

DOCKETED
USNRC

Dated: January 22, 1988

'88 JAN 26 A11:45

UNITED STATES OF AMERICA
NUCLEAR REGULATORY COMMISSION

OFFICE OF SECRETARY
DOCKETING & SERVICE
BRANCH

before the
ATOMIC SAFETY AND LICENSING BOARD

In the Matter of)	
)	
PUBLIC SERVICE COMPANY)	Docket Nos. 50-443-OL
OF NEW HAMPSHIRE, <u>et al.</u>)	50-444-OL
)	
(Seabrook Station, Units 1)	(Offsite Emergency
and 2))	Planning Issues)
)	

APPLICANTS' REBUTTAL TESTIMONY NO. 3

(REBUTTAL TO THE TESTIMONY OF ZEIGLER, JOHNSON
AND COLE REGARDING THE SDA TELEPHONE SURVEY
CONDUCTED FOR THE COMMONWEALTH OF MASSACHUSETTS)

Witnesses: Bruce D. Spencer
Dennis S. Mileti

Applicants' rebuttal testimony regarding the Telephone Survey conducted by Social Data Analysts, Inc., ("SDA") at the request of the Attorney General for the Commonwealth of Massachusetts was developed from two viewpoints. First, a study was done with regard to external validity, or the ability of the Survey findings to be generalized to the general population which did not participate in the Survey. The second area of review looked at internal validity or the examination of the questions within the questionnaire with

regard to the ability of those questions and answers, actually to measure what they purport to measure with associated freedom from systematic error or bias. These two viewpoints are presented below.

However, first, and perhaps of foremost importance, is the fact that the SDA Telephone Survey is a study of behavioral intentions. Pre-emergency intentions have little if anything to do with actual behavior. The lack of relationship between behavioral intentions and actual future behavior in a real emergency is as true for the public as for special sub-groups such as emergency workers. This basic and profoundly important point must not be lost in the context of the critique of the technical aspects of SDA's poll which follows. Even a behavioral intentions poll that was not troubled by factors which would detract from its external and internal validity would not produce data indicative of actual public response to an actual future emergency which has not been experienced. Human response in an actual emergency is largely directed by factors which prevail during the emergency as it is being experienced. These factors, for example, would include the frequency with which emergency warnings are heard and confirmed, interaction with other persons as people engage in response decision-making, and other such factors which cannot be taken into account by a pre-emergency poll. Behavioral intentions regarding future emergency response by a segment of the public would not be

". . . roughly representative of what the EPZ population would do in an accident at the Seabrook Station" (Testimony of Zeigler, Johnson and Cole, p. 18) again, even if the SDA poll were free of external and internal validity problems. Such intentions, in other words, can be nothing more than what interviewees thought on the day that they were interviewed taking into account only what they may or may not have had in mind when they answered the Survey questions. In contrast, actual public behavior in an actual future emergency is the consequence of factors and relationships which cannot be simulated in pre-emergency polls or surveys. These factors and how they affect behavior are well known from actual studies of actual behavior in actual emergencies. These, not behavioral intention polls, should guide and determine emergency planning for actual emergencies at Seabrook.

I. Analysis of External Validity

The sampling methodology employed in the Telephone Survey conducted by SDA described in Attachment 5 to the Testimony of Donald J. Zeigler, James H. Johnson, Jr., and Stephen Cole on behalf of the Attorney General for the Commonwealth of Massachusetts, "Behavior During a Radiological Incident: Reactions of EPZ Residents to a Possible Accident at the Seabrook Nuclear Power Station", cannot, in our opinion, ensure accurate descriptions and

predictions for the population that the Survey purports to describe. Claims that:

The results of the survey were generalizable to all households with telephones within the EPZ (Attachment 5, p. 3),

With the exception of the few households who do not have residential telephones, the sample is an accurate way to generalize to all households living in the EPZ" (Testimony of Zeigler, Johnson and Cole, p. 16), and

we can be confident that the results we obtained are roughly representative of what the EPZ population would do in an accident at the Seabrook Station (Testimony of Zeigler, Johnson and Cole, p. 18)

are unfounded. The problems with the design and execution of the sampling procedures are so serious that the Survey data and interpretations of that data should not be trusted.

The Survey is described as a random sample of households with residential telephones, not a random sample of individuals (Attachment 5, p. 43). The sample was drawn with a "complex procedure" (Attachment 5, p. 40). A summary of the design of the sampling procedure is provided below. Although the details of the design are technical, examination of those details will show four things.

First, the sample design systematically excluded some unknown proportion of EPZ households. Not only were households without a residential telephone excluded (Attachment 5, p. 40), but an unknown proportion of households with residential telephones who lived near the

boundary of the EPZ were systematically excluded from the Survey.

Second, the Survey did not seek a random sample of heads of households. The Survey is not representative even of the households that participated in the Survey because the responding heads of the participating households may differ from the other heads of those households.

Third, many households -- perhaps more than half of the households in the EPZ -- had no chance of participating in the Survey. It is inappropriate to claim that the Survey represents households that had no chance of participating in the Survey.

Fourth, the sampling errors appear to have been calculated as if the sampling design were a far less complicated one. The consequence of ignoring the complexity of the design is to understate the sampling errors, i.e., the sampling errors described in Attachment 5, pp. 52-53 and the Testimony of Zeigler, Johnson and Cole, pp. 16-17 are too small numerically and give a misleading impression of more reliability than was actually attained. (Sampling errors do not reflect validity.)

The first step in drawing the sample was an effort to list all telephone exchanges containing telephone numbers of residents of the EPZ. However, exchanges for which less than 15% of the numbers were determined to be within the EPZ were excluded from the list. Since those excluded exchanges were

"areas which straddle the boundaries of the EPZ" (Attachment 5, p. 41), the sampling procedure systematically excluded some proportion of the EPZ residents who lived near the boundary of the EPZ. The magnitude of the exclusion is not discussed in Attachment 5 or in the Testimony but simply opined on cross-examination to be a very small number. (December 16, 1987, Tr. 7954)

Telephone numbers were selected from the listed exchanges "in such a way so that the proportion of numbers in the sample in a particular exchange would be the same as the proportion of numbers in the population in that exchange. The sample utilized is a random digit dial sample in which the last two digits in the telephone number are selected at random by a computer from among all those working blocks in a particular exchange" (Attachment 5, p. 42).

Once a telephone number was selected and a "contact" was made, the interviewers were instructed to ask to speak to the male or female head of household (Attachment 5, p. 43). Since sex quotas were employed (Attachment 5, p. 43), it is presumed that the interviewers were not seeking household heads of one sex or the other, but rather they would speak to a head of either sex up until that point in the Survey when they had met their quota of males (or females), after which point they would only speak to females (or males). Table A3, "Failure to Complete", Attachment 5, page 57, identifies 170 New Hampshire and 79 Massachusetts calls which were not

completed because callers "could not obtain correct sex". No random sampling was performed within the households (December 16, 1987, Tr. 7960), therefore the sample is not a random sample of heads of households.

The non-random selection of the respondent within the selected households is critical and extremely unfortunate because it means that of those households containing more than one head, the sample over-represents those heads who were home and willing to answer the phone. Beliefs, knowledge, and attitudes can vary markedly between different heads of the same household, and thus the typical attitudes of the responding heads would not be the same as typical attitudes within their households. Indeed, recognition of this variation between two heads of the same household appears to have led SDA to use sex quotas:

A sex quota was used to insure that the final sample would represent the population in terms of sex. It was important to make sure that women were not over represented as it is well-known from prior surveys that the attitudes of men and women toward issues like nuclear power generally differ. (Attachment 5, p. 43)

If the sample were truly a random sample of households and of their heads then no quota sampling would have been necessary. Not only do men and women have different attitudes, but so may heads of households who are home and willing to be interviewed and heads of households who are not home or not willing to be interviewed. The use of quotas by

sex certainly does not avoid this problem. Doctor Cole has remarked about quota samples:

In my opinion it is dangerous to generalize from this type of sample [a quota sample] to a population. Another flaw which might have created bias is the failure . . . to use a systematic procedure for selecting the member of the household to be interviewed. (Testimony of Zeigler, Johnson and Cole, pp. 30-31).

Thus, the Survey is not representative even of the households that participated in the Survey because the responding heads of the participating households may differ from the other, non-responding heads of those same households.

A further problem with the use of quotas is that the quotas used by SDA were set according to estimates of the proportions of men and women in the population, and not according to the proportions of male or female heads of households in the various towns. To the extent that those proportions differ from each other, the use of quotas ensures a maldistribution of respondents by sex.

It is obvious that the sample could not represent those households in the EPZ lacking residential telephones. It is claimed that "data . . . indicate that more than 95% of the residents of the EPZ have telephones in their homes" (Testimony of Zeigler, Johnson and Cole, p. 14). However, no estimate of the proportion of households (as contrasted with persons) with telephones is offered in the Testimony, although Doctor Cole interpreted the "data" in cross-examination to the effect that "Somewhat less than 5 percent

of households do not have telephones." (December 16, 1987, Tr. 7948)

In addition to the problems of exclusion of households in the EPZ and non-random sampling previously noted, the Survey suffered yet another major problem -- nonresponse. As Doctor Cole has correctly pointed out: "There is no way to be certain that the people who refused to participate in the Survey would have answered the questions in the same way as those who did participate", and later, "The lower the response rate the less confidence we could have in the Survey results." (Testimony of Zeigler, Johnson and Cole, p. 18) In addition to people who directly refuse to participate in the Survey, we must also consider those who were denied the chance to participate because they were not at home, their line was busy, the interviewers had difficulty communicating with them (for example, persons who did not speak English well), or their telephone was out of order.

Therefore, the Survey could only represent those households that had a chance of participation in an interview. Further, it is likely that less than half of the households in the EPZ had a chance of participating in the Telephone Survey. A total of 6,611 telephone numbers were selected for the Survey. These numbers are classified by SDA as follows (Attachment 5, pp. 47ff):

1,055 =	no answers after 3 callbacks
457 =	continuously busy or head of household unreachable

- 2,270 = not working residential numbers (and some businesses)
- 93 = communication too difficult ("language or psychological problem")
- 249 = interviews were not conducted because could not obtain correct sex (quota filled for available sex)
- 793 = refusals
- 100 = households outside EPZ
- 190 = interviews were not conducted because quota for town was filled
- 1,404 = interviews were conducted.

In order to calculate precisely the fraction of the households in the EPZ that had a chance of participating in the Telephone Survey, we need more information. How many of the 1,055 "no answers" were residential telephones? (Some undoubtedly were business phones.) How many of the "continuously busy" numbers were residential? How many of the 2,270 "not working" numbers were residential numbers? Since this information was not available, we will consider a range of alternative assumptions. In the extreme case that all of these numbers were residential as the cross-examination testimony seems to imply (December 11, 1988, Tr. 7954-56), the fraction of households with telephones covered by the Survey would be less than 30%. Even if none of the 2,270 "not working" numbers were residential, the fraction of households with telephones covered by the Survey would be less than 40%. Those assumptions are extreme, but they yield lower bounds on the coverage of the EPZ households with

residential telephones. If we assumed that half of the 1,055 "no answers" were really residential numbers, 75% of the 457 "continuously busy or head of household unreachable" were residential, and as stated by Doctor Cole in cross-examination (December 16, 1987, Tr. 7954-56) none of the 2,270 "not working" numbers were residential, the fraction of households with residential telephones covered by the Survey would still be less than 50%.

In the Testimony of Zeigler, Johnson and Cole (p. 18), a "completion rate" of 64% is calculated as the ratio of the number of completed interviews to the sum of the completed interviews and the refusals. That rate ignores the 93 interviews that could not be completed because the respondent did not speak English or for some other communication or "psychological" problem. The rate also ignores residential phones that were not working or not answered or busy during the initial call and the three call-backs. Considering such cases suggests that the proportion of the households with residential telephones covered by the Survey is surely less than 60% and could well be less than 50%. Consideration of the additional households with no possibility of selection into the sample further diminishes the chances that as many as half of the households in the EPZ are represented by the Survey.

Statistical theory provides no basis for generalizing results from the Survey to persons or households with no

chance of participation into the Telephone Survey. One can try to make assumptions that those who were excluded from or refused to participate in the Survey are similar to those who did participate, but those assumptions cannot be trusted unless they can be tested empirically. It is claimed by Doctor Cole that despite massive amounts of nonresponse, ". . . given past surveys we have conducted utilizing the same methods, we can be confident that the results we obtained are roughly representative of what the EPZ population would do in an accident at the Seabrook Station" (Testimony of Zeigler, Johnson and Cole, p. 18) but no empirical evidence is provided to support that claim. Indeed, Doctor Cole has admitted that ". . . important in assessing the adequacy of the survey results are the number of no answers, busy signals, or no eligible respondent at home. There can be no way of knowing whether these people would have answered differently than those interviewed." (Testimony of Zeigler, Johnson and Cole, pp. 18-19). One possible way of trying to see whether those eligible to be interviewed would have answered in the same ways as those who actually were interviewed is to compare the Survey results with known statistics, such as census statistics. Not all of the Survey statistics can be compared because not all of the questions on the Survey are asked in the census or in another high-quality data source. However, a demonstrated agreement between some proportion of the questions on the Survey and

census (or other external criteria) would certainly lend more credibility to the Survey's results, even if it would not be proof that the Survey was representative with respect to the questions that could not be matched against census (or other) benchmarks.

The low coverage of the households (less than 60% or maybe less than 50% of those with phones and even less than that of all households) in the EPZ is so inadequate that the Survey cannot support statistical generalizations to all the households in the EPZ. The quality of the Survey is too low for the results to be trusted for use in important decision-making. The accuracy of the statistics based on the Survey is simply too suspect.

Sampling theory provides a means of estimating the variability in statistics that would occur from one sample to another as a result of the randomization that was used in the sampling. The term "sampling error" is used in Attachment 5, pp. 52-53, and the Testimony of Zeigler, Johnson and Cole, pp. 16-17, to describe the typical size of the variability. Sampling error does not reflect the magnitude of other sources of error in the Survey, such as nonresponse, lack of randomized selection of head of household, response biases due to question wording and ordering, and so forth. The interpretation of sampling error in Attachment 5, p. 53, suggests that it is computed as approximately twice the standard error. (The square of the standard error of a

statistic equals the average squared difference between a statistic and its average value, where the average refers to the average over hypothetical independent repetitions of the sampling procedure under identical conditions. The standard error may also be interpreted as the typical size of the difference between a statistic and its average value.) To calculate standard errors applicable to complex sampling procedures is rather complicated. For certain kinds of simple sampling procedures, however, the standard error of a percentage, say P , may be easily computed as the square root of the ratio of P times $100\% - P$ to the number of interviews. The standard error is largest when the percentage P is equal to 50%, in which case the standard error is equal to 50% divided by the square root of the number of interviews. With 915 interviews (the number of completed New Hampshire interviews) the standard error would then be 1.65% and the sampling error would be 3.30%; with 489 interviews (the number of Massachusetts completed interviews) the sampling error would be 4.45%. In essence, this simple formula was used to calculate the sampling errors used in the Zeigler, Johnson and Cole Testimony. (December 16, 1987, Tr. 7990-92, 8021)

In order to estimate sampling errors correctly (i.e., accurately), one must take into consideration the exact manner in which the sample was selected. The sample is described as "a stratified random sample of households with

residential telephones" (Attachment 5, p. 40). Doctor Cole at (December 16, 1987 Tr. 7949) agreed that the sample was stratified in essence into 23 strata. In addition, the description of the sampling procedure suggests that multistage sampling was used. It is important to know that multistage sampling was used because, other things being equal, sampling errors for multistage samples tend to be larger than sampling errors for one stage samples. Lacking a more detailed account of how the sample was selected, we cannot say for certain that the sample was indeed a multistage sample, but we believe that it was. However, the Zeigler, Johnson and Cole Testimony does not address these matters and the sampling errors reported in the Zeigler, Johnson and Cole Testimony were calculated as if no multistage or stratified sampling were used.

Thus, the simple formula for calculating standard errors appears to be inappropriate for the Telephone Survey. The actual sampling errors quite possibly are considerably larger. Furthermore, certain statistics are calculated on small subgroups of the interviews, and the standard errors for those statistics are enormously larger.

In particular, the sampling errors for the statistics on emergency workers in the New Hampshire EPZ are often far greater than 3%, even with the simplified formula described earlier in this rebuttal. For example Table 2 in the Zeigler, Johnson and Cole Testimony , p. 51, estimates that

19.4% of the emergency work roles are assigned to police, but the sampling error under the simplified formula is 14%, so the sampling variability is almost as large as the calculated statistic! If the complexity of the design were taken into account, the sampling error would probably be larger than 14%.

For the same reason, the statistics on the behavioral intentions of emergency workers are also extremely unreliable. Table 1 of the Zeigler, Johnson and Cole Testimony , p. 49, presents statistics for emergency personnel showing that 52% would perform emergency work, 39% would check on their families, 3% would leave the area, 3% would do something else, and 3% did not know what they would do in an evacuation advisory. However, those statistics are based on a sample of only 31 emergency workers and the sampling errors are large. Even using the simplified formula (which underestimates the sampling errors), the sampling error for the percentage saying they would check on their families is more than 17% and the sampling error for the percentage who would perform emergency work is more than 19%.

II. Analysis of Internal Validity

A large amount of systematic error or bias exists in the questionnaire used in this Telephone Survey. In other words, the answers which Survey respondents gave to the questions they were asked without doubt have been systematically colored or influenced by factors (for example, the wording

and ordering of questions) beyond their actual judgments. Sufficient sources of systematic measurement error (bias) exist to such a degree that we must conclude that the results of this Survey lack a basis for internal validity; we do not trust, therefore, that Survey findings represent a reasonably accurate representation of the actual views, judgments or opinions of persons interviewed. The many reasons why we have reached this conclusion follow.

Most of the bias (or sources of systematic measurement error) in the questionnaire are located in the first parts of the instrument. This is unfortunate because bias early in an instrument not only affects answers to the biased questions but can carry forward to subsequent questions which, taken alone, may not themselves be biasing.

The first topical question in the present Survey is numbered question 14. This question was worded as follows.

In general, how dangerous do you think it
would be to live near a nuclear power
plant?

The structured response categories read to the respondent were limited to the three which follow.

- 1 = very dangerous
- 2 = dangerous
- 3 = not dangerous at all

The question, "In general, how dangerous do you think . . ." implies an answer to the respondent before the question is even finished being read by use of the word "dangerous". It therefore leads the respondent to an opinion of "dangerous".

In addition, the range of possible answers for this question also contains a source of bias particularly when one considers how the range of answers read to the respondent would interact with the biasing question wording. This entire question and its answers take only a few seconds to read to the respondent, yet before the respondent has a chance to offer his/her opinion they have heard the word "dangerous" four times. This is a source of systematic measurement error or bias since it would lead respondents to an opinion of "dangerous". This question and the answers read to respondents at the conclusion of the reading of the question more resemble a lecture on how dangerous nuclear power plants are than social science measurement relatively free of systematic error, or at least social science measurement which has made a reasonable attempt to minimize systematic error or bias.

The second question on the questionnaire and its answers as read to respondents forces the respondent to select a general value position on nuclear power:

15. Would you describe yourself as

1. = a supporter of nuclear power plants as a means of providing electricity.
2. = an opponent of nuclear power plants, or
3. = you haven't made up your mind yet on this issue?

The answers given to this question would contain bias since respondents heretofore have been instructed that nuclear power plants are dangerous due to bias introduced in the first (number 14) question. Respondents are here forced to become a "supporter" of nuclear power, and "opponent," or else claim that their minds are not yet made up. This dichotomization of opinion on an issue on which opinions range along a continuum is biasing, because whichever position is chosen, respondents will remember their selection and labor to be as consistent as possible with their choice in answering all subsequent questions.

The next question (number 16) was "Do you think that the Seabrook Nuclear Power Plant should be allowed to operate to generate electricity?" This question shows no major internal sources of systematic error. However, its position in the questionnaire is after questions 14 and 15 which do bias results. Interactive bias would operate from questions 14 and 15 on answers to question 16. For example, question 14 "teaches" people that nuclear power is dangerous and would serve to bias answers to question 16 toward "no" (the answer consistent with the bias introduced in question 14). A similar interactive bias on answers to question 16 would have been operating from question 15.

Question 17 was "Given where you live, do you think you would be affected by a release of radiation if a serious problem developed at the Seabrook nuclear power station after

it started operating?". This question likely elicited measurement influenced by systematic error because of its position in the questionnaire. For example, question 14 biased persons to say nuclear power was "dangerous"; once that position was adopted, it would bias persons away from the answer "no" to question 17 (an admission that nuclear power is not dangerous, for all practical purposes).

However, the more important concern to be had with the first four questions (numbers 14, 15, 16 and 17) in the questionnaire is not that the answers given by respondents to these questions were themselves subject to systematic error; the prime problem that questions 14, 15, 16 and 17 present to the internal validity of this questionnaire is the effect they have by introducing systematic error or bias into subsequent question answers in the remainder of the questionnaire. Taken together, the first four questions in the questionnaire serve to create unique subsets of study respondents, for example, respondents who voiced the following perceptions to their interviewer: nuclear power is dangerous (question 14); I am an opponent of nuclear power because it is dangerous (question 15); I am an opponent of nuclear power because it is dangerous and, therefore I do not think Seabrook should be allowed to operate (question 16); and finally, I am an opponent of nuclear power because it is dangerous, and therefore I do not think Seabrook should be allowed to operate, therefore, of course I think I would be

affected if Seabrook had a serious problem after it began operating (question 17). After just four questions, this questionnaire has created study respondents so boxed into a corner as to significantly guarantee that answers to subsequent questions would be influenced (colored, biased, and so on) by the box in which respondents must have found themselves. Interviewees desperately try to be consistent during interviews. How now, for example, can a person already committed to the above illustrative position select "go about your normal business" as an answer to a question about emergency response after reading a scenario in which a release of radiation were asked to be assumed (see, for example, question and answers number 274)? The answer is obviously that the respondent would have been biased toward another answer more consistent with the above illustrative position, for example, "leave your home and go somewhere else". Conversely, how could a respondent in the opposite polar box (nuclear power is not dangerous at all, I am a supporter of nuclear power, I think Seabrook should be allowed to operate, and given where I live I do not think I would be affected by a release of radiation if Seabrook had a serious problem) select "leave your home and go somewhere else" as an answer to, for example, question 274?

Questions 14, 15, 16 and 17, however, would have biased the sample of respondents in the direction of being in the former "box" and away from the latter; among other reasons

because of the directional bias contained in the first question. Once respondents had completed hearing and providing answers to the first four questions, enough systematic error would have been introduced into this study to lead to the clear conclusion that subsequent question answers (particularly those on behavioral intentions) would lack internal validity and inflate intended evacuation estimates. This would be the case because of interactive bias introduced by the first four questions and answers, and the "box" into which they would have placed respondents. Answers by respondents to protective action behavioral intention questions (numbers 20, 31, 274 and 312, for example) would have been subject to this interactive bias.

The answers read to respondents to these same protective action behavioral intention questions (numbers 20, 31, 274 and 312, for example) were as follows.

1. = go about your normal business, or
2. = stay inside your home (or where you are) or
3. = leave your home (the place where you are) and go somewhere else

These response categories are neither mutually exclusive nor exhaustive -- it is possible to go about normal business by staying home or by leaving and going somewhere else.

A final problem exists in the questionnaire regarding internal validity in reference to protective action behavioral intentions questions numbered 20, 31, 274 and 312. People were asked to speculate about their intended behavior

in response to simulated emergency information. The information simulated for study respondents, however, does not mirror the emergency information which the public would actually encounter in the event of an emergency at Seabrook. As a consequence, therefore, as noted earlier, answers about behavioral intentions to the emergency information presented to study respondents in this Survey can shed no light on how people might behave in response to the actual form and type of emergency information that would characterize an actual emergency at Seabrook.

Other sources of systematic error or bias exist in the questionnaire. Question 42, for example, reads as follows.

When you heard this message on the radio
how likely do you think it would be that
you and your family would be exposed to a
dangerous level of radiation?

The answers read to the respondents were: (1) very likely, (2) somewhat likely, and (3) very unlikely. Interactive bias from questions 14, 15, 16 and 17 would also direct answers to this perceived risk question. Interactive bias from the first four questions would also direct answers to question 311 which follows:

Suppose there was an accident at the
Seabrook Station and the State Civil
Defense officials said that everybody
living within ten miles of the plant
should evacuate but that everybody who
lived more than 10 miles away from the
plant was safe. Would you believe the
State Civil Defense officials that people
living more than 10 miles away were safe?

Answers to this question would be colored by stated perceptions given as answers to questions 14 and 17, for example.

Questions 344, 345 and 346 in the questionnaire were directed only to respondents who admitted in the interview (see question 342) to having an assigned role in the Seabrook evacuation plan; and these questions were only asked on the New Hampshire portion of the sample. Question 344 reads as follows:

Suppose that the Seabrook Nuclear Power Station is licensed and begins to operate. If there were a problem at the plant and you heard that a ten-mile zone had to evacuate, what would you do first?

The answers read to the respondents for this question follow:

- 1 = report to my assigned place to help the evacuation
- 2 = make sure my family was safely out of the evacuation zone
- 3 = leave the evacuation zone to make sure I was in a safe place
- 4 = something else

Question 345 is the next question asked of emergency workers and it reads as follows:

How would you make sure that your family was safely out of the evacuation zone?

The answers read to the respondents so they could select one follows:

- 1 = go home and drive your family
to a safe place out of the
evacuation zone
- 2 = call home and tell your family
to leave without you
- 3 = some other way

Question 346 is the next question asked of emergency workers and it reads as follows.

If there was a nuclear accident at Seabrook Station requiring the evacuation of people within a ten mile zone, how dangerous do you think it would be for you to spend several hours in your emergency assignment?

The answers read to the respondents follow.

- 1 = so dangerous that it would be
life threatening
- 2 = very dangerous
- 3 = somewhat dangerous
- 4 = not dangerous

The answers obtained to these questions would have been subject to bias for several reasons. In reference to question 344, for example, no choice is provided the respondent regarding what extensive emergency behavioral research illustrates as what most trained emergency workers actually do in the emergency mobilization period (for example, answers 1 and 2 are typically done at the same time). The do "something else" option in the answers to questions 344 and 345 does not correct for this deficiency as interviewees typically select answers from the list they are provided. Answer to question 344 would be systematically

directed toward unrealistic choices about behavioral intentions; answers to questions 344, 345 and 346 would also have been biased interactively because of the "dangerous" bias in question 14, for example. Additionally, questions 344 and 345 and their answers are constructed in such a way that respondents are forced to choose between work, family and personal safety. It overlooks that each can and typically is served at the same time in actual emergencies. Questions 344 and 345 and their answers are more generic value measures of which object (job versus family) is, in general, of higher priority to the respondent. These answers lack internal validity as accurate behavioral intentions which, even if accurately measured, have little if anything to do with actual behavior in an actual emergency.

Question 348 reads as follows:

Currently plans are to have Civil Defense officials supervise an evacuation if this should become necessary. If as a result of an accident at Seabrook, you decide to leave the area and a Traffic Control official who was assigned to prevent traffic congestion told you not to drive on a road that you wanted to use, do you think you would:

1 = go where you wanted to go, or

2 = go where you were told to go

This question is a text book example of how not to ask questions in questionnaires. It illustrates measurement without reliability (different answers would be obtained if measurement were reattempted). Answers would depend on what

people had in their minds when they heard the words "told you" (over radio?, personally, as they directed street traffic?), "official" (someone in a uniform?, someone else?), "prevent traffic congestion" (for purposes of safety?, for convenience?), and so on. The question also presumes a conflict in the minds of evacuees (you want to go one way, and "they" want you to go another) which is a scenario which ignores actual human perception in actual emergencies as they are being experienced -- a "collective will" with "collective safety" as the prime motive for individual behavior. Finally, this question would elicit biased answers from respondents by its questionnaire position; it comes directly after the "Chernobyl" question and some respondents would answer this question with that type emergency mind.

Finally, this Survey was performed over the telephone during which family "spokespersons" were interviewed. Family "spokespersons" were individuals who were not able to take into account in the interview the input from other family members, for example, the discussions between family members leading up to protective action decisions. The correct unit of analysis for the interview should have been the entire family and not just one self-selected "spokesperson", since in a real emergency protective action decisions would be made in the process of family interaction. Dr. Cole's colleagues Drs. Johnson and Zeigler well understand this family evacuation decision-making process; they have, in fact, even

diagrammed it (see Stanley D. Brunn, James H. Johnson, Jr., and Donald J. Zeigler. 1979. Final Report on a Social Survey of Three Mile Island Area Residents, East Lansing: Michigan State University, Dept. of Geography, page 46.)

Interviewing individual "spokespersons" rather than the family unit, therefore, significantly deflates the internal validity of the study design. The Survey gathered behavioral intentions data from individuals, yet it is largely groups (for example, families) who respond to actual emergencies. Family behavioral intentions and "spokesperson" behavioral intentions are not the same.