Estimation of Measurement Error and Identification of Causes: Linking Measurement Error to Nonresponse, Interviewers, and Interviewer Characteristics

Andy Peytchev, University of Michigan
Institute for Social Research, 426 Thompson St., Ann Arbor, MI 48104

Keywords: Measurement Error, Unit Nonresponse, Interviewer Effects, Two-stage Model

1 Abstract

In order to identify survey design factors that induce measurement error, validation data are needed. Gold standards are difficult to obtain and are of limited types. In the absence of truth, analyses commonly revert to differences in means, such as the Kish and Hansen, Hurwitz, and Bershad models for interviewer variance, and similar ANOVA models for comparisons of response scales.

Proposed is a two-stage iterative approach to estimating and modeling causes of measurement error – the measurement error is first estimated through a predictive model and then it is modeled using hypothesized causes.

This two-stage model is demonstrated with data from the 2002 National Election Study to examine interviewer-induced measurement error, and its relationship to proxies for unit nonresponse. Effects of interviewer experience and characteristics on measurement error are estimated. Results show that measurement error variance induced by the interviewers is quite unrelated to interviewer variance that is defined as differences in their means. Furthermore, estimates of differences in their means and their standard errors are biased when interviewer measurement error variances are assumed to be the same across interviewers. Findings are consistent with prior studies of interviewer effects, as the measurement error variance was more than four times larger for black respondents interviewed by white interviewers than when interviewed by black interviewers, on thermometer ratings for blacks. The estimated measurement error was also found to be positively associated with indicators of unit nonresponse.

2 Introduction

Interviewers act as clustering agents, affecting the collected survey data. Studies have shown that responses collected by the same interviewer are more similar than the responses collected by all interviewers, whether estimated through interpenetration of interviewers to sample assignments in face to face surveys (e.g., O'Muircheartaigh and Campanelli, 1998) or interviewers in centralized telephone facilities (e.g., Groves and Magilavy, 1980). A common estimator is the Intraclass Correlation Coefficient (ICC), which is the proportion of the total variance in a survey variable that is attributed to differences in the interviewers. To the extent that the differences in means across interviewers are large relative to the total variance, the variance of a particular statistic is inflated.

This model has been instrumental in the evaluation of interviewer variance due to its simplicity and ease of implementation to almost any survey data that has somewhat random assignment of respondents to interviewers.

However, there are significant deficiencies to this approach to modeling interviewer variance, as has been acknowledged by many authors. For example, interviewers who induce measurement error without altering the expected mean of the responses will not be identified as introducing error. The traditional models do not even attempt to identify particular interviewers. Furthermore, interviewers who induce random measurement error only make mean differences between interviewers even less likely to be detected. In the case of examining interviewer-induced survey error, some critical issues that need to be overcome can be summarized as:

1. Interviewer biases can go in the same direction, hence concealing interviewer error, as $x_j=x_k$ and $X_j=X_k$, but $x_j\neq X_j$, where $j\neq k$.
2. Interviewer variances can be very different, that is $\sigma_j\neq\sigma_k$. Groves and Magilavy (1980) note that this could have affected the precision of their estimates of interviewer variance, but it could also affect the bias of these estimates.
3. Ultimate interest is in estimating e_i and e_j, where e_i is the estimated error for respondent "i" and e_j is the estimated average error for interviewer "j."

This paper presents an approach that relaxes some of the common assumptions like equal interviewer variances, makes estimation of measurement error possible in nonexperimental settings, produces respondent- and interviewer-level estimates of measurement error, and has the ability to examine other covariates of measurement error.

3 Background

Measurement error induced by the interviewers is commonly estimated as the proportion of variance attributable to the interviewers – the between-interviewer variance relative to total variance. This is interviewer variance, with estimators originally constructed by Kish (1962) and Hansen, Hurwitz, and

Bershad (1961). This is inevitable in most surveys as true values on the respondents for survey variables are not known, whether because this is the actual goal of the survey to collect this unavailable information, it is prohibitively expensive, the frame variables are different from the survey variables, or the variable of interest is not directly measurable, such as attitudes and beliefs.

Without an estimate for measurement error at the respondent level, analyses are limited to this type of estimators and cannot go beyond the identification of sources of variability, such as the proportion of between-interviewer variation that is explained by interviewer race, gender, or experience.

Hence one limitation of this approach is that the difference between interviewers is narrowly defined as difference in means. But many hypotheses about interviewer errors refer to differences in the variability of responses. For example, if some interviewers deviate more from the script, they may vary more in the responses they elicit, but the means may still be the same. At the very minimum, differences in within-interviewer variances have to be accounted for in the statistical estimation of interviewer effects, as estimators like Kish's ρ_{int} assume that they are the same.

Another limitation is that we would want to say that not only are more experienced and less experienced interviewers different, both in means and variances, but also which produce less measurement error. And which interviewers produce more measurement error could be of interest in terms of interviewer characteristics and behaviors, or in terms of identifying individual interviewers for retraining, for instance.

In the ideal circumstances, true values for key variables of interest would be known for every respondent, and interviewers would be randomly assigned to respondents. While the latter can be achieved to some degree in centralized telephone survey centers, especially if accounting for the shifts in which particular interviewers work, truth is seldom known for key variables as is usually the reason for conducting the survey. This leads to the use of ANOVA and multilevel ANOVA models for estimating interviewer variance.

If measurement error could be estimated for each respondent, the potential causes and their interactions could be examined through multivariate models. Using data from 1942, Hyman and colleagues (1954) found that interviewer race is associated with mean differences in responses by black respondents, Schuman and Converse later examined on white respondents (1971), and a test of the respondent interviewer race interaction by Hatchet and Schuman (1975). Most of the studies on race of interviewer use racial attitudes as outcome measures.

These effects are highly dependent on how related the variable of interest is to the interviewer characteristic. Similarly to race, researchers have examined the effect of respondent and interviewer gender on gender-related variables, again finding effect on mean responses (Kane and Macaulay, 1993). Like the studies on race, the most forwarded explanation is social desirability in the interaction between the interviewer and respondent. If other mechanisms also act to produce these observed mean differences, no inference could be made about measurement error - what combinations are preferable, as this is not an error estimate. Furthermore, none of these studies have examined distributional differences. Subgroup variances affect the standard errors of the estimates, and can also affect regression coefficients and their standard errors through attenuation (measurement error reducing the coefficients) or inflation (correlated measurement error increasing the coefficients).

Testing these interviewer effects in a telephone survey, especially race, could provide a conservative test as the race of the interviewer or the respondent is not always correctly identified by other party (Davis, 1997), and because the social distance is not the same as in face to face interviewing.

Incentives have long been feared to affect measurement error, yet that fear likely stems from the operationalization of measurement error – mean differences between those who receive incentives and those who do not. While an argument that incentives induce better responses can and has been forwarded, in studies where the amount of the incentive is contingent on the degree of cooperation it needs to be included at least as a necessary control.

There are further such links that could be made between nonresponse and measurement error through other measures of the respondent's level of cooperation. Such evaluations by interviewers in an earlier survey could be used as independent observations linked to measurement error in a subsequent survey administration.

To the extent that survey weights are dominated by nonresponse adjustments, comparison of unweighted and weighted estimates of measurement error should provide a multivariate approach to examining the link between nonresponse propensities and measurement error.

The pace of the interview can also affect measurement error, as respondents may feel rushed by the interviewer and fail to provide thoughtful answers, as argued by Cannell and colleagues (1981).

A correlate of pace is interviewer experience, both within a single survey cycle, and with the survey organization (Olson and Peytchev, 2004), while experience has been linked to lower reporting of drug use (Turner, Lessler and Devore, 1992; Hughes, Chromy, Giacoletti and Odom, 2002). The question remains whether each of these types of experience lead to more or less bias and lower or higher measurement error variance.

When defined as variance rather than bias, measurement error has been linked to respondent characteristics. Andrews and Herzog (1986) found older respondents and less educated respondents to be more susceptible to the response scale used, and to exhibit higher residual variances in a substantive structural equations model. However, the estimation for these residual variances did not account for differences in means, the approach could only examine error variances by one or two variables at most as a separate model was fit for each cell of the independent variable, the estimated residuals are based on inherently different models as the factor loadings (regression coefficients between the latent and observed variables) were allowed to be different across the models, and measurement error was a simple average across all the residuals in each model.

4 Data and Methods

First, measurement error is estimated through a substantive model. For a variable such as the respondent's thermometer rating of blacks, covariates need to be selected to maximize the model's explanatory ability. These could include measures of the respondent's personality, other behaviors, and even measures of social desirability. Ideally, this model would have been used in the construction of the questionnaire. This is the "Mean" model as it predicts respondents' means, and provides a residual – an estimate of measurement error for each respondent under the model. Implicitly, the better the model in terms of including relevant covariates and predictive ability, the better the estimate of measurement error will be.

Second, the estimated measurement error could then be modeled as a function of various hypothesized causes. Rather than concluding that matching the respondent's race to that of the interviewer, this allows inferences about which condition produces the least measurement error. This is the Variance model as it models the error term in the Mean model. It could include all, some, or none of the covariates used in the Mean model. Rather than simply dealing with different variances with methods for heteroscedasticity, this approach allows modeling of the heteroscedasticity achieving interpretable parameters of interest.

Third, there is an inherent interdependence between the estimated means and the estimated variances. For example, black interviewers could be associated with higher ratings of Colin Powell, but a smaller error variance. That is, a larger bias based on the assumption of social desirability, but smaller measurement error variance. Statistically, the estimate of the measurement error is dependent on the estimated mean and vice versus. Therefore, the two models need to be estimated simultaneously. One possible solution is to do this iteratively, so that the Mean model is estimated, the estimates are used in the Variance model, then those estimates are used to go back to re-estimate the mean model, and so forth, using an optimization method such as the Lagrange Multiplier. In industrial engineering and quality control dealing with optimization of production processes, these are called Mean-Variance models (regardless of whether an iterative procedure is implemented).

Formalizing these objectives, the Mean model is the familiar linear regression model:

$$y_i = \beta' X_i + \delta_i \qquad (1)$$

Where y_i is the response from respondent i, X is a vector of covariates, α and β are the intercept and parameter estimates, and the residual is $\delta_i \sim N(0,\sigma^2)$.

Note that the β coefficients here are the biases that are typically examined when looking for the effect of a factor hypothesized to be affecting the variable Y. If this factor was interviewers, one would obtain beta coefficients for each interviewer and the R2 of this model would represent the proportion of variance that is attributable to the interviewers. Similarly, interviewer characteristics could be used, and the effect of gender of interviewer, for instance, could be estimated on the mean predicted Y. This is the same as the ANOVA approach, focusing on the means, assuming the error variance is constant ($\text{cov}(\beta,\delta)=0$), and failing to model causes of the error, δ.

The model for heterogeneity (differences in variance) as proposed by Cook and Weisberg (1982):

$$\ln(\text{var}(\delta_i)) = \lambda' Z_i + \varepsilon_i \qquad (2)$$

Where Z is a vector of covariates that typically includes the variables in X but does not have to, and it could include additional variables.

An iterative procedure then repeatedly fits (1) and (2), until a convergence criterion is satisfied.

This method is demonstrated using data from the 2002 National Election Study (NES). This NES is the first one to include a pre- and post-election survey in a non-presidential election year and both interviews done by telephone. Eighty-eight percent of the sample is the panel component drawn as an area probability sample, the remainder is a list-assisted RDD sample. In the pre-election survey, 1511 responded (55.8% response rate), and 1346 of them responded to the post-election (89.1%).

First, a thermometer rating (i.e., 0-100 scale) for blacks is regressed on the selected substantive predictors from the pre-election interview, respondent characteristics, interview length, interviewer's interview number, whether it was the same interviewer in the pre- and post-election interviews, and interviewer identifier. Then the same model is fit to the log of the error variance. A plot of the interviewer parameter estimates and their standard errors from each model will help examine whether there is any link between the interviewer bias estimates and the interviewer measurement error variance estimates, as well as reveal their relative magnitudes.

The next step is to fit a model that instead of using interviewer identifier, uses interviewer characteristics that are either expected to or at least have been found to explain some interviewer differences in the past, such as interviewer experience, race, gender, age, education, and languages. To test hypotheses about the effect of race on measurement error to the blacks thermometer rating, race of the respondent and the interviewer are fully cross-classified and entered as an interaction effect. Additional interviewer observations from the pre-election interview are also used in the measurement error model, such as the respondent's level of cooperation, sincerity, and interest in politics, allowing linkage of measurement error and potential causes, and further links between measurement error and indicators of nonresponse. This model also included an indicator for whether the same interviewer did the pre- and post-interview. While interviewer matching in multiple-wave studies is done primarily for increasing response rates, even if it does not affect measurement error it is still needed in the model as interviewers may induce correlated measurement error across responses in different waves.

Measurement error may be differential for subgroups in the population. For example, sampled individuals with low response propensities may exhibit higher interviewer effects for a variety of reasons, such as the topic being sensitive to some respondents and hence inducing measurement error and likelihood of nonresponse. To have an estimate of measurement error that is focused on its effect on population estimates rather than just the mechanisms at the respondent and interviewer levels, these models are re-estimated with population weights. These weights have three components: selection probability weight, nonresponse weight, and post-stratification weight. A second objective of estimating the models with survey weights is that a comparison between the unweighted and the weighted results would provide further insight into the link between nonresponse and measurement error, to the extent that population weights are mostly an adjustment for nonresponse.

The mean model is also estimated without the variance model to demonstrate differences in the linear regression parameter estimates when heterogeneity of variances is ignored, which would imply biased estimates of the intra-class correlation coefficient (ICC) or the proportion of variance attributable to the interviewers.

The analyses are repeated with thermometer ratings towards feminists, in which relevant substantive predictors are added, and instead of interviewer and respondent race being estimated as an interaction effect, it is interviewer and respondent gender.

5 Results

Figure 1 shows the interviewers who did at least 5 post-election interviews,[1] by their estimated bias and measurement error variance. The vertical and horizontal lines from each interviewer are the confidence intervals for the bias and error variance, respectively. Interviewers were not a significant correlate in either model[2] and the standard errors for the interviewer estimates for bias and measurement error variance are relatively similar. However, an important observation is that the interviewer bias and error variance estimates are not highly correlated - only 6.7% of the variability in interviewer measurement error variance is explained by interviewer bias.

Table 1 presents the model in which the thermometer rating towards blacks is the dependent variable.[3] The mean-only model provides the estimates that would be achieved through a classical regression approach that assumes a single homogenous error variance. Some of the substantive variables measured in the pre-election interview: interest in politics, thermometer ratings for Bill Clinton, Colin Powell, and Jesse Jackson, and poor people's chance for a fair trial. Respondent gender and income were also significant in this model. Note that none of the interviewer and interview characteristics were significant at the .05 level, nor was the respondent and interviewer race.

Most of the regression coefficients in this linear model increased when heterogeneity of the error variances was accounted for in the mean-variance model, with their standard errors changing slightly in either direction. The mean-variance model fit marginally better than the mean-only model ($\chi^2_{(46)}$=62, p=.070). More notably, some of the key expected mean differences became significant – interviewer and respondent race interactions, and whether the same interviewer conducted the pre- and post-election interviews.

The key results are yet in the measurement error variance part of the mean-variance model. The respondent's level of cooperation rated by the

---

[1] This criterion is used in order to provide somewhat stable interviewer-level estimates of both point and error variance estimates.

[2] Although interviewer was significant in the mean model when only interviewers with 10 or more interviews were kept, rather than 5 or more.

[3] Given the focus on interviewer effects, interviewers who did only one interview were excluded – there were only three such interviews.

interviewer in the pre-election survey was negatively correlated with the measurement error estimate – with every point lower rating, the measurement error variance increases by 46% (1/.687). Among the respondent characteristics, female respondents were predicted as having 31% more measurement error variance. Younger interviewers (under 35 years old) were estimated as inducing 35% more measurement error variance than their older colleagues, even though the model controlled for interviewer experience with the survey organization and experience during this survey.

As expected, interviewer and respondent race had an interaction effect on the rating of blacks. When blacks were interviewed by white interviewers, the measurement error variance was 420% (1.306/.311) greater than when they were interviewed by black interviewers.

Some of these correlates of measurement error variance also interact with nonresponse, to the extent to which post-survey adjustments are dominated by nonresponse. The interviewer's rating of respondent cooperation in the pre-election interview was not a significant predictor when the model was estimated without survey weights, and similarly for the race of the interviewer and respondent.

In the similar model in which the thermometer rating of feminists is the dependent variable and the key interaction is gender of the respondent and of the interviewer, the interaction effect on measurement error variance was not significant.[4] However, there was the tendency for both males and females to provide lower thermometer ratings for feminists when interviewed by males, an effect that was much larger when the equal error variances assumption was relaxed.

6 Discussion and Conclusions

A method for the estimation and modeling of measurement error was presented. The presented method was shown to provide:

1. The ability to obtain estimates for causes of *measurement error bias*, such as differences between interviewers or survey protocols, *without the assumption of equal variances*.
2. Estimates for how these and other causes affect *measurement error variance*.
3. The ability to *link* effects on bias and on error variance from a particular design characteristic, incorporating the conditionality of the two measures.
4. Estimates that can be used to evaluate which designs and conditions are *preferable,* that can be used during and post data collection.

Estimates of measurement error bias and their standard errors changed when the assumption of equal variances was relaxed, and in the presented example the estimated interviewer biases were larger when this assumption was relaxed.

The use of differences in biases as a proxy for total measurement error was not supported – there was almost no association between measurement error bias and measurement error variance.

In addition to being a different quantification of measurement error, the estimate of measurement error as a residual in a substantive model provides guidance on what design characteristics and conditions lead to less measurement error, not just differences. For example, black respondents exhibited far less measurement error when interviewed by black interviewers than by white interviewers, to a thermometer rating for blacks. This is in line with prior literature that measurement error will be found when the cause is related to the variable of interest, such as race and gender, but also provides guidance on what conditions are preferable.

Similarly, there was a tendency for both male and female respondents to rate feminists higher when interviewed by female interviewers, but the effect on measurement error variance was not found in this model.

These estimates of measurement error also allow the examination of the link between measurement error and nonresponse. Higher respondent cooperation rated by an interviewer in a previous wave of the survey was found to be associated with less measurement error. This link and its direction were further supported through the survey weights – some of the predictors of measurement error that were significant in the weighted analysis were not significant in the unweighted analysis. To the degree that the interviewer rated respondent cooperation and that the survey weights are dominated by adjustments for nonresponse, these findings support that respondents with lower propensities to respond provide more measurement error.

The goal of this paper was to present the methodology and provide a limited empirical test. As with all research, findings are subject to replication. Future replications should also attempt to incorporate complex survey design in the variance estimation procedures.

There are limitations to this approach to estimation of measurement error. These estimates are only as good as the specification and explanatory power of the mean model, a weakness of the current demonstration. Like models for every other source of survey error, these are statistic-specific, as even here it was shown to be differentially effective for different variables. Some may only use it for behavioral frequency questions and avoid making the argument

[4] Results available upon request.

that attitudes and opinions could have true values, which is beyond the scope of this paper. And yet others could use only the parameter estimates from the mean model to produce the well-accepted estimates of interviewer variance that at least adjust for respondent characteristics and for heterogeneity of interviewer variances - relax the assumption of random assignment of respondents to interviewers and the assumption of equal interviewer variances.

Having a respondent-level estimate for measurement error could be invaluable, especially in the absence of validation data. However, further research is needed to: (1) build better substantive models for key variables, so that these measures could be incorporated into surveys, (2) perform simulation studies to determine the required strength of associations between variables that would produce meaningful residuals, (3) identify other correlates of measurement error, and (4) further develop the statistical models to include various distributions, link functions, and complex survey designs.

7 References

Andrews, F. M., and A. R. Herzog. (1986). "Respondent Age and Survey Measurement Error." Journal of the American Statistical Association **81**:403-10

Cannell, C., P. Miller, et al. (1981). "Research on Interviewing Techniques." Sociological Methodology: 389-437.

Cook, R. D. and S. Weisberg (1982). Residuals and Influence in Regression. New York, Chapman and Hall.

Davis, D. W. (1997). "Nonrandom Measurement Error and Race of Interviewer Effects among African Americans." Public Opinion Quarterly **61**: 183-207.

Groves, R. M. and L. Magilavy (1980). Estimates of Interviewer Variance for Telephone Surveys. Proceedings of the Survey Research Methods Section of the American Statistical Association.

Hansen, M. H., W. N. Hurwitz, et al. (1961). "Measurement errors in censuses and surveys." Bulletin of the ISI **38**: 351-374.

Hatchet, S. and H. Schuman (1975). "White respondents and race of interviewes effects." Public Opinion Quarterly **39**: 523-528.

Hughes, A., J. Chromy, et al. (2002). Impact of Interviewer Experience on Respondent Reports of Substance Use. Redesigning an Ongoing National Household Survey. J. Gfroerer, J. Eyerman and J. Chromy. Washington, DC, Substance Abuse and Mental Health Services Administration**:** 161-184.

Hyman, H. H., W. J. Cobb, et al. (1954). Interviewing in Social Research. Chicago, University of Chicago Press.

Kane, E. and L. Macaulay (1993). "Interviewer Gender and Gender Attitudes." Public Opinion Quarterly **57**: 1-28.

Kish, L. (1962). "Studies of Interviewer Variance for Attitudinal Variables." Journal of the American Statistical Association **57**: 91-115.

O'Muircheartaigh, C. and P. Campanelli (1998). "The Relative Impact of Interviewer Effects and Sample Design Effects on Survey Precision." Journal of the Royal Statistical Society **161**: 63-77.

Olson, K. and A. Peytchev (2004). Interviewer Experience, Interview Order, and Interviewer Behaviors and Attitudes. Paper presented at the Joint Statistical Meetings, American Statistical Association, Toronto, Canada.

Schuman, H. and J. Converse (1971). "The Effects of Black and White Interviewers on Black Responses in 1968." Public Opinion Quarterly **35**: 44-68.

Turner, C. F., J. T. Lessler, et al. (1992). Effects of Mode of Administration and Wording on Reporting of Drug Use. Survey Measurement of Drug Use: Methodological Studies. C. F. Turner, J. T. Lessler and J. Gfroerer. Rockville, MD, National Institutes of Health.

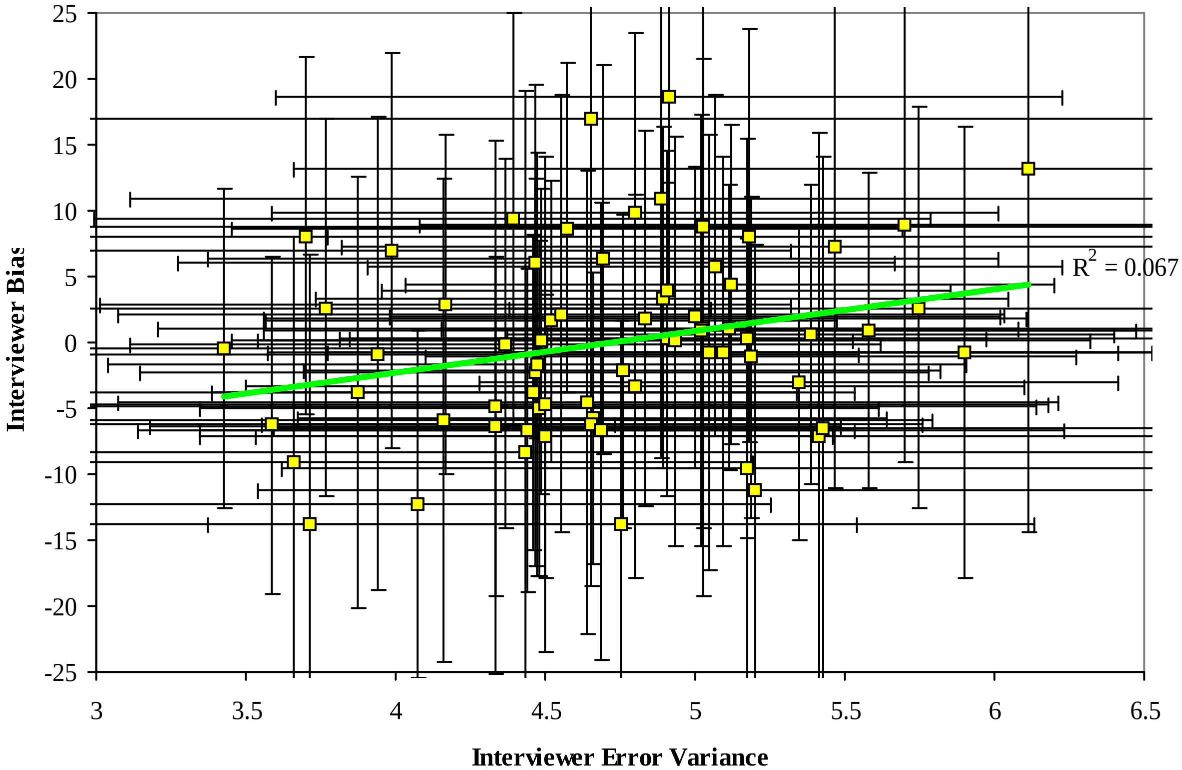

Figure 1: Mean Bias and Mean (Natural Log of) Measurement Error Variance for Each Interviewer.
* Includes only interviewers with at least 5 interviews, n=770
** For this variable ρ_{int}=.0007

Table 1: Weighted Bias and Measurement Error Models for Thermometer Rating for Blacks (n=805).

Variable	Category	Mean-Only Model Means Param. Estimate	Mean-Only Model Means Std Error	Mean-Only Model Means Approx Pr > |t|	Mean-Variance Model Means Param. Estimate	Mean-Variance Model Means Std Error	Mean-Variance Model Means Approx Pr > |t|	Mean-Variance Model Variances Param. Estimate	Mean-Variance Model Variances Std Error	Mean-Variance Model Variances Approx Pr > |t|	Var. Ratio
Intercept		50.953	12.560	<.0001	50.327	11.717	<.0001	5.748	0.017	<.0001	313.596
Amount of Prepaid Incentive	$20	0.223	1.698	0.896	-1.188	1.633	0.467	0.028	0.403	0.945	1.029
	$40	-1.570	4.169	0.707	-3.196	3.177	0.314	-0.484	0.569	0.395	0.616
	$50	0.000	--	--	0.000	--	--	-0.141	0.454	0.756	0.868
Substantive Responses											
Interest in Politics	Very much interested	5.764	2.339	0.014	7.176	2.284	0.002				
	Somewhat interested	1.553	1.697	0.360	1.518	1.638	0.354				
	Not much interested	0.000	--	--	0.000	--	--				
George Bush	Thermometer 0-100	0.064	0.036	0.078	0.069	0.037	0.063				
Bill Clinton	Thermometer 0-100	-0.065	0.030	0.030	-0.082	0.030	0.007				
Colin Powell	Thermometer 0-100	0.160	0.042	0.000	0.183	0.041	<.0001				
Jesse Jackson	Thermometer 0-100	0.134	0.034	<.0001	0.163	0.033	<.0001				
Do the poor have same chance at fair trial as the wealthy?	Agree	-5.218	1.942	0.007	-4.591	1.914	0.016				
	Neither agree/disagree	2.001	2.518	0.427	2.476	2.611	0.343				
	Disagree	0.000	--	--	0.000	--	--				
Is rich/poor gap in US larger/smaller than 20 yrs ago?	Larger	-1.223	1.876	0.515	-1.729	1.928	0.370				
	Smaller	-0.973	2.771	0.726	-4.028	2.865	0.160				
	About the same	0.000	--	--	0.000	--	--				
How important is the tax cut issue to you?	Scale 1-3	-0.032	0.566	0.955	-0.316	0.567	0.577				
Is religion an important part of your life?	Important	0.819	1.749	0.639	0.046	1.687	0.978				
	Not important	0.000	--	--	0.000	--	--				
Interviewer Observations in Pre-election Survey											
R cooperation	Scale 1-5							-0.376	0.136	0.006	0.687
R info. about politics	Scale 1-5							-0.070	0.105	0.508	0.933
R apparent intelligence	Scale 1-5							0.216	0.118	0.067	1.241
R interest in the I'w	Scale 1-5							-0.138	0.094	0.144	0.871
How sincere did R seem to be in his/her answers	Completely sincere							0.000	--	--	1.000
	Usually sincere							-0.004	0.216	0.987	0.996
	Often seemed sincere							0.241	0.954	0.801	1.272
Respondent Characteristics											
Age	18-34 years	-0.094	1.572	0.952	1.416	1.501	0.346	0.000	--	--	1.000
	35 or older	0.000	--	--	0.000	--	--	-0.212	0.137	0.121	0.809
Gender	Male	-3.836	1.421	0.007	-3.448	1.463	0.018	-0.267	0.128	0.036	0.765
	Female	0.000	.	.	0.000	--	--	0.000	--	--	1.000
Education	High school or less	-1.225	1.474	0.406	-1.880	1.509	0.213	0.000	--	--	1.000
	College degree	0.000	--	--	0.000	--	--	0.152	0.124	0.222	1.164
Religion	None	-1.888	2.258	0.403	-0.800	2.176	0.713				
	Provided	0.000	--	--	0.000	--	--				
Income	Refused/Don't Know	-5.026	7.501	0.503	-4.640	4.798	0.334	-0.973	0.633	0.124	0.378
	$0 -$14,999	-2.185	2.855	0.444	0.010	3.116	0.997	0.124	0.257	0.631	1.132
	$15,000-$34,999	2.862	1.750	0.102	3.772	1.652	0.022	-0.020	0.156	0.900	0.981
	$35,000-$49,999	3.955	1.948	0.042	4.874	2.094	0.020	0.210	0.173	0.226	1.233
	$50,000+	0.000	--	--	0.000	--	--	0.000	--	--	1.000
Interviewer Characteristics											
Age	18-34 years	0.686	1.645	0.677	0.711	1.550	0.646	0.000	--	--	1.000
	35 or older	0.000	--	--	0.000	--	--	-0.299	0.143	0.037	0.741
Gender	Male	2.260	1.454	0.120	2.494	1.432	0.082	-0.029	0.126	0.821	0.972
	Female	0.000	--	--	0.000	--	--	0.000	.	.	1.000
Education	High school or less	-1.209	3.240	0.709	-0.786	3.338	0.814	-0.061	0.274	0.825	0.941
	Some college	0.328	1.524	0.830	-0.124	1.519	0.935	0.000	--	--	1.000
	College degree	0.000	--	--	0.000	--	--	-0.019	0.140	0.894	0.981
Languages	No other languages	-0.497	1.752	0.777	-0.312	1.689	0.853	0.000	--	--	1.000
	Spanish	-0.343	2.443	0.888	0.340	2.611	0.896	0.022	0.204	0.914	1.022
	Other	0.000	--	--	0.000	--	--	-0.192	0.155	0.215	0.825
Experience	None	-1.622	2.938	0.581	-1.084	3.069	0.724	-0.010	0.169	0.954	0.990
	1 Year or less	-4.171	2.581	0.106	-3.479	2.762	0.208	0.000	--	--	1.000
	More than 1 year	0.000	--	--	0.000	--	--	0.131	0.231	0.571	1.140
Length of Interview	Minutes	-0.043	0.123	0.724	-0.037	0.138	0.787	0.018	0.011	0.118	1.018
Interviewer's Interview	Sequential number	0.017	0.315	0.956	0.053	0.316	0.868	0.034	0.028	0.217	1.035
Pre- & Post-election Interviewer	Different	9.937	5.664	0.079	10.691	3.179	0.001	0.000	--	--	1.000
	Same	0.000	--	--	0.000	--	--	-1.353	0.567	0.017	0.258
Respondent by Interviewer Characteristics											
R Race*Iwer Race	Black*White	-3.625	9.221	0.694	-4.249	9.137	0.642	0.267	0.230	0.245	1.306
	Black*Black	-12.082	10.708	0.259	-15.569	9.102	0.087	-1.168	0.541	0.031	0.311
	Black*Other	-2.302	13.554	0.865	-0.795	16.293	0.961	0.460	0.880	0.601	1.584
	White*White	-9.573	8.886	0.281	-12.431	8.677	0.152	0.000	--	--	1.000
	White*Black	-14.284	9.357	0.127	-17.079	8.960	0.057	-0.304	0.265	0.252	0.738
	White*Other	-10.759	9.589	0.262	-11.397	9.313	0.221	-0.097	0.368	0.792	0.907
	Other*White	-5.017	9.085	0.581	-5.857	8.940	0.512	0.310	0.189	0.101	1.363
	Other*Black	-22.286	13.864	0.108	-26.751	11.483	0.020	-1.033	0.893	0.247	0.356
	Other*Other	0.000	--	--	0.000	--	--	0.143	0.765	0.852	1.153