate
17 to 25 years	13%	39%	2	31%	6%	4.16
26 to 35 years	9	19	1.11	33	16	1.06
36 to 50 years	7	20	1.85	30	20	.50
51 to 65 years	7	12	.71	28	18	.55
65 and above	17	24	.41	20	8	1.5

The different age group show very different levels of DK responses. The youngest and the oldest are the highest in that order. The different

groups vary greatly in their reaction to probing. The 26 to 35, and the 51 to 65 seem to react by guessing. Their DKs decrease rate is very symillar to their "guess" increase rate. The 36 to 50 years old give much better quality data than expected by guessing. The 65 and older show and increase of the "guess" category higher than expected by the reduction of DKs. The 17 to 25 present a steep DKs decrease rate, 2; and a "guess" increase rate of 4.16. I am at loss to explain this discrepancy.

Income

Table 10

DKs And Guess Distribution By Income and Mode

	DKs			Guess		
	T	FTF	Dec Rate	T	FTF	Inc Rate
0 - 4999	34%	44%	.29	20%	4%	4
5000 - 9999	12	20	.66	23	13	.76
10000 - 14999	12	21	.75	31	13	1.38
15000 - 19999	8	23	1.87	33	11	2
20000 - 29999	9	19	1.11	34	17	1
30000 and over	6	13	1.16	30	18	.66

The relationship between income and four DK responses is not perfectly linear. They reach their highest level in the lowest income group and tend to decrease as we move toward the highest. Again the result of probing, measured by the DKs decrease rate and the "guess" increase rate, varies from group to group. Respondents in the \$5000-\$29999 range (with the exception of \$10000-\$14999) seem to react by guessing. Both "rates" are reasonable close. The \$5000-\$10000 shows a "guess" increase higher than expected as effect of probing. The \$30000 and over show a "guess" increase rate lower than expected by guessing. The \$0-\$4999 present a

widely high "guess" increase rate, the number of cases in this category is probably too small (45 T, 46 FTF) for a reliable estimate.

Political Interest

Table 11

DKs And Guess Distribution By Political Interest And Mode

	DKs			Guess		
	T	FTF	Dec Rate	T	FTF	Inc Rate
Most Interested	2%	9%	3.5	26%	15%	.73
Somewhat Inter	9	16	.77	30	19	.57
Not and Hardly	18	36	1	32	11	1.9

As we could have expected, the higher the level of interest expressed by the respondent the lower the level of DK answers. In keeping with the previous analyses, the Most Interested category shows much better data than expected as result of probing. The Somewhat Interested seem to be guessing and the Least Interested show a "guess" increase rate higher than expected.

Vote Validation

Table 12

DKs And Guess Distribution By Vote Validation And Mode

	DKs			Guess		
	T	FTF	Dec Rate	T	FTF	Inc Rate
Voted	6%	15%	1.5	28%	19%	.55
Said vot,didn't	3	8	1.66	40	19	1.1
Not val,didn't vote	27	46	.7	22	4	4.5
Record not found	13	36	1.76	28	5	4.6
Ambiguous	9	18	1	43	15	1.86

Predictably, respondents who said they voted, and were validated as having voted, give better data than expected when probed. Respondents who reported "didn't vote" and were not validated, show a "guess" increase rate much higher than expected. The other group cells have too small a number of cases to give stable estimates.

The data presented in the previous tables show an expected distribution pattern of DK responses. It is not surprising to find a higher proportion of DKs for women, blacks, low education, low income people, the very young, the very old, and the least interested in politics. The groups traditionally associated with high interest, hence high information in politics, (males, whites, high education, high income, etc.) behave in a predictable manner. They seem to answer DK when they are not sure about the accuracy of their information. As a result, the data they provide when probed is better than expected if they were guessing. The groups generally associated with low interest in politics, react to probing with two different response patterns. One is the typical of guessing; the "guess" category increase in the same proportion as the DKs decreases. DKs in this case seem to reflect a genuine lack of information. This group comprises: women, low education, 26-35 and 51 to 65 years old, respondents in the

\$5000-\$29999 brackett and people Somewhat Interested in Politics. Yet, another low interest, "alienated" set of respondents: the blacks, the very young, the very old, the least interested, the non voters, show in the T an increase in the "guess" category much higher than the corresponding DKs decrease. The pattern shown by the "alienated" group points toward a factor other than probing, generating the unexpectedly high level of the "guess" category.

CONCLUSION

Telephone respondents in the Post display an overall lower level of DKs, as shown by the comparison of T and FTF frequencies for all questions.

Different questions generate very different level of DKs. The highest level is produced by factual items with a true or false answer. The evidence generated by the analyses on the Information Index by different demographic groups, suggests the notion of two kind of DKs in operation. High education-high political interest groups seem to mean "I am not sure" by their DKs. These groups give a lower level of DKs to start with. Under systematic probing they show a much higher reduction rate of DK responses than their low interest counterpart. They produce better quality data than expected if they had guessed. Most of the low interest groups (women, low education, middle income respondents) present a pattern characterized by a high level of DK responses, lower DKs decrease rate when probed followed by a parallel increase of "guess" responses. They seem more reluctant to give a substantive response, even when probed, and their DKs seem to reflect a real lack of information. Several low interest groups (the very young, the very old, the very poor, blacks) start out with a very high DK level, when probed the DKs decrease at various rates, depending on the group involved; and the "guess" category increases at a rate much larger than expected from the DKs decrease. This unexpetedly large effect suggests some other factor playing a part.

Next in line are questions implying some sort of evaluation of a candidate or institution. Are these DKs the product of ignorance on the subject at hand? And, are the judgement past as result of probing, random evaluations? Other questions in the Post interview generate a lower level of DKs (3 to 5% average), so the quality of the data produced by probing is

less of an issue.

Probing doesn't reduce the level of DKs at a similar rate for all questions. The rate of DK decrease is mediated by the familiarity and generality of the subject matter. The more general and familiar, the lower the level of DKs and the highest the rate of decrease when probed, as supported by the evaluation questions.

Probing seems to emerge as a powerful explanation for the DK differences across mode. But from the above analyses we can not untangle what role the different social setting of the T and FTF plays. Is the precipitous DKs drop (or part of it) to the "moral" traits items a function of the anonymity of the T situation? Does the relative anonymity of the T situation allow the respondent to "guess more" or to give less socially desirable answers instead of DK?

It is clear that we need to know more about the effects of probing, as well as the real effects of the differences in social situation between personal and telephone interviewing.

Careful thought should be given to what questions to probe. We think it is ultimately desirable to have **Probe** instructions written in the questionnaire rather than in the instruction book, particularly in a mixed mode situation. Ideally, the questionnaire would be laid out so that is possible to code when a probe was used, i.e., a response category would exist labeled "Check here if probe was used."

It may be possible to begin by experimenting with a couple of questions on the PILOT. It would be useful if we could set up a real experiment by buying some time in the Monthly Survey using probes systematically with only half of the respondents. We would be controlling for differences in staff and social situation, helping to isolate the effect of Probing.