Many survey researchers believe that the quality of answers to attitude questions may be improved by using no-opinion filters to remove those responses that are not based on true opinions. The authors investigate whether reliability, one aspect of data quality, is greater for filtered questions. Split-ballot experiments from two telephone surveys of the Akron, Ohio metropolitan area are used to estimate the effect of filtered questions on the reliability of items from Srole's anomia scale, Rosenberg's self-esteem scale, and a scale of attitudes toward lawyers. There is no evidence that filtered questions improve reliability, and in the case of self-esteem, reliability may be reduced on the filtered form.

No-Opinion Filters and Attitude Measurement Reliability

McKEE J. McCLENDON
University of Akron
DUANE F. ALWIN
University of Michigan

A standard practice in surveys of attitudes and opinions is to discourage, or not encourage, respondents from saying "don't know" (DK). When they occur, however, DK responses are often related to low education (e.g., Converse 1976-1977), suggesting that they tend to be a function of lack of information rather than ambivalence or indifference. When DKs reflect ambivalence or indifference, they might be validly scored at the midpoint of an attitude scale. But when no-opinion responses are based on lack of information, however, they do not belong on the scale at all. If survey practice discourages persons with insufficient information to give meaningful opinions from saying "don't know," then it might reasonably be

AUTHORS' NOTE: *This is a revised version of papers presented at the International Conference on Measurement Errors in Surveys, November 11-14, 1990, Tucson, AZ, and the 1990 meetings of the American Association for Public Opinion Research, Lancaster, Pennsylvania, May 17-20. We would like to thank the three anonymous referees for their helpful comments. Any remaining errors are our responsibility.*

SOCIOLOGICAL METHODS & RESEARCH, Vol. 21, No. 4, May 1993 438-464

expected that the quality of the data should suffer in some way. Thus some would argue that the first question of respondents should not be what is their opinion, but whether they have an opinion at all (Bogart 1967).

There is some information that bears on this question. Numerous split-ballot experiments have found that when respondents are explicitly asked if they have an opinion—the no-opinion filter form—the number of DK responses is significantly larger than when they must volunteer a no-opinion response—the standard question form (Schuman and Presser 1981; Bishop, Oldendick, and Tuchfarber 1983; McClendon 1986). This filter effect on DKs indicates that there is a group of respondents which expresses an opinion on the standard form of a question but would choose the DK option if offered the filtered form. Schuman and Presser (1981) have termed these respondents "floaters." Not all of the DKs on the filter form, however, are caused by the filter; some would have been volunteered anyway. Little is known about the identity of the respondents who are affected by no-opinion filters, except that their composition appears to shift from topic to topic. Whether these floaters express different opinions on the standard question form than those who are not affected by the filter (nonfloaters) appears to also vary from question to question, for the filtered distribution of opinions sometimes differs from the unfiltered (standard) distribution and sometimes does not (Schuman and Presser 1981; Bishop, Oldendick, and Tuchfarber 1983; McClendon 1986). Even less is known about how the filtered form affects the quality of responses in surveys. This article investigates the effects of no-opinion filters on the *reliability* of substantive responses. Using a split-ballot experimental design, we compare the estimated reliability of standard form responses with the reliability of filtered form responses after floaters have been removed.

THEORY

There are two diametrically opposed explanations for why filters increase no-opinion responses, explanations that have opposite implications for reliability. The most commonly held viewpoint is that persons who do not have true attitudes are under pressure to respond on the standard form (Converse 1964). This perspective forms the basis for the manifest intent of using opinion filters, that is, to screen out persons without opinions.

Respondents who do not have attitudes or opinions may feel pressure to respond to survey questions because they assume that interviewers want them to respond or because persons with opinions are held in higher esteem than those without opinions (Converse 1964). Similarly, Hippler and Schwarz (1989) argue that one of the norms of conversation specified by Grice (1975) implies that when an interviewer asks the respondent a question, there is a presupposition that the respondent is able to answer the question. Thus even respondents without true opinions may feel obliged to respond. Converse (1964) has termed the responses of such persons "nonattitudes" and has reported evidence consistent with the interpretation that they are essentially made on a random basis. If so, the responses of persons who do not have an opinion would increase random error variance and consequently reduce the reliability of the attitude measures.

We would argue that by explicitly asking if the respondent has an opinion, no-opinion filter questions should counter these social factors that induce the reporting of nonattitudes. The filtered form should make it clear to those without opinions that it is legitimate to give a DK response. Therefore, persons without true opinions who would give substantive responses on the standard form may choose DK on the filtered form of the question. If Converse is correct that these nonattitudes represent nothing more than random variance, reliability should increase on the filtered form of the question. This perspective on filtering will be referred to as the *random-response hypothesis*.

In order to facilitate theoretical predictions, we treat the hypothesis as an ideal type. Therefore, it is assumed that all of the floaters are persons who would respond randomly on the standard form; the filter is screening out only random responses. If so, the proportion of total variance that is random should be lower on the filtered form, which means that reliability should be higher on the filtered form. The hypothesis does not imply, however, that all random responders would necessarily opt for the no-opinion alternative on the filtered form. Thus the increase in reliability on the filtered form will be inversely related to the amount of random responding among nonfloaters. The increase in reliability will also be directly related to the percentage of persons who float. Therefore, the higher the percentage of floaters and the lower the random responding among nonfloaters, the greater will be the increase in reliability on the filtered form.

Consistent with this hypothesis, Andrews (1984) found greater reliability on the filtered form of questions. Andrews's conclusion, however, was based on aggregate results for a large pool of items containing both attitude questions and factual questions. Thus it is unclear whether it is equally valid for both subjective and objective phenomena. Furthermore, Alwin and Krosnick (1991) found no form differences in reliability for agree-disagree attitude items and a significantly higher reliability among 7-point rating scales when DK was not offered, a finding that contradicts both Converse's (1964) argument and Andrews's (1984) findings. The Alwin and Krosnick (1991) results, however, were not based on split-ballot experiments and thus the effect of question form was confounded to an unknown degree with the effects of question content.

A second explanation for the filter effect on DKs assumes that filters actually discourage some people who do have attitudes from expressing them. For example, Schuman and Presser (1981, p. 159) provide evidence suggesting that respondents who do not give an opinion on the filtered form answer the standard-form question on the basis of a general underlying attitude that they infer is relevant to the question. The general attitude allows them to formulate responses to questions that may be more specific in content. Respondents who pass through the filter, on the other hand, may be responding on the basis of information relevant to the specifics of the question, specifics that may even lead them to a different response than would be given on the basis of the general attitude. They suggest that the standard form may be a more valid approach to measuring general attitudes, whereas the filtered form may be better for measuring informed opinions (Schuman and Presser 1981, p. 160).

In line with this hypothesis, Hippler and Schwarz (1989) have argued that the filtered form induces people who do have opinions, and who express them on the standard form, to choose the DK alternative on the filtered form. This occurs because the filter (especially strongly worded filters) suggests that only "informed opinions" are desired. If so, respondents who have a global attitude on the issue, but not much specific expertise, may feel that their opinion may not be desired. Hippler and Schwarz (1989) found, for example, that the filtered form may lead respondents to expect more follow-up questions on the issue. Thus some people may respond with a DK because, lacking either sufficient expertise or motivation, they want to avoid the follow-up questions. In sum, according to this view, floaters can have global opinions on an issue, but these are not based on much specific expertise. We refer to this as the *generalized-response hypothesis*.

This hypothesis has different implications for understanding the effects of filters on reliability. If people who are filtered out have real opinions, even though they might be based on relatively uninformed global attitudes, then by definition, they are responding systematically, not randomly, on the standard form. Thus, under the assumptions of this hypothesis, there is no reason to expect the filtered form to reduce random error variance and increase reliability. If anything, responses driven by global attitudes reflect reliable variance. Therefore, assuming there are respondents who create random error variance on both the standard and filtered forms, screening out general attitudes using the filtered form will probably reduce the proportion of variance that is systematic and thus lower reliability.

There is an additional perspective that leads to the same predictions as the generalized-response hypothesis. It may be that the filtered form screens out responses that do not represent valid attitudes or opinions, but that are nevertheless reliable. Instead of responding randomly, persons without real opinions may introduce other types of response errors on the standard form such as acquiescence to agree-disagree questions, a bias in favor of either the last or the first response alternative, or a social desirability response bias. Such response effects create invalid but reliable variance on the standard form. Persons responding in this way, however, might well opt for a no-opinion response when explicitly offered that opportunity on the filtered form. Although this would increase validity, reliability would decrease on the filtered form, assuming that random response variance persists on the filtered form. McClendon (1991), however, found little evidence of reduced acquiescence and response-order effects on the filtered form.

We have stated the random-response hypothesis and the generalized-response hypothesis as though they were mutually exclusive. If this is the case, we would expect reliability to either be lower on the filtered form or higher on the filtered form, depending on which hypothesis is true. It is possible, however, that some floaters are random responders and others are generalized responders. If so, the filter might screen out approximately equal amounts of random variance and systematic variance, and thus there may be little or no difference in reliability between the standard and filtered forms.

MEASUREMENT MODEL

We examine these hypotheses by comparing the variance-covariance structures for survey measures across the two forms, standard and filtered. The utility of this approach can be seen by considering the following model for survey responses:

$$y = \mu + \tau + \varepsilon. \tag{1}$$

In Model 1, y is the response variable, μ is the population mean of y, τ is the underlying variable (latent "true" score) intended to be measured by y, and ε is a random error term. This model is based on classical true-score theory, which assumes that errors are random with respect to the underlying variable being measured (see Alwin [1989, 1991] for a discussion of this model in the context of survey data). In this article, we disregard the population means, dealing with variables that are measured as deviations from the population mean $(y - \mu)$, and our focus is entirely on the variance components of y.

Under the assumption that these response variance components are uncorrelated, response variance in the population can be written as

$$\sigma^2_y = \sigma^2_\tau + \sigma^2_\varepsilon \tag{2}$$

and the covariance between repeated measures of τ can be written as

$$\sigma_{y_1 y_2} = \sigma^2_\tau. \tag{3}$$

If we were able to obtain replicate (repeated) measures of y, then it would be possible to estimate the reliability of measurement by manipulating the above quantities, as follows:

$$\text{Reliability} = \sigma_{y_1 y_2}/\sigma^2_y. \tag{4}$$

But it is very rare that survey measurement designs permit this estimation because items are rarely replicated within the same survey (Alwin 1989). Indeed, it is often the case that panel designs are preferred to cross-sectional studies for reliability estimation because they offer the exact replication of questions.

If the y's are not measures of the same τ, that is, the y's measure different τs that are not perfectly correlated (i.e., the measures are not congeneric), then the traditional form of reliability estimation is inappropriate. We can, however, modify the model to entertain this possibility. In this case, one can more comfortably assume that each "true" variable (the τs) measures a "factor" in common, as well as a specific factor, as follows:

$$\tau_i = \beta_i\xi + \upsilon_i, \tag{5}$$

where τ_i is the "true" score measured by y_i, ξ is a latent common factor, and υ_i is the specific factor. The β_i coefficient is simply a linear scaling constant that converts units of ξ into units of τ_i.

By substituting this formulation in the above model for response variance in y_i, we have

$$y_i = \beta_i\xi + \upsilon_i + \varepsilon_i. \tag{6}$$

According to this modified version of the model, each response variable measures a factor in common with other similar variables, a specific component of variation unique to each measure, and a random error of measurement. In other words, by relaxing the "univocal" assumption of classic true-score theory, we arrive at a common factor representation (Lord and Novick 1968; Alwin and Jackson 1979; Bollen 1989). In this case, the covariance between the y's is

$$\sigma_{y_iy_j} = \sigma_{\tau_i\tau_j} = \beta_i\sigma^2_\xi\beta_j, \tag{7}$$

and the response variance is a function of the following components:

$$\sigma^2_{y_i} = \beta_i^2\sigma^2_\xi + \sigma^2_{\upsilon_i} + \sigma^2_{\varepsilon_i}. \tag{8}$$

For the present, we assume that υ_i and ε_i are both uncorrelated with ξ. Because in practice it is most often impossible to separate these variance components, we combine them as

$$\delta_i = \upsilon_i + \varepsilon_i \tag{9}$$

and we write

$$\sigma^2_{\delta_i} = \sigma^2_{\upsilon_i} + \sigma^2_{\varepsilon_i}. \tag{10}$$

Thus our decomposition of the response variance in y_i is

$$\sigma^2_{y_i} = \beta_i^2\sigma^2_\xi + \sigma^2_{\delta_i}. \tag{11}$$

These two components of variance can be estimated with a confirmatory factor analysis of a set of y's. The first component, $\beta_i^2\sigma^2_\xi$, is the common factor variance and the second component, the unique variance, is the sum of the item-specific variance and random response variance. If the common factor variance estimates the reliable variance, and if the unique variance estimates the measurement error variance, the reliability of y_i is

$$\text{Reliability} = \beta_i^2\sigma^2_\xi/(\beta_i^2\sigma^2_\xi + \sigma^2_{\delta_i}) \tag{12}$$

Under our model, however, we have defined the unique variance as containing both reliable specific variance and error variance, so this "reliability" coefficient will generally underestimate true reliability. This estimate is a lower-bound estimate because part of $\sigma^2_{\delta_i}$ is the nonrandom $\sigma^2_{\upsilon_i}$, the true-score specific variance. However, this is the best estimate that is possible with cross-sectional data. It is also consistent with widely used measures of the reliability of linear composites, such as an equally weighted sum of y's (Σy_i). Cronbach's alpha for Σy_i, which is computed from the variance/covariance matrix, is based on the assumption that $\beta_i = 1$ for all y_i, that is, it assumes a tau-equivalent measurement model. Heise and Bohrnstedt's (1970) reliability coefficient for Σy_i (i.e., omega) is based on a common-factor model like ours:

$$\Omega = 1 - \sum_i \sigma^2_{\upsilon_i}/\sigma^2_{\Sigma y_i}. \tag{13}$$

Our measure of reliability (equation 12) can be used to test the difference between the estimated reliability on the standard and filtered forms hypothesized by the two explanations given above on the effects of filtering on reliability. As noted above, the random-response hypothesis predicts that $\sigma^2_{y_i} = \sigma^2_{\varepsilon_i}$ for floaters' responses on the standard form; that is, the filtered form screens out only random variance. If so, both common variance $\beta^2_i\sigma^2_\xi$ and specific variance $\sigma^2_{\upsilon_i}$ should be greater proportions of $\sigma^2_{y_i}$ on the filtered form. Because Reliability = $\beta^2_i\sigma^2_\xi/(\beta^2_i\sigma^2_\xi + \sigma^2_{\delta_i})$, reliability should be greater on the filtered form.

These predictions have been expressed in terms of variance *proportions*. Predictions about form differences in the absolute amounts of the variance components are less certain. It is possible, for example, that floaters might have a larger random-variance proportion than nonfloaters but create a smaller amount of random variance than nonfloaters. The random-response hypothesis says that floaters respond entirely in a random fashion, but it does not predict whether these random responses are concentrated tightly around the mean or vary widely around the mean. Nevertheless, the most reasonable expectation to us is that floaters would contribute a greater absolute amount of random variance on the standard form than nonfloaters, and thus random variance would be less on the filtered form. Because floaters do not contribute to specific variance according to the random-response hypothesis, specific variance would be greater on the filtered form. Since we can only estimate $\sigma^2_{\delta_i} = \sigma^2_{\upsilon_i} + \sigma^2_{\varepsilon_i}$, we cannot predict whether unique variance is greater on the filtered form or less on the filtered form. However, because floaters do not contribute common variance, we can predict that $\beta^2_i\sigma^2_\xi$ will be greater on the filtered form. And last, because $\sigma_{y_iy_j} = \beta_i\sigma^2_\xi\beta_j$, the random-response hypothesis predicts that the covariances will be greater on the filtered form.

The generalized-response hypothesis predicts that floaters are responding superficially to the items on the basis of a general underlying attitude that is represented by the common factor ξ in our model. If the set of items is unbalanced (i.e., all items worded in the same direction), this common factor might also represent response sets such as acquiescence or social desirability, or a mixture of substance and response sets. This common factor creates correlated responses among a set of items.[1] Because floaters contribute only common variance, according to the generalized-response hypothesis, $\beta^2_i\sigma^2_\xi$ will be a smaller proportion of total variance on the filtered form. Thus Reliability = $\beta^2_i\sigma^2_\xi/(\beta^2_i\sigma^2_\xi + \sigma^2_{\delta_i})$ will be less on the filtered form. Again, although the generalized-response hypothesis speaks mainly to variance components as proportions of total variance, we would predict that floaters would have a greater absolute amount of common variance than nonfloaters. Thus the amount of common variance and the covariances that are created by the common variance should be less on the filtered form. On the other hand, because floaters contribute only common variance, the amounts of random variance and specific variance should both be greater on the filtered form, and thus unique variance $\sigma^2_{\delta_i}$ should be greater on the filtered form.

To summarize, the random-response hypothesis leads us to anticipate a greater proportion of common variance, a greater amount of common variance, and larger covariances on the filtered form. The generalized-response hypothesis leads us to expect the opposite results plus greater unique variance on the filtered form. If there happens to be approximately equal numbers of random-response floaters and generalized-response floaters, there may be little or no difference between forms in reliability and variance components.

DATA

The data analyzed were collected as parts of the 1984 and 1986 Akron Area Surveys, which were telephone surveys of Summit County, Ohio, conducted by the Department of Sociology at the University of Akron. Households were selected by random digit dialing and one respondent (18 years of age or older) was randomly selected within each household. Out of 1,350 households contacted in 1984, there were 799 completed or partially completed interviews (59.2 percent), 33.9 percent refused, 5.2 percent were unable to be interviewed, and 1.7 percent were "call backs" who could not be contacted again. The results for the 1,387 households contacted in 1986 were 759 completed or partially completed interviews (54.8 percent), 36.6 percent refusals, 5.4 percent unable to be interviewed, and 3.3 percent "call backs."

The questions consisted of four items from the 1984 survey measuring opinions about lawyers, four items from the 1986 survey measuring anomia (Srole 1956), and four items from the 1986 survey measuring self-esteem (Rosenberg 1965). The items are shown in Figure 1 in both an agree-disagree form and a forced-choice form, to which respondents were randomly assigned. Approximately one-third of the respondents in each survey received the agree-disagree form and two-thirds received the forced-choice form. Additionally, those respondents who received the forced-choice form were randomly assigned either a version with the response alternatives ordered as shown in Figure 1 or a version with the order of the response alternatives reversed.

The lawyer questions are dichotomous items in both the agree-disagree and forced-choice forms. The anomia and self-esteem items, however, used a four-category Likert-type scale for the agree-disagree form (strongly disagree, somewhat disagree, somewhat agree, strongly agree), which we scored from 1 to 4. In order to make the forced-choice form as equivalent to the Likert form as possible, respondents were asked if they felt strongly or only somewhat strongly about their choice after they had selected one of the two response alternatives. This allowed us to also score the forced-choice responses from 1 to 4.

Respondents were also randomly assigned to either a standard or a filtered version of the agree-disagree and forced-choice forms, thus creating a 3×2 factorial design. A respondent received the same form for each of the anomia and self-esteem items in the 1986 survey and the same form for each of the lawyer items in the 1984 survey. In the filtered version, "or don't you have an opinion about that?" was appended to each question. This has been called a quasi-filtered form by Schuman and Presser (1981). It is the weakest form of filtering, but as will be seen, it still can lead to substantial increases in DK responses.

This study will examine only differences in reliability between the standard and filtered forms of the questions. Theoretical predictions for the effects of these forms have been given above. Because there are not any compelling theoretical hypotheses concerning reliability differences between the agree-disagree and forced-choice forms, or between the different orders of the response alternatives on the forced-choice form, any effects that these factors might have on reliability will not be explored in this article. The effects of these forms and response orders on the validity of the univariate responses are reported by McClendon (1991).

Lawyers

HONEST: Lawyers are less honest and ethical than most other professionals.
Would you say that lawyers are more honest and ethical than most other professionals, that lawyers and other professionals are equally honest and ethical, or that lawyers are less honest and ethical?

CHARGE: Lawyers charge too much for their services.
Would you say that the fees lawyers charge for their services are fair, or that lawyers charge too much for their services?

CONFLICT: Lawyers increase conflict and trouble between people.
Would you say that lawyers decrease conflict and trouble between people, that they neither decrease nor increase conflict and trouble, or that lawyers increase conflict and trouble?

WINNING: Lawyers are more interested in winning cases than in the welfare of their clients.
Would you say that lawyers are more interested in the welfare of their clients than in winning cases, that they are equally interested in both, or that they are more interested in winning cases than in the welfare of their clients?

Anomia

LIVETOD: Nowadays a person has to live pretty much for today and let tomorrow take care of itself.
Would you say that nowadays a person can't just live for today, you have to make some plans for tomorrow, or that nowadays a person has to live pretty much for today and let tomorrow take care of itself?

CHILDFUT: It's hardly fair to bring children into the world with the way things look for the future.
Do you think it's all right to bring children into the world with the way things look for the future, or do you think it's hardly fair to bring children into the world with the way things look for the future?

COUNTON: These days a person doesn't really know whom he can count on.
Would you say that these days a person can know whom to count on, or that these days a person doesn't really know whom he can count on?

PUBOFF: There's little use writing to public officials because they often aren't interested in the problems of the average person.
Do you think that it can be useful to write to public officials about the problems of the average person, or that there's little use writing to public officials because they often aren't interested in the problems of the average person?

Self-Esteem

WORTH: I feel that I am a person of worth, at least on an equal basis with others.
Do you feel that you are a person of less worth than others, or do you feel that you are a person of worth, at least on an equal basis with others?

DOTHING: I am able to do things as well as most other people.
Would you say that you are *not* able to do things as well as most other people, or that you are able to do things as well as most other people?

RESPECT: I wish I could have more respect for myself.
Would you say that you have enough respect for yourself, or do you wish you could have more respect for yourself?

USELESS: I certainly feel useless at times.
Would you say that you seldom if ever feel useless, or that you certainly feel useless at times?

Figure 1: Agree-Disagree and Forced-Choice Forms of the Question[a]

a. Items 1-2 for the lawyers questions and all of the anomia and self-esteem items were also asked in quasi-filtered form by adding "or don't you have an opinion about that?" to the end of the item.

LISREL (Jöreskog and Sörbom 1989) will be used to test for form differences in reliability. The standard and filtered forms of the covariance matrices will be analyzed by the maximum likelihood (ML) method to estimate the parameters of the common factor model on each form of the questions. Cross-form equality constraints on the common factor variances and the unique variances will be imposed to test for form differences in these variance components. Acquiescence (agreeing-response bias) may be a problem on the agree-disagree form, especially for the unbalanced lawyer items and the unbalanced anomia items. Therefore, the tests for differences in reliability between the standard and filtered forms will be repeated using only the forced-choice forms of the questions, which comprise about two thirds of the cases. This will allow us to examine the effect of filtering on reliability in the absence of any influence of acquiescence.

The ML method assumes that the observed variables are measured at least approximately on an interval scale. The use of ML for ordinal variables may lead to greatly distorted parameter estimates and incorrect chi-square goodness-of-fit measures and standard errors (Jöreskog and Sörbom 1989, p. 223). When the observed variables are ordinal, Jöreskog and Sörbom (1989) recommend analyzing a matrix of polychoric correlations (or tetrachoric correlations for dichotomous variables) with a weighted least squares (WLS) method available in LISREL that provides asymptotically distribution-free estimates. Because our anomia and self-esteem variables are ordinal and the lawyer variables are dichotomous, we have also estimated the common factor model with the WLS method as applied to matrices of polychoric correlations for the anomia and self-esteem variables and tetrachoric correlations for the lawyer variables (these correlation matrices were computed by the PRELIS program).

FINDINGS

Table 1 gives both no-opinion percentages for each question and "listwise" no-opinion percentages for each of the three sets of questions. A "listwise" percentage is the percentage expressing no opinion for one or more items within a set. As expected, the filtered form caused an increase in the percentage of no-opinion responses for each

TABLE 1: Percentage of No-Opinion Responses by Question Form

Questions	*Standard*	*Filtered*
Lawyers		
HONEST	6.9	27.0
CHARGE	10.4	29.2
CONFLICT	8.8	22.4
WINNING	6.4	20.3
Listwise	24.5	60.0
n	384	380
Anomia		
LIVETOD	1.0	6.7
CHILDFUT	2.9	13.8
COUNTON	2.4	16.4
PUBOFF	1.8	11.7
Listwise	5.5	34.5
n	381	341
Self-Esteem		
WORTH	1.0	6.7
DOTHING	1.0	4.1
RESPECT	0.5	6.5
USELESS	0.5	10.0
Listwise	1.6	21.1
n	381	341

question in each of the three sets of items (Table 1). DKs were highest for the lawyer questions and lowest for the self-esteem questions on both the standard and filtered forms. The estimated percentage of respondents who would give an opinion on the standard form but would choose DK on the filtered form (i.e., floaters)—the difference in DK percentages between the standard and filtered forms—is also greatest for lawyers and smallest for self-esteem. The fact that the self-esteem items show the smallest filter effects has important theoretical implications for form differences in reliability that we will return to later. The listwise percentages are dramatically higher than those for the individual items, reaching a maximum value of fully 60.0 percent on the filtered form of the lawyer items. The estimated percentage "listwise filter effect" is 35.5 percent for lawyers, 29.0 percent for anomia, and 19.5 percent for self-esteem. These estimates are relevant for our comparisons of reliability between the standard and filtered forms that are based on listwise deletion of nonresponses.

TABLE 2: Covariances, Means, Variances, and Sample Sizes (n) by Question Form[a]

	1	*2*	*3*	*4*	*Mean*	$\sigma^2_{y_i}$	*Alpha*
Lawyers							
1 HONEST	—	.048	.034	.061	.783	.171	.554
2 CHARGE	.045	—	.030	.064	.362	.232	
3 CONFLICT	.020	.053	—	.059	.745	.191	
4 WINNING	.035	.074	.060	—	.597	.242	
Mean	.737	.368	.750	.605	—	(290)	
$\sigma^2_{y_i}$	.195	.234	.189	.241	(152)	—	
Alpha	.532						
Anomia							
1 LIVETOD	—	.322	.272	.261	1.994	1.198	.618
2 CHILDFUT	.144	—	.419	.337	1.864	1.020	
3 COUNTON	.209	.244	—	.352	2.297	1.162	
4 PUBOFF	.183	.274	.409	—	2.306	1.182	
Mean	1.841	1.786	2.050	2.318	—	(360)	
$\sigma^2_{y_i}$	.975	.863	1.171	1.168	(222)	—	
Alpha	.549						
Self-Esteem							
1 WORTH	—	.208	.135	.058	3.720	.411	.500
2 DOTHING	.146	—	.097	.103	3.587	.580	
3 RESPECT	.022	–.034	—	.356	3.157	1.085	
4 USELESS	.095	–.006	.237	—	2.893	1.117	
Mean	3.665	3.580	3.264	3.048	—	(375)	
$\sigma^2_{y_i}$	.418	.610	1.023	.994	(269)	—	
Alpha	.309						

a. The standard form is above the diagonal and the filtered form is below the diagonal.

The means, variances, and covariances of the items for each form, based on listwise deletion of nonresponse, are given in Table 2. The only significant mean difference between the standard and filtered forms is for COUNTON (McClendon 1991).

LISREL (version 7.16) was used to test the differences between forms in the variance and covariance matrices. First, for each set of questions, a model was specified in which the four variances were constrained to be equal across forms. The chi-square goodness-of-fit statistic for the maximum likelihood solution ($df = 4$) was used as the test of this equality constraint (Alwin and Jackson 1979). Second, the six covariances were constrained to be equal across forms, while leaving the variances unconstrained on each form. Again, the chi-

TABLE 3: Tests for Cross-Group Differences in Variances and Covariances

	Chi-Square	*df*	*p*
Lawyers			
$Eq.\sigma^2_{y_i}$	.93	4	>.90
$Eq.\sigma_{y_iy_j}$	3.47	6	>.70
Anomia			
$Eq.\sigma^2_{y_i}$	4.59	4	>.30
$Eq.\sigma_{y_iy_j}$	7.95	6	>.20
Self-Esteem			
$Eq.\sigma^2_{y_i}$	1.47	4	>.80
$Eq.\sigma_{y_iy_j}$	14.17	6	<.05

square goodness-of-fit statistic for the maximum likelihood solution ($df = 6$) was used to test the equality of the covariances.

There are not any noticeable differences in the variances across forms, and none were predicted. None of the three chi-square tests for equality of variances approached even marginal significance (Table 3). Thus we cannot conclude that there are any differences in the variances between the standard and filtered forms of either the lawyer, anomia, or self-esteem items.

The covariance matrices (Table 2), however, do show some noticeable differences across forms. The covariances for the lawyer items show a mixed picture, with two covariances somewhat larger on the filtered form and two somewhat larger on the standard form. That these differences tend to offset one another is shown by the fact that Cronbach's alpha (which is a positive monotonic function of the average covariance) is very similar on both forms (Table 2). The chi-square for the form difference in covariances was not significant for the lawyer items (Table 3).

The covariance differences for the anomia items tend to be larger than for the lawyer items, and with one exception, all are greater on the standard form. Consistent with this trend, alpha is slightly larger on the standard form (Table 2). Nevertheless, the chi-square test did not approach significance for the anomia items either (Table 3).

The covariances of the self-esteem items are also larger on the standard form, again with one exception. Alpha is also substantially

TABLE 4: Estimated Parameters of Measurement Model for Lawyers

	β_i	$\sigma^2_{\delta_i}$	R^2_i
Standard			
HONEST	.202	.130	.240
CHARGE	.207	.189	.185
CONFLICT	.177	.160	.164
WINNING	.312	.144	.402
σ^2_ξ	1.000		
	GFI = .998	AGFI = .990	
Filtered			
HONEST	.134	.177	.092
CHARGE	.274	.159	.320
CONFLICT	.199	.149	.210
WINNING	.277	.164	.319
σ^2_ξ	1.000		
	GFI = .997	AGFI = .985	
	Chi-square = 1.90	*df* = 4	$p = .755$

NOTE: $p < .05$ for all parameter estimates. GFI and AGFI are the goodness-of-fit index and the adjusted goodness-of-fit index, respectively.

greater on the standard form. Chi-square was significant for the self-esteem items (Table 3), indicating that the covariances were generally larger on the standard form. These tests support the predictions of the generalized-response hypothesis for the self-esteem items. There is no support, however, for either hypothesis for the lawyer and anomia items.

The variance-covariance matrices based on pairwise deletion of nonresponses were also inspected for the standard and filtered forms of each set of items (not shown). The patterns of the form differences in covariances were the same as those described above for listwise deletion.

LISREL was also used to estimate our common-factor measurement model for each set of items, separately for the standard and filtered forms. Table 4 shows the ML estimated parameters of a single-factor model for the lawyer items. The common factor was standardized to have a variance of unity and the factor loadings β_i and the unique variances $\sigma^2_{\delta_i}$ were estimated for each form. The squared correlation (R^2) between each indicator and the common factor (Table 4) is the ML estimate of the item's reliability.[2] The chi-square

TABLE 5: Tests for Parameter Differences Between the Standard and Filtered Forms

			Chi-Square		
Model	*Constraint*[a]	*df*	*ML*	*ML′*	*WLS*
		Lawyers			
1	none	4	1.90	3.49	1.74
2	Eq.β_i	8	4.77	8.45	6.24
3	Eq.$\sigma^2_{\delta_i}$	8	6.66	10.16	—
2-1	—	4	2.87	4.96	4.50
3-1	—	4	4.76	6.67	—
		Anomia			
1	none	4	1.09	3.15	0.94
2	Eq.β_i	8	8.64	6.30	7.74
3	Eq.$\sigma^2_{\delta_i}$	8	3.52	5.16	—
2-1	—	4	7.55	3.15	6.80
3-1	—	4	2.43	2.01	—
		Self-Esteem			
1	none	4	2.02	1.68	3.29
2	Eq.β_i, Eq.$\sigma_{\xi_1\xi_2}$	7	11.78	9.41	13.95*
3	Eq.$\sigma^2_{\delta_i}$	8	7.28	5.22	—
2-1	—	3	9.76**	7.73*	10.66**
3-1	—	4	5.26	3.54	—

a. Model constrains estimated parameters to be equal across the standard and filtered forms.
b. ML equals the maximum likelihood solution for the sample covariance matrix (ML′ is for the forced-choice form only). WLS is a weighted least squares solution for the polychoric correlations (anomia and self-esteem items) and the tetrachoric correlations (lawyer items).
*$p \leq .10$; **$p \leq .05$.

goodness-of-fit statistic ($df = 4$) is not significant ($p = .76$) for the lawyer items (Table 4), indicating that this single-factor model cannot be rejected.

There is a mixed pattern of form differences in the estimated parameters for the lawyer items, with some estimates being larger on the standard form and some being larger on the filtered form. Differences in the βs across forms were tested by constraining them to be equal across forms and using the change in chi-square ($df = 4$) as the test statistic. The change in chi-square was not significant (ML in Table 5). This indicates that we cannot conclude that there are any cross-form differences in the common-factor variance component of the lawyer items. This is consistent with the test for form differences in the covariances. The unique variances were also constrained to be equal across forms and the change in chi-square ($df = 4$) was again not significant (Table 5). Thus the prediction of the generalized-response hypothesis that unique variances would be greater on the filtered form is not supported for the lawyer items.

The same tests as above were conducted using only the forced-choice forms. There were also no significant form differences in common variance or unique variances (ML′ in Table 5).

The common factor model was also estimated using the tetrachoric correlations and the WLS method. When the βs were constrained to be equal across groups there was not a significant increase in chi-square (WLS in Table 5). When analyzing the correlation matrices (with σ_ξ^2 fixed at unity), we are estimating standardized solutions because the variances of the observed variables are unity. The βs are correlations between ξ and y and the $\sigma_{\delta_i}^2 = 1 - \beta_i^2$. Thus Reliability = β_i^2 and when the βs are constrained to be equal across forms, the estimated reliabilities are also equal across forms. Consequently, the change in chi-square reported under WLS in Table 5 is a test for form differences in reliability, based on a WLS analysis of the tetrachoric correlations. The WLS solution does not allow us to test for differences in the absolute amounts of common variance and unique variance. The single test of form differences in βs indicates that the reliabilities of the lawyer items are not significantly lower on the filter form. Because the outcomes of the tests for the WLS solutions are the same as for the ML solutions, we report only the ML estimates because they are unstandardized.

The parameter estimates for an identically specified single-factor model for the anomia items are given in Table 6. The fit of this model also cannot be rejected. There is again a mixed pattern of parameter differences across forms, although on the average the βs are slightly larger on the standard form. Using the same testing procedure as for the lawyer items, we find no significant differences between forms in either the factor loadings or the unique variances (Table 5). The same conclusions are reached using only the forced-choice form and using the WLS analysis of the polychoric correlations (ML′ and WLS in Table 5). Thus there is no support for either hypothesis for the anomia items.

Difficulties were encountered in estimating a single-factor model of self-esteem, such as negative estimates of unique variances. The pattern of covariances (Table 3) suggests that a two-factor solution,

TABLE 6: Estimated Parameters of Measurement Model for Anomia

	β_i	$\sigma^2_{\delta_i}$	R^2_i
Standard			
LIVETOD	.470	.977	.184
CHILDFUT	.659	.586	.425
COUNTON	.632	.763	.344
PUBOFF	.534	.897	.241
σ^2_ξ	1.000		
	GFI= .999	AGFI= .995	
Filtered			
LIVETOD	.315	.875	.102
CHILDFUT	.412	.693	.197
COUNTON	.624	.781	.333
PUBOFF	.648	.748	.359
σ^2_ξ	1.000		
	GFI = .999	AGFI = .995	
	Chi-square = 1.09	$df = 4$	$p = .896$

NOTE: $p < .05$ for all parameter estimates. GFI and AGFI are the goodness-of-fit index and the adjusted goodness-of-fit index, respectively.

consisting of one factor for the two positively worded items (WORTH and DOTHING) and one factor for the two negatively worded items (RESPECT and USELESS), might be more valid. (Note that the two highest covariances are those for the two positively worded items and the two negatively worded items, on both the standard and filtered forms.) A two-factor model for self-esteem had also been found by Alwin and Jackson (1981). Difficulties were also encountered in estimating a model with two correlated factors; the solution did not converge after 500 iterations and a negative error variance resulted.[3] These problems were eliminated, however, by constraining the two factor loadings to be equal on each factor, and the goodness-of-fit test indicated the model could not be rejected (chi-square = 10.56, $df = 6$, $p = .103$). When the covariance between the two factors was fixed at unity to create the equivalent of a single-factor model, the chi-square test showed a significant decrease in the fit (not shown). We thus conclude that a two-factor model is appropriate for the self-esteem questions. The fit was significantly improved, however, by allowing the errors for Items 1 and 4 (WORTH and USELESS) to be correlated on each form (change in chi-square = 8.54, $df = 2$, $p < .025$). The

TABLE 7: Estimated Parameters of Measurement Model for Self-Esteem

	β_i			
	ξ_1	ξ_2	$\sigma^2_{\delta_i}$	R^2_i
Standard				
WORTH	.457	.000	.199	.512
DOTHING	.457	.000	.378	.356
RESPECT	.000	.601	.722	.334
USELESS	.000	.601	.761	.322
$\sigma^2_{\xi_j}$	1.000	1.000		
$\sigma_{\xi_1\xi_2}$	.431			
$\sigma_{\delta_1\delta_4}$	−.063			
	GFI = .998	AGFI = .990		
Filtered				
WORTH	.382	.000	.271	.350
DOTHING	.382	.000	.464	.239
RESPECT	.000	.480	.793	.225
USELESS	.000	.480	.760	.233
$\sigma^2_{\xi_j}$	1.000	1.000		
$\sigma_{\xi_1\xi_2}$	.007			
$\sigma_{\delta_1\delta_4}$	.089			
	GFI = .998	AGFI = .990		
	Chi-square = 2.02	$df = 4$	$p = .732$	

NOTE: $p \leq .05$ for all parameters except $\sigma_{\xi_1\xi_2}$ the filtered group. GFI and AGFI are the goodness-of-fit index and the adjusted goodness-of-fit index, respectively.

estimated parameters and goodness-of-fit for this model are given in Table 7.

It is possible that a two-factor solution might be caused by acquiescence when there is a balanced scale composed of two positive self-esteem items and two negative self-esteem items (Mirowsky and Ross 1991). An alternative two-factor specification might be used involving a single bipolar self-esteem factor and an acquiescence factor. We have used the model shown in Table 7 because the two-factor solution also occurs when only the forced-choice form is used, a form on which acquiescence is not possible. Although response-order effects on the forced-choice form might also cause a two-factor solution, McClendon (1991) found no evidence of response-order effects for the self-esteem items.

The estimated factor loadings are higher on the standard form for each self-esteem factor (Table 7). The two factors are also significantly positively correlated on the standard form, whereas the covariance on the filtered form is near zero and nonsignificant. The estimated unique variances, on the other hand, are somewhat higher on the filtered form. All of these differences are consistent with the generalized-response hypothesis of filtering.

Constraining both the factor loadings and the covariance between the two factors to be equal across question forms significantly reduced the goodness-of-fit (Table 5). (Constraining only the βs was not significant, $p = .11$, whereas constraining the covariance of the factors was significant.) This is consistent with the prediction that common variance would be reduced on the filtered form, and is also consistent with the previous results showing smaller covariances between the items on the filtered form. The same results occurred when only the forced-choice form was analyzed, although the decreased fit was only marginally significant ($p < .10$), possibly due to smaller sample size. Constraining the βs to be equal across forms when using WLS also significantly reduced the fit.

Thus we conclude that on the standard form there is a general self-esteem factor underlying the positive and negative self-esteem dimensions, as evidenced by the significant positive covariance between these two dimensions. The general factor, however, has been eliminated by the filtered form. This, in conjunction with the possibility of smaller factor loadings on the filtered form results in smaller common variance components for the filtered items and lower reliability on the filtered form.

Constraining the unique variance components of the items to be equal on the standard and filtered forms does not significantly reduce the goodness-of-fit of the model (Table 5). Thus, even though the estimated unique variances are higher on the filtered form, as predicted by the generalized-response hypothesis, we cannot conclude that these differences are reliable.

CONCLUSIONS

We have found absolutely no support for the conventional view—which we have labeled the random-response hypothesis—that no-opinion filters improve the reliability of measurement. None of the three sets of items (lawyers, anomia, or self-esteem) had greater common variance, covariances, or reliability on the filtered form. These results suggest that Converse (1964) and others may have overestimated the amount of random responding that was occurring in public opinion surveys and/or neglected the presence of common factor responders.

Our results, however, have indicated some support for the generalized-response hypothesis of filtering. That is, for the self-esteem items there is more common variance on the standard form than on the filtered form, thus indicating greater reliability for this set of items on the standard form. We should be cautious, however, about concluding that the filtered form may reduce reliability because this outcome occurred for only one set of items and might represent a Type 1 error. On the other hand, Alwin and Krosnick's (1991) findings also indicated that the filtered form sometimes reduced reliability but never increased reliability. Thus our experimental design and cross-sectional estimates of reliability produced similar conclusions to those of Alwin and Krosnick (1991), which were based on a nonexperimental design and longitudinal estimates of reliability.

It is probably much more frequently assumed that the standard form pressures persons without true attitudes, or with less meaningful opinions than other respondents, into expressing an opinion than it is assumed that the filtered form discourages the expression of real attitudes. This research provides no support for the conventional view, and thus our findings do not indicate any support for using the filtered form of questions. Although reliability is not the only criterion by which the filtered form might be evaluated, it does suggest that the burden of proof should be on those who argue for the merits of opinion filters.[4] Even though we provide only limited evidence that filtering might reduce reliability, it is obvious that it can drastically reduce the effective sample size, and thus reduce the efficiency of estimates of population parameters.

It should be stressed, however, that the hypotheses tested in this article assume that any effects of filtering on reliability come about because the composition of the sample included in the analysis changes. This assumption is much different than another possibility, which is that the same individuals might respond either more reliably or less reliably to the wording of the filtered form than to the wording of the standard form.

Our conclusion that the reliability of the filtered items is not greater than the reliability of the unfiltered items is restricted to the subsamples of respondents who do not give any "don't know" responses to the four items in a scale. It is possible, however, that the DK responses given by floaters on the filtered form might be very reliable. If so, then the reliability of all responses on the filtered form (both substantive and DK responses) might be greater than the reliability of all responses on the standard form. The reliabilities of all responses cannot be estimated with our LISREL models because the "don't know" response is a nominal category that cannot be ordered among the other responses. Furthermore, the reason for using a filtered form of a set of related items is to separate nonattitudes from attitudes so that only the latter will be included in the measure of the construct. Thus the fundamental question is whether the substantive responses on the filtered form are more reliable than the substantive responses on the standard form. This is the question that we have addressed.

We must again add the caveat that we are not able to separate nonrandom item-specific variance from random variance. Thus we are probably underestimating reliability by some unknown amount. Whether this potential bias in our estimation of reliability biases our inference about the difference in reliability between the standard and filtered forms we cannot say. But it should also be emphasized that our method of estimating reliability, by decomposing variance into common and unique components, is consistent with the conventional method used with cross-sectional survey data. Therefore, it is of interest to know how the standard versus the filtered form of questions affects the conventional estimates of reliability.

Why is the set of items with the smallest percentage of floaters the one where we find the greatest effect of the filter on our reliability estimates? As has been discussed, no-opinion responses are less frequent on the self-esteem items (Table 1). Presumably this is because respondents are more likely to feel that they have sufficient information to answer questions about themselves than they are to feel that they have sufficient information to answer questions about lawyers, for example. As we suggested in our theory section, there may be two types of respondents who are susceptible to filter effects for any set of items, random responders and generalized responders. Whether one finds support for the random-response hypothesis or the generalized-response hypothesis may depend on their relative proportions in the population, and this relative frequency may vary from topic to topic. A finding of no difference in reliability (such as for the lawyers and anomia items) may indicate that there are relatively equal numbers of each type of floater.

One factor that may explain which type of floater is more numerous is that as specific questions become more difficult to answer because more individuals do not have the expertise to answer them, the proportion of floaters who are random responders may increase relative to the proportion who are generalized responders. Thus the susceptible respondents on the lawyer items may be more equal in numbers in terms of random and generalized responders, as compared to the self-esteem items where generalized responders may be disproportionately represented among them. Thus the information difficulty of the questions may help to predict when filtering will decrease reliability. This is proposed as a hypothesis meriting further investigation.

NOTES

1. If consideration of the specifics leads to a pattern of responses that is congruent with the general factor, then such responses are included in the common factor variance instead of in $\sigma^2_{\upsilon i}$.

2. As we noted earlier, this is a lower-bound estimate of reliability. See Alwin (1989) for a discussion of this problem.

3. These problems also occurred when analyzing the forced-choice form. The negative error variances were not significant in either the single-factor model or the two-factor model. When the variances were fixed at zero in an attempt to achieve an admissible solution, other negative variances occurred, for both the one- and two-factor models. Thus this technique of fixing an inadmissible solution was not successful in this case.

4. The filtered form might also be evaluated, for example, in terms of its ability to increase validity. McClendon (1991), however, found little or no evidence that use of the filtered form would reduce two types of invalidity, that is, acquiescence and response-order effects.

REFERENCES

Alwin, Duane F. 1989. "Problems in the Estimation and Interpretation of the Reliability of Survey Data." *Quality and Quantity* 23:277-331.

———. 1991. "Research on Survey Quality." *Sociological Methods & Research* 20:3-29.

Alwin, Duane F. and David J. Jackson. 1979. "Measurement Models for Response Errors in Surveys: Issues and Applications." Pp. 68-119 in *Sociological Methodology 1980*, edited by K. F. Schuessler. San Francisco: Jossey-Bass.

———. 1981. "Applications of Simultaneous Factor Analysis to Issues of Factorial Invariance." Pp. 249-79 in *Factor Analysis and Measurement in Sociological Research*, edited by D. J. Jackson and E. F. Borgatta. London: Sage.

Alwin, Duane F. and Jon A. Krosnick. 1991. "The Reliability of Attitudinal Survey Measures: The Role of Question and Respondent Characteristics." *Sociological Methods & Research* 20:139-81.

Andrews, Frank M. 1984. "Construct Validity and Error Components of Survey Measures: A Structural Modeling Approach." *Public Opinion Quarterly* 48:409-42.

Bishop, George F., R. W. Oldendick, and A. J. Tuchfarber. 1983. "Effects of Filter Questions in Public Opinion Surveys." *Public Opinion Quarterly* 47:528-46.

Bogart, Leo. 1967. "No Opinion, Don't Know, and Maybe No Answer." *Public Opinion Quarterly* 31:331-45.

Bollen, Kenneth. 1989. *Structural Equations With Latent Variables*. New York: Wiley.

Converse, Jean M. 1976-1977. "Predicting 'No Opinion' in the Polls." *Public Opinion Quarterly* 40:515-30.

Converse, Phillip E. 1964. "The Nature of Belief Systems in the Mass Public." Pp. 206-61 in *Ideology and Discontent*, edited by D. D. Apter. New York: Free Press.

Grice, H. P. 1975. "Logic and Conversation." Pp. 41-58 in *Syntax and Semantics: Speech Acts*, vol. 3, edited by P. Cole and J. L. Morgan. New York: Academic Press.

Heise, David R. and George W. Bohrnstedt. 1970. "Validity, Invalidity, and Reliability." Pp. 104-38 in *Sociological Methodology 1970*, edited by E. F. Borgatta. San Francisco: Jossey-Bass.

Hippler, Hans-J. and Norbert Schwarz. 1989. "No Opinion Filters: A Cognitive Perspective." *International Journal of Public Opinion Research* 1:77-87.

Jöreskog, Karl G. and Dag Sörbom. 1989. *LISREL 7 User's Reference Guide*. Mooresville, IN: Scientific Software, Inc.

Lord, F. M. and M. R. Novick. 1968. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley.

McClendon, McKee J. 1986. "Unanticipated Effects of No-Opinion Filters on Attitudes and Attitude Strength." *Sociological Perspectives* 29:379-95.

———. 1991. "Acquiescence and Recency Response-Order Effects in Attitude Surveys." *Sociological Methods & Research* 20:60-103.

Mirowsky, John and Catherine E. Ross. 1991. "Eliminating Defense and Agreement Bias from Measures of the Sense of Control: A 2 × 2 Index." *Social Psychology Quarterly* 54:127-45.

Rosenberg, Morris. 1965. *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Srole, Leo. 1956. "Social Integration and Certain Corollaries: An Exploratory Study." *American Sociological Review* 30:709-16.

Schuman, Howard and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys*. New York: Academic Press.

McKee J. McClendon is a professor of sociology at the University of Akron. He received his Ph.D. in sociology from the University of Kansas. From 1982-1988 he served as the director of the Akron Area Survey, an annual telephone survey of Summit County, Ohio. He currently teaches

a course on cognitive aspects of survey methodology in the University of Michigan's Summer Institute in Survey Research Techniques. His recent research has focused on formal aspects of questionnaire construction such as question-order effects, response-order effects, no-opinion filters, acquiescence, and mode of administration. He is the author of Multiple Regression and Causal Analysis *(Peacock, forthcoming).*

Duane F. Alwin is a professor of sociology and the program director of the Survey Research Center at the Institute for Social Research, University of Michigan, where he directs the SRC's Summer Institute in Survey Research Techniques. In addition to his interests in survey methods and their use, he is mainly interested in the family, socialization, and the relationship between human development, society, and social change.