Quality of Prisoner Self-Reports

Arrest and Conviction Response Errors

Kent H. Marquis
with the assistance of Patricia A. Ebener

Rand

Prepared under Grant Number 79-NI-AX-0055 from the National Institute of Justice, U.S. Department of Justice. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the U.S Department of Justice.

The Rand Publications Series: The Report is the principal publication documenting and transmitting Rand's major research findings and final research results. The Rand Note reports other outputs of sponsored research for general distribution. Publications of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

Published by The Rand Corporation

R-2637-DOJ

Quality of Prisoner Self-Reports

Arrest and Conviction Response Errors

Kent H. Marquis
with the assistance of Patricia A. Ebener

March 1981

Prepared for The National Institute of Justice,
U.S. Department of Justice

Rand
SANTA MONICA, CA. 90406

PREFACE

Under Grant No. 79-NI-AX-0055 from the National Institute of Justice, The Rand Corporation is conducting basic and policy-oriented research on criminal careers. Part of that research involves a large-scale questionnaire survey that was administered to prison and jail inmates in three states. This report evaluates the quality of arrest and conviction information furnished by prison inmates.

The report provides estimates of measurement biases and error variance based on record checks and retests. The estimates should be of particular interest to policy analysts and survey methodologists concerned with the effects of measurement error on subject matter statistical estimates.

A description of the large-scale survey is given in Mark Peterson and others, The Rand Offender Survey: Background and Method, The Rand Corporation, N-1635-DOJ (forthcoming, 1981). Additional methodological explorations of sampling and measurement issues are planned for future publication.

SUMMARY

This report evaluates the quality of answers to selected questions in a survey of adult male prisoners in three states. The strengths and weaknesses of prisoner self-reports of arrests and convictions are investigated by comparing them to official record information, and the hypothesis that prisoners systematically deny certain facts about their criminal careers is tested.

Our general findings include:

- o On the average, prisoners do not deny arrests and convictions. Amounts reported in the questionnaire are usually equal to or greater than the amounts coded from the records.
- o Response reliability is moderately high for self-reports of convictions, but uncertain for reports of arrests.
- o Discrepancies between survey and record values are not predicted well by ability, memory, and demographic variables, so we did not identify the "kinds" of prisoners prone to lying or to other response errors.

METHODS

A sample of approximately 1500 convicted male felons residing in prisons in California, Michigan, and Texas filled out questionnaire booklets; a retest subsample of 252 men completed the questionnaire a second time. Then the questionnaire reports of arrests within a defined time period and the reports of current conviction offenses were compared to information from official records. Bias scores based on average survey-record discrepancies were computed, and estimates of reliability made based on retest correlations and correlations between records and questionnaire responses.

Our reliability estimate expresses the amount of expected attenuation in a two-variable correlation estimate due to errors in the measure. Error scores were regressed on selected variables to identify possible correlates of response problems.

EVALUATION OF CURRENT CONVICTION REPORTS

Across items, the prisoners reported about the same number of current conviction categories that we found in the records (±6 percent). Several individual conviction items had statistically significant bias scores, but no item showed a consistent bias (same sign and statistically significant) across all three states. Several misclassification error patterns were found in the Michigan and Texas data; for example, robbery and theft convictions were often interchanged, either by the respondents, records, or coders. Item bias scores not explained by misclassification errors were usually positive.

The conviction item responses are moderately reliable. Using the record method, we find that correlations of conviction items with other, perfectly measured variables might be attenuated by 20 to 35 percent on the average, due to unreliability. The retest estimate of unreliability suggests only approximately 10 percent correlation attenuation, but the retest method probably overstates true reliability. Single conviction item reliabilities were generally satisfactory. No item showed reliability problems in more than one state.

The ability, memory, and demographic variable sets do not predict the respondent bias and total error scores uniformly across the three state samples.

EVALUATION OF ARREST REPORTS

Prisoners described their arrest history for a defined time period in two different questionnaire sections. Although each section is evaluated separately in the report, we summarize them jointly here. Both the surveys and the records appear to contain errors, and drawing conclusions by comparing the two sources is difficult.

When each arrest item is scored dichotomously and the bias scores summed over times, the net bias is close to zero (sometimes slightly negative, other times slightly positive). The dichotomous variable reliabilities (record method) are low, but the errors tend to "average out" over items and people. When the number of arrests are compared to the records, however, the average bias scores are usually positive (very large in Michigan) and the reliabilities get even lower. The item bias

scores show a mixture of large and small, positive and negative values that are seldom consistent across states. The pattern of results cannot be explained by the presence of a few extreme cases, or by the kind of item misclassification errors found in the current conviction evaluation. The pattern of arrest bias and reliability results is consistent with one or both of the following structural hypotheses:

- Respondents overreport arrests in proportion to the true number, and/or
- Records underreport in proportion to the true number.

In addition, either the questionnaire responses and/or the records contain substantial amounts of random error variation, or else substantial amounts of error were introduced while matching, coding, and entering the data on computer tapes.

We also estimated arrest response reliability by the retest method on smaller samples of respondents. These reliability estimates were much higher than those obtained by the record method. We suspect that respondents repeat their errors on the retest, thus inflating the reliability estimate in proportion to the amount of error covariation. An additional reliability estimation strategy to reduce the uncertainty about arrest item reliabilities is proposed in the report.

RECOMMENDATIONS FOR ANALYSIS

We recommend that analysts consider several steps when analyzing arrest data and similar questionnaire data in the future:

1. Repeat analyses in each of the three state samples separately. The item biases are usually localized (not the same sign and size in all states). Findings that replicate are unlikely to be effects of uncorrelated item biases.

2. To test the sensitivity of findings to the possible existence of a questionnaire bias that is positively correlated with true scores use transformed data and reanalyze.

3. Combine unreliable variables, thought to measure the same thing, into scales to reduce the effects of random response error variation.

RECOMMENDATIONS FOR FUTURE RESEARCH

We suggest that future surveys of this type include not only record checks but also additional data collection to describe the record error structure. Future surveys should also consider the costs, benefits, and feasibility of personal interviewing and the use of more redundancy in the questions.

ACKNOWLEDGMENTS

We wish to acknowledge the helpful comments on earlier drafts from Jan Chaiken, Marcia Chaiken and Carl Hensler of The Rand Corporation, from Joseph Weis of the University of Washington, Seattle, and from an anonymous review commissioned by the National Institute of Justice.

We also extend our sincere thanks to Susanne Polich and Paul Honig of Rand for valuable technical assistance throughout the project and to Mark Peterson for his continuing advice, encouragement, and substantive contributions during the analysis.

CONTENTS

TABLES

EXHIBITS

I. INTRODUCTION

Untruthful answers (response errors) are a major concern in the analysis of questionnaire data and in decisionmaking based on the results of such analysis. Unrecognized response errors can distort estimates of the true population mean, and can make it difficult to discover the true relationships among the variables.

In this report we evaluate the quality of selected data from The Rand Corporation's Second Inmate Survey. We emphasize measurement errors, assessing how closely the observed values from the questionnaires correspond to values obtained from official records. We do not address sampling, nonresponse, processing, or analysis sources of error.

The variables examined are prisoners' reports of arrests during a defined period of street months and prisoners' reports of the conviction offenses for which they were serving time when the questionnaire was completed.

In keeping with contemporary practice, the evaluation was conducted by persons not connected with the original design of the research and not involved in the subject matter analyses of the data.

We conclude that the arrest and conviction data are unbiased on the average but contain moderate to large amounts of random errors that will yield attenuated estimates of the true population associations.

The remainder of this introductory section reviews the relevant measurement literature, develops the methods used to evaluate the data, and describes the basic features of the Rand Offender Survey. The next three sections evaluate findings for three parts of the questionnaire: Section II, the reported conviction offenses, and Sections III and IV, the self-reports of arrests obtained in two different ways. Section V discusses the evaluation results and their implications for research analysis and future measurement design.

REVIEW OF THE METHODS LITERATURE[1]

Most researchers hypothesize that survey responses about arrest and conviction history will contain a negative bias. The conventional reasoning asserts that respondents prefer to deny, omit, or cover up undesirable facts about themselves. In this section we review the available literature concerning response errors in self-reported arrests and convictions. We conclude, contrary to conventional expectations, that arrests and convictions have been reported without a negative response bias in most well-designed studies. In fact, validity studies that are designed and analyzed properly usually find that the average response bias has a positive sign.

Approach

On the surface, the existing literature yields conflicting inferences about the quality of self-reported criminal and delinquent history. However, by reclassifying the studies and limiting attention to criterion validity research on arrest and conviction reporting, we found evidence that respondents usually reveal more arrests and convictions in questionnaires or interviews than can be found in official records.

In our review of existing methods literature, we focused on criterion validity studies, limiting ourselves to evaluations of self-reported arrests and convictions. We separated full design from incomplete design assessments and distinguished between direct and indirect estimates of average response bias.

Criterion Validity Studies. We wanted to assess the truthfulness of self-reports (their concurrent validity) rather than their predictive or construct validity. So we sought published research that compared survey answers to a criterion (or standard of truth) such as official records of arrests and convictions. Ideally, the survey questions and record entries referred both to the same concept and to the same time periods.

---

[1]This literature review is based on work done by The Rand Corporation on the reliability and validity of responses to sensitive survey questions (see Marquis et al., forthcoming). The authors gratefully acknowledge the help of M. Susan Marquis, Jan Meshkoff, J. Michael Polich, Donna Schwarzbach, and Cathleen Stasz in preparing the material.

Other research objectives might call for different kinds of validity assessments; for example, a researcher interested in future recidivism might study the ability of the answers to predict future rates of arrests and convictions. We do not address prognostic validity issues in this report. Another researcher may wonder whether arrest and conviction answers are good proxy indicators of underlying crime or delinquency rates; that researcher is interested in construct validity, or the extent to which the proxy measure possesses the same measurement properties and "behaves" as the true construct would in its cause-and-effect relationships with other variables. Because we do not address construct validity in our research, we omit construct validity studies from our literature review.

Full and Incomplete Design. We distinguish between full and partial record check designs because partial designs can result in misleading conclusions. The most common partial design evaluates answers only if records indicate an arrest, conviction, or other official action for a particular person. No evaluation of answers given by other people is made. The only thing this design can detect, in its pure form, is the respondent's apparent failure to report something already known about him. From such results, researchers sometimes infer that self-reports contain a negative response bias (e.g., denial, omission, lying, cover-ups). As discussed elsewhere (Marquis, 1978), however, these inferences can be incorrect for two reasons: (1) This kind of partial design can miss any positive response biases, so the direction of the average response bias may be incorrect. (2) It will cause random errors made by respondents, records, transcribers, coders, and data entry personnel--either on the main question or on items used to link answers and records--to appear as a negative response bias. Thus, the size of the bias will be overestimated.

Full designs select respondents independently of their record values and check answers of all respondents, regardless of whether the records indicate any official action. This allows estimates of errors in both directions (false positives and false negatives) and makes it possible to distinguish random errors from net biases.

Direct and Indirect Estimates. We distinguish between direct and indirect estimates of response bias. Several studies report the proportion of respondents whose answer was above the recorded value and the proportion who gave an answer below the criterion value. We call these indirect estimates. By noting the difference in proportions we can tentatively infer something about the sign of the response bias but nothing about its size. If the response variable is dichotomous, we also need to know the proportions of matched positives and matched negatives to get a direct estimate. For continuous variables we also need to know the average size of the response record differences for a direct bias estimate.

Full Design Studies of General Adult Populations

Knupfer (in press) and Bridges (1979) find that adults in general populations report more arrests in surveys than can be found for them in records. Jessor et al. (1975), however, indicate that the conviction response bias may be negative, on the average.

In 1964, Knupfer asked a sample of 970 noninstitutionalized men and women in San Francisco about drinking-related arrests. She searched police records for "any mention of drunk driving or other drunkenness related offenses." Truncating each source at three arrests and deleting 94 respondents who did not answer the arrest questions resulted in a mean number of .09 reported arrests per person. The recorded number was .08, implying a positive average response bias of 12 percent [(.09 - .08)/.08]. We computed a correlation of .52 between the self-reports and recorded values. The author did not mention any special attempts to obtain complete juvenile record data.

Bridges (1979) located and interviewed 567 adult males in a 1945 Philadelphia birth cohort who had resided in Philadelphia during their school years (age at interview was 25-26 years). Respondents were asked how many times they had been arrested ("picked up or taken to the police station for any offense"). Responses were compared to information extracted from Philadelphia police records. An analysis based on 558 cases shows 1.37 self-reported arrests per person and 1.17 in the records.[2]

[2]The per-person record mean is inferred from Bridges' data.

We infer a +18 percent average bias from these numbers. The correlation between the two data sources was .60.

Jessor et al. (1975) furnish a full design interview study using a sample of males and females in the 1945 birth cohort of a Colorado community. Respondents were asked about alcohol-related problems followed by the ambiguous question: "What (else) have you gotten into trouble with the law about?" Adult and juvenile records covering a ten year period were obtained from courts, prisons, the Bureau of Indian Affairs, a mental hospital, and probation offices. The authors focus on recorded "convictions," excluding those arising from civil suits or traffic and game law violations. They had to infer which answers to the ambiguous question referred to convictions.[3]

The comparison of inferred survey answers to records showed 16 false negatives, 3 false positives, 26 true positives, and 173 true negatives. Our estimate of the average bias is -32 percent of the record frequency: [(26 + 3) - (26 + 16)]/(26 + 16) = -.32. No estimate of the bias in reporting the number of convictions could be made. The lesson from this research is probably to ask explicit questions about topics of interest. It is not surprising that some respondents would fail to report details of convictions to an ambiguously worded question with no clear time referent.

Full Design Studies of Special Groups

Full design studies of student, alcoholic, and former drug addict groups also suggest the absence of a denial (net negative) response bias.

Nine hundred fourteen 7th through 9th grade boys filled out questionnaires in a delinquency study by Hardt and Peterson-Hardt (1977). One item was "I was given a ticket or arrested by the police." Answer categories were (1) In the last seven days, (2) in the last two months, (3) over a year ago, and (4) never. These answers were checked against

[3]The authors do not say how the inference was made. In their text, they refer to reporting "their police record" when discussing true positive and false negative interview reports. They refer to "reported convictions" when discussing survey overreports (false positives). This raises the possibility that different definitions were used that might bias the estimate of average response bias in a negative direction.

files of the county central registry of juvenile offenders for 862 boys. Using dichotomous scoring, 22 percent of the boys reported being arrested or ticketed, but records indicated only 20 percent were so categorized. This is a positive 10 percent average response bias. The self-reports correlated .67 (phi coefficient) with the registry information. The authors suggest and offer empirical support for the hypothesis that records were incomplete in two respects: arrest data for nonjuvenile offenses were sometimes missing, and they did not contain a complete count of arrests occurring in other jurisdictions.

Ball (1967) evaluated the questionnaire (program application) responses of 58 ex-addicts originally convicted on drug charges and incarcerated at the Lexington treatment facility. FBI records were used in the analysis of 54 cases (4 persons who reported "numerous arrests for breach of peace and drunkenness" were deleted, presumably because the FBI records did not substantiate the self-reports). For the 54 cases the self-reported mean per person was 3.8 and the comparable estimate from records was 3.5. We infer a positive net response bias of 9 percent from these results.[4]

Ex-addicts are found to give reliable arrest and conviction reports. Bale (1979) obtained a 78 percent retest agreement rate on the item of whether the person had been convicted. Stephens (1972) obtained an agreement rate of 94 percent between self-reports and counselor reports of whether the respondent had been arrested. Stephens also found an 88 percent agreement between the arrest self-reports and information from relatives. Unfortunately, agreement scores cannot be used to assess the net response bias.

Walsh (1967) asked 270 male undergraduates about the number of contacts with the Iowa City Police and about the number of traffic tickets received from the campus police. He does not give means or state the direction of disagreements but reports only 1-2 percent disagreement with city police records and a 13 percent disagreement with

[4] Amsel et al. (1976) provide another criterion validity study of self-reported arrests by former drug addicts (but using local police records). Detailed data are not reported so we were unable to obtain a direct estimate of the response bias. Results that are presented are sufficient for an indirect estimate of the sign of the response bias (it is negative), but the authors' text suggests the bias is positive.

campus police records. These could represent large or small biases (depending on the size of the criterion mean) and could contain positive, negative, or no average biases (depending on the relative frequencies of overreport and underreport errors). Sobell (1976) asked 33 male college undergraduates about the number of their California speeding tickets in a two-year period. Compared to state records, more students overreported the number than underreported. This furnishes an indirect estimate of the sign of the bias (positive) but no information about its size.

Linda and Mark Sobell have contributed many response validity studies of reported alcohol-related arrests, convictions, and other official actions for small samples of male alcoholics and former alcoholics.[5] Self-reports are obtained either by personal interviews or questionnaires. Records include those from the California Department of Motor Vehicles, the California Bureau of Criminal Identification and Investigation, and the FBI. Unfortunately, few of the studies present enough detail to estimate either the average response bias or the correlation between self-reports and official records. We discuss the indirect estimates here to infer something about the sign of the average response bias. The sign is positive if relatively more respondents report amounts greater than (vs. less than) the criterion values.

We constructed an alcohol arrest index consisting of at least two of the following: public drunk arrests, drunk in auto arrests, and open container arrests. The index has a positive sign in four of the five studies from which it could be computed (in the fifth study, no respondent report disagreed with the recorded values).[6] Similar positive signs are found for the reported number of California speeding tickets in a 2 year period (4 out of 4 studies), the number of auto accidents in 3 years (4 out of 4 studies), number of reckless driving convictions (2 of 2 studies), and frequency of driver's license suspensions or revocations (2 of 2 studies). Five studies provide direct estimates of

[5] Sobell and Sobell (1978); Sobell (1976); Sobell and Sobell (1975); Sobell, Sobell, and Samuels (1974).

[6] Sobell, Sobell, and Samuels (1974) indicate self-reported arrests for DWI or public intoxication correlate .65 with official record numbers (n = 70).

admitting current probation or parole status. Probation has a positive response bias in all the studies; apparently, none of the respondents had a record indicating he was on parole, and in only 1 of the 5 studies did anyone claim he was on parole.

We also constructed an index that combined other arrest charges, including drugs, burglary, petty theft, robbery, assault with a deadly weapon, disturbing the peace, assault and battery, and "other crimes." The index score was negative in 4 out of 4 studies, suggesting that alcoholics may be prone to denying non-alcohol-related arrests.

Sobell et al. (1979) report a high retest reliability estimate for self-reports of lifetime alcohol arrests ($r = .97$, $n = 12$). No record data were available. Guze et al. (1963) obtained a .43 correlation between paroled convict self-reports of drinking-related arrests and similar information from relatives (informants).

The full-design evaluations, then, tend to find average response biases that are usually positive for self-reports of arrests and convictions. The finding holds for general as well as special population groups and for both direct and indirect estimates of the bias.

We review the partial design studies next. They tend to yield estimates of a negative response bias. However, partial design studies fail to detect positive response errors and so the findings are probably caused by the faulty design rather than the true reporting tendencies of respondents.

Partial Design Studies

Petersilia (1977) reports the only validity study we found using imprisoned felons as respondents. She interviewed 49 males whose records indicated were currently serving time for armed robbery and who had also served at least one prior prison term. The self-reports of arrests and convictions were compared to state and federal records. Respondents reported a lifetime average of 5.1 arrests on nine felony charges while records indicated an average of 7.7 arrests (a negative 21 percent average response bias).

In a study of 600 ex-addicts enrolled in a jobs program, Wyner (1976) compared self-reported arrests of 79 participants to New York

City Police records. He found a small negative reporting bias that was not statistically significant.

Locander et al. (1976) drew a sample of people involved in bankruptcy proceedings from records and another sample of people arrested for drunk driving. They found almost 50 percent underreporting of drunk driving and 25 percent denial of bankruptcy. These are the largest negative biases reported in this literature. The published study gives few methodological details so it is difficult to tell what, besides the partial design, might account for the results.

The juvenile delinquency literature contains quite a few measurement studies.[7] Most focus on the construct validity of self-report behavior indexes as indicators for underlying crime rates of individuals, groups, and other collectivities. Their focus on behavior and construct validity excludes them from consideration here, but there are a couple of exceptions. Erickson and Empey (1969) executed a full design but discuss only the responses of juveniles known (from records) to have appeared in court or been officially charged with an offense. Among other things, the respondent was asked if he had ever been caught, arrested, or taken to court for each of 22 offenses. None of the recorded court appearances or charges was denied in the interview. The authors do not comment on how many arrests, etc. were reported but not verified by the records.

Voss (1963) also started with a full design but chose to report only part of the results. Six hundred twenty students were asked if they had ever been arrested (taken to the police station for something you have done). Fifty-two students had been apprehended by the police according to records. Of these, only one student failed to report being arrested. We are not told how many reported arrests were not found in the records.

---

[7] For example, Nye and Short (1957), Belson (1968), Kulik et al. (1968), Hirschi (1969), Gould (1969), Hackler and Lautt (1969), Gibson et al. (1970), Williams and Gold (1972), Farrington (1973), Elliott and Voss (1974), and Blackmore (1974). Most of the research uses both self-reported behavior and official records. Several discuss whether the juvenile admitted committing the offense with which he was charged. This is a partial design approach to analysis.

Summary

While many partial design studies in the literature seem to show a negative response bias, this may be due to using a research and analysis strategy that can detect errors in only one direction. The consequent negative bias, then, may be a research artifact rather than a true behavioral characteristic of respondents. Full design studies, on the other hand, can detect errors in both directions. Most of these studies show a positive response bias; that is, they obtain more reports of arrests and convictions in questionnaires than can be found in official records. There are exceptions, however. Jessor et al. find a negative bias in reporting convictions and the Sobells consistently find that alcoholics or ex-alcoholics underreport their felony arrests. To add to the uncertainty, none of the existing full design studies evaluated responses of imprisoned felons.

We approach our own questionnaire evaluation research, realizing the importance of the full design analysis strategy and knowing that although most well-designed studies do not find a negative response bias, there are exceptions and there are special circumstances that may be unique to the prisoner population that we will study.

RESPONSE ERROR MODELS, ESTIMATES, AND EFFECTS

Our goal is to describe response errors in Rand Offender Survey data that might adversely affect the analysts' statistical estimates. We know that analysts will estimate population means and correlations. So our description will focus on the kinds of response errors that can distort these statistics, specifically, average response bias and response reliability. In this section, we mention some key concepts in modeling and estimating response errors and in linking their effects to statistics. We state our response model, explain how its parameters are estimated, and show relationship between these parameters and the basic statistics that analysts will estimate.

There is a lack of agreement among practitioners on what constitutes good survey data. One approach is to assume the data are good if they are collected by a recognized organization that employs highly trained questionnaire administrators, extensively pretested questionnaires, and state-of-the-art data reduction techniques. This approach

assumes that perfect measurement can be achieved and that the professionals know how to do it. Unfortunately, while the approach cannot be entirely wrong, there is very little research evidence to indicate that perfection is possible to achieve. A second approach is to purge the data set of "bad" observations. This is done after data are collected by recoding each questionnaire and discarding responses or forms that fail to conform to expectations of reasonableness and consistency. Again, the approach probably has some benefits, but it rests on intuitive judgments and it is difficult to know about the quality of the observations that remain.

After employing a professional approach to designing, pretesting, and administering the questionnaires, and after discarding the most obvious instances of confusion and deceit, we still need a way of knowing the quality of the remaining observations. In this report we adopt a quantitative approach to describing measurement quality. The technique has the advantages of resting on explicit models and assumptions and of having a direct connection to statistical estimates. It does, however, rest on assumptions and other simplifications that may be imperfect depictions of reality.

Response and Record Models

Our response model contains parameters that can be linked to estimates of means and correlations. It is based on classical test theory (e.g., Guilford, 1954), the Census Bureau Models (e.g., Hansen et al., 1961), and on more recent structural equation approaches (e.g., Henry, 1973; Joreskog and Goldberger, 1975; Marquis and Marquis, 1977).

We assume that a person's response to an item can be described as follows:

$$S_{ij} = T_{ij} + b_j T_{ij} + B_{ij} + e_{ij} \qquad (1)$$

where

S_{ij} = response on the survey of person i to item j,
T_{ij} = true value of item j for person i,
b_j = parameter indicating the correlation of response bias with the true score for item j,

B_{ij} = response bias uncorrelated with true score for item j,
e_{ij} = random response error for person i on item j. The expected value $E(e_{ij})$ is zero; e is assumed to be uncorrelated with other variables.

The model of the record value contains the same kinds of terms:

$$R_{ij} = T'_{ij} + b'_j T'_{ij} + B'_{ij} + e'_{ij}, \qquad (2)$$

e_{ij} and e'_{ij} are assumed to be uncorrelated.

The Bias Scores

Response biases that do not average out to zero can have important effects on estimates of first-moment statistics such as means, proportions, and rates.

We evaluate bias in the questionnaire responses with the bias scores. They summarize the amount and direction of the average disagreement between the survey observations and the record observations. Each bias score is a correction factor, that is, the bias score is the amount subtracted from the survey mean to make it equal to the record mean:

Survey Mean - Bias Score = Record Mean.

Positive bias scores indicate that the survey value is larger than the record value. If the record value is "true," a positive bias score means that there is a questionnaire overreporting bias (and a negative score means a questionnaire underreporting bias). If the records are imperfect, a positive bias score may indicate underreporting in the record.

We use three bias scores, two to describe the bias for questionnaire items and one to portray the bias tendencies of individual respondents.

The basis for the three bias scores is the difference between the questionnaire and record values for an item. Define the basic bias score, D_{ij}, for person i and item j as follows:

$$D_{ij} = (S_{ij} - R_{ij}) \qquad (3a)$$

where i indexes the respondent and j indexes the item.

Using eqs. (1) and (2), the difference may be rewritten as

$$D_{ij} = (T_{ij}-T'_{ij}) + [(b-b')(T_{ij}-T'_{ij}) + (B_{ij}-B'_{ij})] + (e_{ij}-e'_{ij}). \qquad (3b)$$

Before developing the error indexes used in this research, the interpretation possibilities for D_{ij} should be discussed; what does a nonzero value of D_{ij} indicate about the quality of the survey and record observations? The answers come from examining three parenthetical terms in eq. 3b:

1. $(T_{ij} - T'_{ij})$: This term represents the possibility that the questionnaire and record definitions of the true value of item j are different. Several definition differences may exist.

Occasionally the questionnaire and coded records disagree about the street months time period (the total number of months involved and which months are involved). A nonzero value of D_{ij} may indicate this kind of difference on the two arrest variables.

Questionnaire and record definitions of arrests may differ and the discrepancies are partly a function of which arrest question set is being evaluated. One set of questionnaire items, Arrests for Crimes Done, elicits reports of arrests for specific categories of crimes that the respondent committed (e.g., In all, how many thefts did you do? How many of these thefts were you arrested for?). There are three kinds of definition differences here: (a) the committed crime category (e.g., theft) may differ from the arrest charge category (e.g., arrested for drug possession while committing theft); (b) the respondent will not report arrests for offenses he did not commit (e.g., a "false" arrest for theft); (c) the respondent cannot report arrests on felony charges if the offense leading to the arrest is not one listed on the questionnaire (e.g., arrested for burglary after breaking a window).

The second set of arrest questions, Arrest Incidents, does not have these problems because the respondent is asked to report arrests in the category of the official arrest charge. Here, however, there are other possible ambiguities concerning the appropriate definition of the number of arrests. The term "arrest" can mean three things: (a) an occasion when a person is taken into official custody, (b) a charge that the

person has committed a crime (or crimes) of a specific type, or (c) a count, which is one particular offense within a set of similar offenses with which the person is accused. An arrest occasion may involve multiple charges and/or multiple counts per charge. Also, a person may be arrested on several occasions on a single count. Charges and counts can be added without incurring a new arrest occasion.

An experienced RAP sheet coder can distinguish and code separate counts, charges, and occasions fairly accurately. For this project, the RAP sheet coder was asked to abstract (code) counts and charges. For the evaluation analysis, only the recorded count data were used. The questionnaire did not instruct the respondent about which definition of arrests to use (counts, charges, or occasions), but the evaluation analysis assumes that the respondent was answering in terms of arrest "counts." If the assumption is violated, the obtained bias scores will be negatively biased. As it turned out, the possibility of negatively biased bias scores would not affect the conclusions drawn from our analysis.

2. $(b - b')(T_{ij} - T'_{ij}) + (B_{ij} - B'_{ij})$: These difference terms represent the systematic questionnaire and record biases. We will simplify for this discussion by assuming $b = b' = 0$ and the terms reduce to $(B_{ij} - B'_{ij})$. If the record is unbiased ($B'_{ij} = 0$), the nonzero difference is an estimate of the questionnaire bias for item j. A positive value suggests that respondents are exaggerating while a negative value indicates an underreporting or denial response bias. For this report, however, we are unwilling to assume that the coded official record data are unbiased. There is a possibility that arrest record data are incomplete and the additional possibility of errors in coding the data and/or linking up the questionnaire and record information for each respondent.

By assuming possibly imperfect records ($B'_{ij} \neq 0$), we made the interpretation of nonzero values of D_{ij} difficult; a positive value could also indicate record omissions and a negative value could suggest "false positive" record entries (e.g., double counting a single arrest).

3. $(e_{ij} - e'_{ij})$: The final set of differences are in the random error terms of the model. Each has an expected value of zero over persons and items, so the expected value of their difference is also zero. On the average, then, nonzero values of D_{ij} are not due to differences in the random error values. For any single comparison of a survey

and record value, however, the obtained difference could be a "random" phenomenon that is unlikely to reoccur in exactly the same way over repeated trials (e.g., retest), across respondents, or across items.

In summary, a nonzero difference between a questionnaire and record value could be due to definitional differences, systematic biases, or random errors. The random error effects are expected to diminish to zero over repeated, independent observations. In general, we will assume no definitional differences so that we conclude that a nonzero value of D_{ij} suggests the operation of a systematic data bias of unknown source.

Single Item Bias Score. The single item bias score is used to portray systematic data biases in single questionnaire-record item pairs, for example, survey and record values for burglary arrests or murder convictions. The single item bias score is the basic item bias score for item j averaged over the n respondents for whom questionnaire and record data are both available:[8]

$$\text{Single Item Bias Score} = \Sigma_{i=1}^{n} D_{ij}/n \qquad (4)$$

where

n = number of respondents for whom questionnaire and record data are available for item j. Implicitly this imputes missing data at the item mean for the observed group.

D_{ij} = basic item bias score (defined in eq. 3a) for the i th person and the j th item.

Respondent Bias Score. We investigate the possibility of predicting data biases using characteristics of respondents such as ability and demographic variables. One dependent variable in these analyses is the Respondent Bias Score, which is the basic item bias score averaged over items in a section of the questionnaire for a single respondent:

---

[8] In all estimation formulas, i refers only to respondents for whom record data were available. The treatment of missing questionnaire item data for a respondent varies with the estimate being made.

$$\text{Respondent Bias Score} = \Sigma_{j=1}^{k} D_{ij}/k \qquad (5)$$

where

k = number of items in the questionnaire section for which survey and record values are available for person i. This is equivalent to imputing a bias score for missing items equal to the average bias for this person on the questions he answered.

D_{ij} = basic item bias score (eq. 3a).

Summary Bias Score. The third bias score is used to summarize the Single Item Bias Scores over items in one section of the questionnaire The summary measure is useful not only because it reduces a lot of separate estimates to a single number but also because it captures the overall trend in the Single Item Bias Scores. The summary Bias Score is defined as

$$\text{Summary Bias Score} = \Sigma_{i=1}^{n} \Sigma_{j=1}^{k} D_{ij}/n \qquad (6)$$

where respondents are indexed by i (i = 1, ..., n), items for which both questionnaire and record data exist are indexed by j (j = 1, ..., k); and D_{ij} = basic item bias score (eq. 3a). We imputed missing item responses at the average bias for the item.

Bias Score Definition Summary. The definition of the three bias scores can be viewed from the perspective in Table I.1. The columns are items in a section of the questionnaire, and the rows are respondents. The entries in the cells are the basic item bias scores.

Estimates of the Reliability Bounds

Two methods of estimating item reliability are used: the record method and the retest method. We present the reliability estimates in terms of the implied attenuation of a subject matter correlation due to the unreliable variance in the data. Since we are concerned with single item measures, we delete the j subscript; for convenience we delete the i subscript also.

Table I.1

SUMMARY DEFINITION OF BIAS SCORES

	Items in One Questionnaire Section						
Respondents	1	...	j	...	k	All Items	
1	D_{11}					$\Sigma_{j=1}^{k} D_{1j}$	*
...							
i			D_{ij}			$\Sigma_{j=1}^{k} D_{ij}$	*
...							
n					D_{nk}	$\Sigma_{j=1}^{k} D_{nj}$	*
All Respondents	$\Sigma_{i=1}^{n} D_{i1}$		$\Sigma_{i=1}^{n} D_{ij}$		$\Sigma_{i=1}^{n} D_{ik}$	$\Sigma_{i=1}^{n} \Sigma_{j=1}^{k} D_{ij}$	
	When divided by n, these are the Single Item Bias Scores					When divided by n, this is the Summary Bias Score	

NOTE: See text for a description of missing data imputation procedures.

*When divided by k, these are the Respondent Bias Scores.

Effect of Unreliability on a Correlation Coefficient.[9] Suppose we are interested in the correlation between the variable T and another variable X. Assuming that X is perfectly measured, we will estimate the correlation between T and X, denoted r(T, X), by the correlation between X and the measure of T, which is S. This latter correlation is denoted r(S, X) where:

$$r(S, X) = \text{cov}(S, X)/(\sqrt{\text{var } S}\ \sqrt{\text{var } X}).$$

Using the measurement model shown in eq. 1, and assuming that the measurement bias B is uncorrelated with X,

$$\begin{aligned} r(S, X) &= (1 + b)\ \text{cov}(T, X)/(\sqrt{\text{var } S}\ \sqrt{\text{var } X}) \\ &= r(T, X)(1 + b)(\sqrt{\text{var } T}/\sqrt{\text{var } S}). \end{aligned} \tag{7}$$

We will call the expression $(1 + b)(\sqrt{\text{var } T}/\sqrt{\text{var } S})$ the Multiplicative Correlation Attenuation Factor (MCAF). The MCAF indicates the degree to which measurement error attenuates the estimate of the true correlation, r(T, X). For example, if the MCAF is .6, the measured correlation is 60 percent of the true correlation; the effect of measurement error is to reduce the correlation by 40 percent.

Record Method Estimates of MCAF. We will use the correlation between the survey and record values, and the square root of this correlation to estimate bounds on the MCAF. The bounds correspond to alternative assumptions about record error.

Using the measurement model in eqs. 1 and 2, and assuming var B = 0, the correlation between survey and record values can be written:

$$\begin{aligned} r(S, R) &= \text{cov}(S, R)/(\sqrt{\text{var } S}\ \sqrt{\text{var } R}) \\ &= (1 + b)(1 + b')\text{var } T/(\sqrt{\text{var } S}\ \sqrt{\text{var } R}). \end{aligned} \tag{8}$$

[9] Other treatments of links between measurement errors and statistics may be found in Cochran (1968, 1970), Blalock, Wells, and Carter (1970), Bohrnstedt and Carter (1971), Marquis (1977), Marquis and Marquis (1977), and Marquis et al. (forthcoming).

Case 1: Assume that the record is a perfect measure of T, that is, var R = var T and b' = 0. Then all observed discrepancies are survey errors and the correlation in eq. 8 reduces to

$$r(S, R) = (1 + b)(\sqrt{\text{var } T}/\sqrt{\text{var } S}) = \text{Lower Bound MCAF.}[10] \tag{9}$$

Case 2: Assume that the record and survey contain equal errors, that is, b = b' and var S = var R. Then the correlation in eq. 8 becomes

$$r(S, R) = (1 + b)^2(\text{var } T/\text{var } S)$$

and

$$\sqrt{r(S, R)} = (1 + b)(\sqrt{\text{var } T}/\sqrt{\text{var } S}) = \text{Upper Bound MCAF.} \tag{10}$$

The correlation between S and R and the square root of the correlation thus are estimates of MCAF under alternative assumptions about record errors. The true MCAF probably lies between the two estimates.

Retest Method Estimate of MCAF. The retest provides a second survey measure. Let S_{i1} and S_{i2} be the two responses to an item for individual i. Assume both measures follow the response model in eq. 1. We will make the following assumptions about the random errors:

$$\text{cov}(e_1, e_2) = 0$$

and

$$\text{var } e_1 = \text{var } e_2$$

so that

$$\text{var } S_1 = \text{var } S_2 = \text{var } S.$$

Under these assumptions, the square root of the correlation between S_1 and S_2 is an estimate of the MCAF:

[10] $r^2(S, R)$ is traditionally used to estimate survey reliability when perfect criterion measures are available. To calculate the correlation attenuation factor, however, one must still take the square root of this reliability estimate.

$$r(S_1, S_2) = \text{cov}(S_1, S_2)(\sqrt{\text{var } S_1}\ \sqrt{\text{var } S_2}) = (1 + b)^2(\text{var T/var S}),$$

and

$$\sqrt{r(S_1, S_2)} = (1 + b)(\sqrt{\text{var T}}/\sqrt{\text{var S}}) = \text{MCAF}. \qquad (11)$$

In the results to be presented, the MCAF estimates based on retest will typically exceed those based on the record check. The explanation for this may be that respondents make consistent errors across trials and so the correlation of the errors e_1 and e_2 is greater than zero. If the assumption that $\text{cov}(e_1, e_2) = 0$ is violated, then the retest will overestimate the MCAF by $\text{cov}(e_1, e_2)/(\sqrt{\text{var } S_1}\ \sqrt{\text{var } S_2})$.

Total Error Score

We need a variable that captures the total amount of error we think each individual has made in each of the questionnaire sections examined. The Total Error Score is used for this purpose but, unfortunately, the score has no direct links to analytic estimates as do the bias and MCAF indexes.

The total error score is similar to the Respondent Bias Score (eq. 5) except that the absolute value of the basic item bias score is involved:

$$\text{Total Error Score} = \Sigma_{j=1}^{k} |D_{ij}|/k \qquad (12)$$

for person i, where

$|D_{ij}|$ = absolute value of the basic item bias score (eq. 3a)

k = number of items in the questionnaire section for which survey and record values are available for person i. In effect, missing questionnaire responses are imputed at the person's error mean.

Measures of Possible Causes and Correlates of the Errors

Some types of prisoners may be more likely to make response errors than others. The empirical literature does not suggest who the error-prone types might be, so we undertake an exploratory analysis of variables that might identify liars or other error-prone respondents.

We use the respondent bias scores and the total error scores as dependent variables. Exploratory predictors were chosen to test our intuitive hypotheses about the respondent and other characteristics that might be systematically related to response quality. We group the predictors into ability variables, memory variables, match problem indicators, and demographic characteristics. They are discussed in more detail below and summarized in Table I.2.

The respondent's ability to respond accurately is captured by three variables: education level, personal questionnaire administration, and the number of skipped question items. The education predictor is based on the prisoner's own report of the number of grades of formal education he completed. The variable is ordinally scaled as indicated in Appendix A, Exhibit A.6.

Table I.2

LIST OF VARIABLES USED AS PREDICTORS
OF RESPONDENT BIAS AND TOTAL ERROR SCORES

Ability Variables
- Education Level (X_1)
- Personal Questionnaire Administration (X_2)
- Skipped Items
  - In Current Conviction Section (X_3)
  - In Arrest Incident Section (X_4)
  - In Arrests for Crimes Done Section (X_5)

Memory Variables
- Elapsed Time Since End of Street Months (X_6)
- Number of Felony Convictions (Indicator of Interference) (X_7)

Record Match Problem Potential
- Recorded Street Months End Later Than Reported (X_8)
- Reported Street Months End Later Than Recorded (X_9)

Demographic Variables
- Respondent Age (X_{10})
- Respondent Race (X_{11})

If the respondent could not read either the English or Spanish version of the questionnaire, the session administrator gave him personal assistance. The personal assistance dummy variable is scored 1 to reflect this; a 0 value indicates the respondent answered under regular group administration conditions.

Three additional predictors indicate the number of items the respondent skipped (should have answered but did not). The first two variables are counts of the number of missing items in the two questionnaire sections about arrests (Arrests for Crimes Done and Arrest Incidents). The third is an indicator of whether the entire conviction offense section was skipped (we cannot detect skipping only part of the section). Some ambiguity exists about interpreting a relationship between an arrest bias score and the number of skipped items in that section. This arises because of how the bias score is constructed when the respondent has skipped items. The numerator of the average item bias score is the sum of the survey-record differences for answered items; the denominator is the number of answered items. The effect of the scoring is to impute the person's bias mean for the missing items. As more items are skipped, more imputation is needed. An association of the number of skipped items with the Respondent Bias Score or the Total Error Score, therefore, can reflect errors in the imputation model and assumptions. However, the association between the number of missing items in question set A with the average bias score for question set B has a straightforward interpretation.[11]

Two variables are used to represent potential memory or recall problems; one reflects memory decay and the other memory interference. The most common hypothesis about memory effects in surveys concerns memory decay, a forgetting tendency that increases with the passing of time. We have constructed a variable to indicate how long ago the last

[11] It may be that missing item rates for question sets A and B are highly correlated, which would raise the interpretation problem again. However, multiple regression methods are used to estimate the association between missing items in Section A and the average bias score in Section B. To some extent, the multiple regression method can "correct for" the correlation between A and B missing item frequencies.

arrest event to-be-recalled took place.[12] The event is the arrest for the conviction on which the prisoner-respondent is now serving time (this event marks the end of the Street Month time period, described later). The respondent's report of the arrest month and year is accepted regardless of later modification by editors or staff record abstractors. The elapsed time variable value is defined as:

Elapsed Time = (year arrested × 12) + month arrested.[13]

Since larger values indicate shorter elapsed time, the reader should mentally reverse the sign of the coefficient to interpret the memory decay effect.

Total number of lifetime felony convictions (reported by the respondent) is used as an indicator of interference memory problems. Interference memory theories view recall as a process of selecting relevant events to-be-reported from a "set" of similar events stored in memory. Selection errors increase as the size of the set of events increases and as the similarity of items within the set increases. We assume that total number of lifetime felony convictions reflects, albeit imperfectly, the number of similar events in the memory set that must be scanned and evaluated during questionnaire recall and responding about arrests and current convictions.

We did not use the number of reported or recorded events (e.g., arrests) as an indicator of set size because the regression coefficients would be biased. The coefficient bias occurs because the measurement errors in the predictor variable are also in the dependent variable. These "correlated errors" would bias the regression coefficient estimate away from zero, as explained in Appendix F.

Two variables are used to indicate possible distortions in the bias score because the respondent and the record abstractor were not using

[12] Although this is the simplest indicator of elapsed time between arrest events and filling out the questionnaire, it is imperfect because the length of the total street month period (total number of months included) varies among respondents.

[13] Coded January = 1, ..., December = 12.

exactly the same street month time period. We compared the respondent's version of the end of the street month period with that of the record abstractor. (The abstractor could alter the ending month on the basis of information in the records such as that the first arrest for the current conviction offense occurred three months earlier than the respondent said it did.) We hypothesized that if the recorded end of the street months was later than the respondent's version, more arrests and convictions would be recorded and a negative bias score would result. On the other hand, a recorded end of the street months that occurred somewhat earlier than reported by the respondent would not have predictable effects. This is because the respondent probably was not arrested during the period in dispute. Because of the directional hypotheses, we created two dummy variables:

X_{10} = 1 if record street months ended after the questionnaire version
= 0 if ended in the same month or earlier

X_{11} = 1 if record street months ended before the questionnaire version
= 0 otherwise.

The final two variables are the demographic characteristics of age and race. We have no hypotheses about their links to the error scores; they are included because analysts may wish to know if age or race-defined subgroups have different amounts of error in their data. Age is entered in years. Race is represented by a dummy variable whose value is 1 if the respondent said he is white and 0 if he said he is chicano, black, or something else.

Persons with missing data on any predictor or dependent variable were excluded from the multiple regression analyses.

The dependent variables are the average Respondent Bias Score and the Total Error Score (eqs. 5 and 12).

DESIGN OF RAND OFFENDER SURVEY

Next we describe the sample, measurement, and field procedures used in the Rand Offender Survey.

Sample Design

The sample design was multistage, and it combined judgment, probability, and quota features. On the basis of judgment, state prisons[14] in three states were selected. The population of prison inmates sentenced from specified counties was stratified on age, race, past prison record, and current conviction offense and selected with probability inversely proportional to sentence length. Substitutes were drawn from the relevant stratum if the selected inmate was unwilling or unable to participate. Approximately 1500 completed questionnaires were obtained from the initial prisoner sample, which included about 2500 primary selections and about 600 replacement candidates.

Prisoner Analysis Groups

For various reasons, the number of cases entering the measurement evaluation analyses is less than the full sample who completed questionnaires. Some questionnaires were excluded during the data cleaning phase,[15] official records were not obtained for some,[16] and others were excluded because respondents failed to answer some or all of the questions in the relevant sections of the questionnaire (item nonresponse). The analysis group sizes are summarized in Table I.3. Because item nonresponse is variable, ranges are shown. Group sizes for specific analyses in this report may be obtained from the relevant text or appendix tables.

The interested reader can learn something about the average characteristics of persons in some of our analyses by consulting Appendix Table C.1. The average age of respondents was around 27; their

[14] Jails were also selected at this stage. Responses from the jail sample are not evaluated in this report because record information for them was not sought.

[15] Questionnaires that were extremely difficult to code or mostly blank were reviewed by the analysts. If the analysts agreed that a questionnaire was unusable, it was excluded.

[16] In Michigan and California, missing records arose primarily because the needed information was not available during the abstracting period. Time constraints limited the number of cases that could be abstracted in Texas so records (except convictions) were systematically sampled.

Table I.3

EVALUATION ANALYSIS GROUP SIZES BY STATE
(Numbers of Questionnaires)

Item	California	Michigan	Texas
Completed Questionnaires	366	461	676
Failed Coding Checks	6	21	7
Official Record Not Obtained	16	79	0[a]-84
Item Nonresponse Deletions	0-44	1-54	37-86
Analysis Group Sizes	300-344	307-360	499-632

[a]Conviction information from the Texas Department of Corrections tape was available for all respondents.

educational attainment average was some high school. The average number of adult felony convictions was over 3, and between 30 and 42 percent of the sample were white.

Survey Field Procedures

The questionnaire was administered to groups of 20 to 30 men at a time in classrooms, visiting rooms, and other facilities available inside the institutions. Eight survey administrators were hired and trained to conduct the survey sessions. Sessions were usually run by two or three administrators, who had received approximately 20 hours of training on the procedures and the questionnaire. They had also participated in two pretests before they began actual fieldwork.[17] All but one of the administrators had previous experience working with felons in correctional institutions.

[17]Three survey administrators were hired after fieldwork began, and they did not have the pretest experience.

The inmates selected for the survey were notified by mail (Appendix E, Exhibit E.1). The institutions divided the men into groups, assigning them to particular sessions depending on where they worked or were housed in the institution and notifying them by ducat[18] to attend.

At a typical session, men usually arrived in small groups and had an opportunity to talk with the survey administrators before the session began. One administrator passed out the questionnaires while the other explained the purpose of the survey and the response task. Then an administrator read the Agreement To Participate form aloud and asked the men who wanted to participate to sign their names on the forms. The refusals were asked to leave at this time. The administrators reviewed the instructions on how to fill out the questionnaire and offered help with any questions or problems. They also offered Spanish versions of the questionnaire to men who could not read English and asked the nonreaders to identify themselves so that other arrangements could be made.[19]

Sessions usually ran smoothly, resembling a room of students taking achievement tests. Respondents worked on their own to fill out the questionnaire. They frequently asked administrators to help figure out the exact street month time period for the criminal behavior section and to explain the crime categories when they did not know which category their crime fell into.

When each respondent finished, he turned in his questionnaire, sealed in an envelope, and signed his name to a list to receive the $5 payment for participating.

Procedures for men in disciplinary segregation varied. In Texas, they were not allowed to participate; in California, administrators went to the man's cell. In one Michigan prison, the men in disciplinary segregation were brought to one group session; otherwise individual administration was arranged.

[18]A hand- or typewritten pass issued to inmates telling them to report to a particular place at a specified time.

[19]Usually one administrator would read the questions to the one or two nonreaders. Sometimes special sessions were held for nonreaders.

The Questionnaire

The self-administered questionnaire, which took about 50 minutes for the average respondent to complete, asked about:

1. Background in crime (e.g., age first arrested, number of felony convictions).
2. Attitudes toward criminal behavior and criminal justice system.
3. Criminal behavior during a specific time period, arrests during the street months, and current convictions.
4. Other behavior during the specific time period, such as use of drugs and alcohol, employment, and changes in residence.
5. Motivations for criminal behavior.
6. Details of crime resulting in the current conviction.
7. Arrests, incarceration, and criminal behavior during two earlier reference periods.
8. Demographics.
9. Participation in prison treatment programs during current term.

This report evaluates the arrest and conviction data in the third area listed above. Relevant pages of the questionnaire are reproduced in Apendixes A and D.

Informed Consent Procedure

Each respondent signed an Agreement to Participate (Appendix E, Exhibit E.2) which he read while the session administrator read it aloud at the beginning of the session. The form described: the purpose of the survey, the questions, the name/questionnaire link, the record check, confidentiality, lack of protection from legislative subpoena, voluntary participation, and the $5 payment benefit. Few questions were asked about the Agreement. When questions were asked, they were usually about the legislative subpoena and what the answers would be used for. Although the form mentioned the record check, few respondents raised questions about it.

Questionnaire Street Months Period

The questionnaire asked about criminal behavior during a time period called the Street Months, which the respondent defined for himself

following questionnaire instructions. The months ended with the (earliest) month of his arrest for his current conviction offense(s). The period began two Januaries before this (or one January before if the arrest month was January) and excluded any months during which the respondent was in jail or prison. The length of the Street Month period varied among respondents from 1 to 24 months. The respondent marked his complete Street Month time period on a large card and could refer to it when necessary while completing the questionnaire. Session administrators frequently helped respondents calculate the time period. Appendix A, Exhibit A.1, shows the set of questions used to determine the Street Months; Exhibit A.2 is a sample filled-in reminder card.

Record Street Months Period

The record abstractor used the Street Months claimed on the questionnaire which was relisted on a separate form (the "Transmittal Sheet," not shown).

In some cases the end date obtained in the survey did not correspond with the dates of arrest and disposition on the official record. Usually the record showed a later date of arrest for the current conviction crime(s). When this occurred, the record coder was instructed to use the date on the official record as the end date for determining the Street Months. In Michigan 19 percent of the cases had differences in Street Months between the survey and record; in California the number was also 19 percent; and in Texas, 14 percent. Later we describe the impact these Street Months differences had on the data quality.

The Questionnaire Retest

A small subsample was asked to fill out the questionnaire again about 7 to 10 days after the first administration. They were paid an additional $5 for participating.

The retest sample was randomly chosen during initial sample selection, with the constraint that retest sessions would not be scheduled in all prisons. "Replacement" respondents were never retested. Three hundred and twenty-three persons who completed the first questionnaire were asked to complete a retest; 252 did (66 in California, 45 in Michigan, and 141 in Texas).

The sessions were comprised exclusively of retest respondents and the procedures were identical to the original sessions except for the explanation that the retest was being conducted in order to find out how good the questionnaire was.

The Record Check

A record check of 1290 of the 1469 completed primary questionnaire cases was conducted by a separate team of coders, using records maintained by the state corrections departments. Michigan and Texas keep these records at the Department of Corrections offices; California keeps them only at the individual prisons.

The record check, which abstracted over 100 variables per case, supplemented the survey data in several areas and provided the evaluation study with the material needed for the survey and record comparison. In Michigan the coders pulled the selected inmates' folders from file stacks, but in California and Texas the coders requested folders from the files, and inmates or department personnel who worked in the files section pulled and delivered the folders.

In California and Texas there was only one folder for each inmate. In Michigan, if an inmate had been released on parole from a sentence but was brought back on a new charge before the former sentence expired, he was given a new folder for the second conviction and both folders would be current until one term expired. This meant an additional check by the coder to be sure he had the most recent folder.

The present evaluation uses arrest and conviction information from these records. Arrest information was abstracted from local, state, and federal RAP sheets; conviction information came from a Cumulative Case Summary Sheet in California, a Basic Information Sheet in Michigan, and directly from computer tapes in Texas.

Coding the evaluation study record variables (current conviction, and arrests during the street months) took an average of six minutes per case. Many cases (for example, first conviction with only a single arrest and conviction charge) took much less time to code because the folder contained fewer forms and the RAP sheets were easy to read. The arrest and current conviction data were coded in all states by the same coder.

Specific abstracting and coding procedures are described later. In the sections to follow, we present the results of the bias, reliability, and individual difference analyses using the methods described above. Section II evaluates the current conviction data; Sections III and IV contain evaluations of the self-reported arrest information.

II. CURRENT CONVICTION OFFENSES

This section examines the quality of the reports of charges on which prisoners were serving prison sentences at the time of the survey. The bias and reliability scores, derived by comparing questionnaire and record data, are used. We conclude that, on a general level, the data are close to unbiased. On an item-by-item basis, however, moderate amounts of bias are sometimes found and there appears to be some systematic confusion about item definitions in two of the states. We estimate reliability to be moderately high, but there appears to be enough unreliable response variance so that correlations of a conviction variable with another variable will be attenuated at least 20 percent on the average. Prisoner demographic characteristics and memory problem indicators do not predict the errors.

SURVEY QUESTIONS

About two-thirds of the way into the booklet, the respondent was given a checklist of fifteen conviction offense categories and asked to check the ones on which he was currently serving time. The checklist is reproduced in Appendix A, Exhibit A.3. The time reference period ("now") is easily understood,[1] and the information sought is not especially sensitive or threatening. Memory decay effects seem unlikely; any memory interference effects may be small since prisoners seldom serve sentences on more than a few different kinds of offenses at any one time. The checklist format is felt by many questionnaire designers to promote accurate recall because it is a recognition task rather than a more difficult recall or reconstruction problem. On the other hand, prisoners may have difficulty deciding which checklist category their conviction offense belongs to.

RECORDS

Information from the official records about the sampled prisoner's current conviction was obtained in slightly different ways across the states.

In California, project staff reviewed inmate records maintained at each institution. The coder used the Cumulative Case Summary Cover Sheet in the inmate's folder to obtain the current conviction offense(s). He matched the current offense information from the record with the appropriate offense type listed on the coding form. The match was made using the statute number of the offense when available and good judgment about the appropriate category when it was not. The number of offenses included in a category varied, ranging from only one to more than twenty. The coder entered a "1" next to each offense category on the form for which the inmate had a current conviction. See Appendix B, Exhibit B.1 for a reproduction of the applicable sections of the California coding form for current conviction.

In Michigan, the current conviction information came from the inmate's folder(s) which are kept at the Department of Corrections Headquarters. Project staff reviewed the current conviction information that was available on the Basic Information Sheet contained in each folder.[2] This sheet provides summary data about the inmate including the current conviction offense(s). In the case of multiple folders, the coder had to be careful to check that these offenses were not over- or undercounted across the multiple current folders. Each unique current conviction offense was noted on the coding form, again using the matching procedure, and coded according to whether it was the most current offense, next most current, etc. See Appendix B, Exhibit B.2 for the Michigan coding form.

In Texas, the official version of current convictions came from computer tapes furnished by the Texas Department of Corrections. Relevant information was transferred from the Department of Corrections tape to the research data base without an intervening abstracting, and coding step. The data from Texas showed how many convictions the inmate had for each offense category.

---

[1] Confusion sometimes arose if the inmate was given consecutive sentences or if he was also serving time on an old conviction for which he violated his parole.

[2] An inmate might have more than one current folder if he is serving a new sentence and a former sentence concurrently.

The Record information from all three states was recoded for this analysis to indicate the presence or absence of a current conviction in each offense category. The analysis is confined to the categories explicitly mentioned in the questionnaire except that a few categories were collapsed. Drug Possession and Drug Sales were combined into one category, and Forgery and Fraud were collapsed into one category.

CURRENT CONVICTION BIAS

How good are the survey data for estimating population means? We use the bias scores for this evaluation issue. An overview is provided by the summary bias scores that reflect the net error in the responses over all items. Net error is small in all three states. The individual offense category bias scores, however, are not all close to zero. We find some confusion about classifying conviction offenses into their proper categories. Some of these misclassification errors "average out" over items but make it difficult to interpret analyses using individual item data.

Summary Bias Scores

The observations contain a small amount of bias (± 6 percent) when summed over conviction offense items in each of the states.

Table II.1 shows that, in California and Texas, respondents reported 6 percent more conviction categories than records. In Michigan, there was a net underreporting of 6 percent. The Summary Bias Score (col. 3) can be viewed as the difference between the number of conviction categories per person reported in the survey (col. 1) and the number found in records for the same people (col. 2).[3] Column 5 shows the score relative to the record values for easier interpretation. Column 4 shows

[3] In theory, the Summary Bias Scores could be derived by summing the Item Bias Scores in Table II.2. However, because the item scores have been rounded, their sum will sometimes be different from the more precisely calculated Summary Bias Scores.

the estimated standard error of the Summary Bias Score. We interpret the bias score as statistically significant (and asterisk the estimated value) if its 2 standard error confidence interval does not include zero.

The bias scores are very small and suggest that on the average neither the questionnaire responses nor the records contain large, systematic distortions of information at this summary level of analysis. However, there is some variation among items and among respondents in how closely the survey and record information agree. This variation is discussed next. The reader should remember, however, that the discrepancies detailed below tend to "average out" and, therefore, do not have large effects on estimating the first moment of the population distribution of combined conviction offense types.

TABLE II.1

PER PERSON SUMMARY BIAS SCORES FOR CURRENT CONVICTION REPORTS BY STATE

State	(1) Survey Sum	(2) Record Sum	(3) Summary Bias Score	(4) Standard Error of Score	(5) Summary Bias Rel.to Record	(6) No. of Respondents
California	1.75	1.65	.10	.06	+6%	329
Michigan	1.31	1.39	-.09	.05	-6%	336
Texas	1.33	1.26	.07*	.03	+6%	632

NOTES:
Excludes respondents for whom record information was unavailable.
Excludes respondents who did not answer (did not provide quantifiable responses) to all current conviction items.
For a description of the Summary Bias Score see Sec. I., equation 6.

*2 Standard Error Confidence Interval does not include zero.

Single Item Biases

Many of the individual item bias scores indicate significant amounts of questionnaire and record disagreements. The patterns of item bias are reviewed here, but we find few consistent effects across the three states. We find what we call "substitution confusion" among the items: misclassification of conviction offenses into related categories. The source of the misclassifying is not clear; it could be the respondents, the records, the coding rules, or the coders. There is some evidence that rules out the research staff coders (who executed the coding rules) as a major source of misclassification error.

The discussion of conviction item bias is based on Appendix Tables C.2., C.3., and C.4. Relevant information is summarized in Table II.2.

Across the three states, there are thirty-nine separate comparisons of survey and record data for conviction offense classes (thirteen in each state). Twenty-four of these comparisons show statistically significant differences, a much greater number than expected by chance.[4] Michigan and Texas have the greatest number of statistically significant differences (nine and ten respectively) and California the fewest (five). Some explanation of these differences is desired, one that accounts for the substantial amount of apparent bias at the single item level and the small or no bias result on the summary indexes.

The first analysis looks for item biases that occur in the same way (same sign and statistically significant) across all three states. Since the Summary Bias Scores are close to zero, it would be surprising to discover many systematic bias patterns across the states.

No item meets the across-state criterion. The absence of consistent bias across states is more easily seen by consulting Table II.2 where the single Item Bias Scores for all three states are listed

[4]Using the two-tailed probability of .01, the expected number of statistically significant differences is less than one on the basis of chance alone.

Table II.2

ITEM BIAS SCORES FOR CURRENT CONVICTION ITEMS BY STATE

Conviction Category	California	Michigan	Texas
Assault	.07*	-.06*	.04*
Auto Theft	.02	.04*	.04*
Burglary	-.02	-.04*	-.06*
Drugs (Possession & Sale)	.03*	-.01	.02*
Forgery, Card & Fraud	.02*	-.01	.03*
Kidnap	.01	.00	.01*
Murder	.01	.01	-.02*
Stolen Property	.00	-.03*	.02*
Rape	-.01	-.03*	-.00
Robbery	.00	.08*	-.00
Other Sex	-.00	.05*	.01
Theft	.03*	-.07*	-.04*
Weapons	-.08*	-.03*	.02*

SOURCE: Appendix Tables C.2, C.3 & C.4

NOTES:
California N = 329
Michigan N = 336
Texas N = 632

*2 Standard Error Confidence Interval does not include zero.

together. Whatever is producing the large number of statistically significant bias scores is not operating consistently across the states. This rules out questionnaire effects, such as the position of the question in the checklist, as important causes of bias.

The second analysis examines the possibility of systematic confusion in assigning a conviction offense to one of the thirteen questionnaire categories. The results suggest that such confusion does take place in Michigan and Texas.

Our misclassification hypothesis is that the data source (questionnaire or record) contains the information about the conviction offense (viz., there is not an omission bias) but that it is sometimes put into the wrong category.

We use the hypothesis to guide the exploratory analysis as follows:

1. We search for negative correlations of the Item Bias Scores across pairs of items.
2. We note whether the negatively correlated items also have oppositely signed Item Bias Scores.

The item pairs that meet these criteria are listed in Tables II.3 to II.5; respondents, records, or record coders are probably misclassifiying offenses between elements of each pair. For example, some offenses that should be classified as "Theft" in Michigan are classified under Robbery by the respondent (or Robbery offenses are misclassified as Theft by the records or record abstractors).

Some evidence for systematic misclassification is found in the Michigan and Texas data. Negative correlations among the Item Bias Scores are found four times (seven items) in Michigan and three times (six items) in Texas. In all cases, the paired items had oppositely signed Item Bias Scores. Thus, the hypothesized systematic misclassification errors probably occurred. The data cannot be used to determine the source of the errors: the respondent, the records, the record coders, or the analysts. It can be noted that Texas record data came directly from a computer tape without intervening record abstracting while Michigan record data were assembled and transcribed by project coders. Since the two states show roughly the same levels of misclassification errors, it seems reasonable to rule out the

Table II.3

SUBSTITUTION MISCLASSIFICATION ANALYSIS
FOR CURRENT CONVICTION OFFENSES
CALIFORNIA PRISON SAMPLE*

Conviction Category	Single Item Bias Score	Negative Interitem Correlation of Item Bias Scores
Assault	.07	NO NEGATIVE INTERITEM BIAS CORRELATIONS
Auto Theft	.02	
Drugs	.03	
Forgery, Fraud	.02	
Theft	.03	
Weapons	-.08	

NOTES:
*Single item bias scores and negative correlations of item bias scores across items

Table II.4

SUBSTITUTION MISCLASSIFICATION ANALYSIS
FOR CURRENT CONVICTION OFFENSES
MICHIGAN PRISON SAMPLE*

Conviction Category	Single Item Bias Score	Negative Interitem Correlation of Item Bias Scores
Assault	-.06	-.25
Robbery	.08	-.16
Theft	-.07	
Auto Theft	.04	-.17
Stolen Property	-.03	
Rape	-.03	-.78
Other Sex	.05	
Burglary	-.04	(None)
Weapons	-.03	

NOTES:
*Single item bias scores and negative correlations of item bias scores across items

Table II.5

SUBSTITUTION MISCLASSIFICATION ANALYSIS
FOR CURRENT CONVICTION OFFENSES
TEXAS PRISON SAMPLE*

Conviction Category	Single Item Bias Score	Negative Interitem Correlation of Item Bias Scores
Assault	.04	-.29
Murder	-.02	
Auto Theft	.04	-.25
Theft	-.04	
Burglary	-.06	-.15
Drugs	.02	
Forgery, Fraud	.03	(None)
Kidnap	.01	
Stolen Property	.02	
Weapons	.02	

NOTES:
*Single item bias scores and negative correlations of item bias scores across items

project staff coders as a major source of the errors.[5]

On the other hand, some items from each state contain statistically significant bias scores that cannot be explained by the systematic misclassification error hypothesis. In California, five or six items are in this category. Texas has four items of this kind. Random misclassification or other error processes may be responsible for producing these biases. Since nine of the twelve remaining biased items have positive Item Bias Scores, we can narrow our speculation to processes that produce an excess of questionnaire over record reports. The process could involve systematic overreporting by respondents on the selected items or systematic omissions in the record data available for analysis. The deliberate exaggeration hypothesis seems counterintuitive. Reasons for falsely claiming a conviction category might be to enhance the prisoner's public image. If he wished to present a "bad guy" image, he would have falsely reported conviction categories that are serious and violent. Few (three of eleven) of the overreported offenses are of this type. The counterargument is that some respondents wish to claim a less serious set of conviction offenses than is true. While the data superficially support this idea (the overreported categories are nonviolent and the underreported category is "weapons"[6]), we would also expect these prisoners to omit mentioning their more violent and serious conviction offenses (and the data we are examining here do not conform to this expectation).

---

[5]Jan Chaiken speculates that the misclassification errors arise because some respondents don't remember the plea bargaining effects on their conviction offense. In Michigan, for example, Robbery is "overreported" (positive single Item Bias Score) and linked to "underreports" of Assault and Theft. It is possible that the plea bargaining resulted in a Robbery charge being reduced to an Assault or Theft charge in return for a guilty plea. The respondent, however, forgets this and reports the original robbery charge as his conviction offense. Chaiken acknowledges that this is not a plausible explanation for the underreporting of Murder and the linked overreporting of Assault in Texas.

[6]California prisoners underreported weapons convictions. Some may have forgotten that use of a weapon during a crime carries an almost automatic weapons conviction.

Respondents may accidentally overreport if, for example, they check all categories on which they have ever been convicted in their lifetime.

A more likely explanation is that the record information available for analysis does not contain all the less serious conviction offense information. We have no empirical evidence to support this but offer the following logical argument instead.

First, documents describing a prisoner's most recent convictions may have been delayed at the local level, not yet filed, or removed from a central file (for temporary administrative reasons) without research staff coders knowing this. If so, then our analysis files will not contain all of the true current conviction offense data.

A second argument is a variant of the original substitution misclassification hypothesis: it says that record coders (or the classification rules they follow) systematically misclassify minor conviction offenses as not belonging to one of the thirteen questionnaire categories (e.g., misdemeanor offenses). This would result in some omissions within the official record data available to researchers and would cause some of the item bias signs to be positive. Since coders did not record every conviction offense they encountered, it is not possible to test the hypothesis.

CURRENT CONVICTION RELIABILITY

Reliability estimates, obtained by two different methods, are presented here for the thirteen current conviction items. Estimates obtained by the retest method are higher than estimates using the record method. Since the retest estimates may be artificially high, we recommend that readers be guided by the results of the record method. The results suggest (1) that current conviction item reliabilities are higher than the arrest item reliabilities but (2) that there is enough response error variance to moderately attenuate subject matter correlations. We provide upper and lower bound estimates of the degree of attenuation to be expected (the MCAF estimates).

Average MCAF Estimates

To get an overview of the results, Table II.6 shows the arithmetic

Table II.6

ARITHMETIC AVERAGES OF CURRENT CONVICTION ITEMS
MCAF ESTIMATES

	Record Method			Retest Method	
State	Lower Bound*	Upper Bound**	Number of Respondents	$\sqrt{r}$ ***	Number of Respondents
California	.69	.83	329	.89	53
Michigan	.66	.81	336	.91	42
Texas	.65	.81	632	.93	122

NOTES:
*Average Correlation of the Survey and Record
**Square Root of the Average Correlation of the Survey and Record
***Average of the Square Roots of the Survey and Retest Correlations

average of three estimates across the thirteen conviction items.[7] MCAF coefficients are Multiplicative Correlation Attenuation Factors described in Sec. I. Like all other reliability coefficients, high values are desirable; low values indicate that estimates of relationships between this and another variable (e.g., correlations, regressions, nonparametric coefficients of association) will be biased toward zero.

Two estimates are derived by the record method based on the correlation of survey and record values. The lower bound number, about .66 in each of the states, assigned all the error variance to the questionnaire measure (it assumes the record values are perfect). The upper bound coefficient, about .81 in each state, assumes that both the questionnaire and record contain equal amounts of error variance. Note that there is almost no difference between the states in the record method estimates. Thus, on the average, when analysts estimate a correlation between a yes-no conviction item and another (perfectly measured) variable, the expected value of the correlation coefficient will be only 65 to 83 percent of the true population value.

The retest method yields one estimate of attenuation due to unreliability. Across the states, it averages between .89 and .93 for current conviction items, which is higher than the estimates from the record method. The retest procedure will overestimate reliability if respondents repeat their individual mistakes on the retest[8] and we suspect that this happens. For example, a respondent is very likely to remember his original response and give it again on the retest, especially if the question causes him difficulty. Repeating "yesterday's" answer is a lot easier than trying to think through the difficulty anew.

We recommend that readers be guided mainly by the upper bound MCAF estimate obtained by the record method. It is based on what are probably the most realistic assumptions about the sources and characteristics of the random error distributions, given the sets of assumptions

[7] To minimize the effect of rounding errors on the new average, the upper bound MCAF estimate is the square root of the average lower bound estimate.

[8] That is $\text{cov}(e_1, e_2) \neq 0$, where e is response error and the 1 and 2 index the two trials.

we have to choose from. This choice implies that the abstracted record data are not perfect and that respondents do repeat some of their original mistakes when answering the retest questionnaire.

Item MCAF Estimates

Next we examine the single item MCAF estimates, looking for consistent problem patterns across the three state samples. We define a "problem item" as an item that has an upper bound MCAF estimate of less than .70. The discussion is based on information in Appendix Tables C.5, C.6, and C.7, and summarized in Table II.6.

Very few problem items are found by either the record or retest estimation method and no item is a problem in two or more states. The California retest problem item is Auto Theft. In Michigan, the record method indicates unreliable responses about Auto Theft and Sex Offenses Other Than Rape. Reports of convictions for Possession or Receiving Stolen Property are very unreliable in Texas (record method) as are reports of Weapons convictions (retest method).

We conclude from these data that no single current conviction item poses consistent reliability problems across the states.

PREDICTORS OF BIAS AND TOTAL ERROR

We explore next the predictive potential of the ability, memory, demographic, and street-time difference variables on the Respondent Bias and Total Error scores. Our conclusion is that none of the variables account for important amounts of variance in the conviction item errors.

Six equations are involved (three states, two dependent variables each, Appendix Tables C.8 and C.9). In three equations none of the dependent variable variance was explained (adjusted R^2 = .00 or negative). Both Texas equations and the Michigan Total Error equation were in this category. The only effect consistent across two states (California and Texas) was the positive relationship between street-time differences and the Total Error Index score. The interpretation of the effect is that when the version of the Street Month period coded from records ends later than the respondent's version, the total

error score (reflecting both systematic bias and random error) is larger. Thus, we may be observing the effects of a coding or survey-record matching problem.

On the positive side, age does not predict the errors, eliminating a possible confounding effect in subject matter analyses that use the age variable.

The general conclusion, again, is that the tested individual difference variables are not important predictors of the small amount of current conviction bias and total error observed.

SUMMARY

The conviction offense data show a mixture of large and small, positive and negative item biases that average out to little or no overall bias. Part of the reason for the mixture is the confusion about classifying a conviction offense among the available categories. The small positive biases remaining, we speculate, are more likely to reflect record coding omissions than questionnaire overreports. The answers to the questionnaire are moderately reliable and no item has severe reliability problems in more than one state. We were unable to account for a meaningful amount of the bias or total error using a set of demographic, ability, memory, or match error variables.

In general, the prisoner respondents do not appear to be systematically denying their conviction offenses in the questionnaire. Their answers are somewhat unreliable, which will cause some attenuation in estimates of correlations.

III. ARREST INCIDENTS

Bias and reliability estimates for Arrest Incident data are given in this section. We compare official record information about arrest charges for each prisoner with the prisoner's questionnaire report. Nine offense categories are considered: Burglary, Robbery, Assault, Murder, Auto Theft, Theft, Forgery, Fraud, and Drugs. Bias is estimated by noting differences between questionnaire and record means. Reliability is assessed first by the correlation of questionnaire and record responses and second by the test-retest correlations on a smaller sample of prisoners.

The major results are:

- o Positive Summary Bias, for reports of the number of arrests in the listed categories (2 states);
- o Minimal Bias, for reporting any arrest in the listed categories (2 states);
- o Inconsistent patterns of statistically significant bias scores for single offense categories; the patterns do not repeat across the three state samples;
- o No evidence of systematically misclassifying arrests among the nine offense categories;
- o Low-to-moderate item reliabilities when assessed by the record method;
- o Moderate-to-high reliabilities when assessed by the retest method;
- o Low ability to predict error scores using demographic, memory, ability, and match problem variables.

QUESTIONNAIRE ITEMS

Each respondent was asked to indicate the number of times he was arrested within a "street months" time period for each of nine categories of charges. He was instructed to count an arrest even if he did not do what he was arrested for.

"Street months on the calendar" refers to a time period, up to twenty-four months long, which the respondent calculated earlier in the questionnaire. The Street Month period ended with the month that the respondent was (first) arrested for the offense he is now[1] serving time on. It began in January of the year before this arrest and excluded any intervening months in prison or jail.

The nine arrest offense categories are listed in Appendix A, Exhibit A.4.[2] The question item asks for the number of times the respondent was arrested for each crime category. Record data (described below) contain arrest charges which may be different from "times."

RECORD INFORMATION

Arrest information for the recorded Street Month period was abstracted onto coding sheets from RAP sheets in each inmate's official record folder. Each inmate folder usually contains at least one RAP sheet, which is a chronological listing of an individual's arrests, convictions, and sentences. The RAP sheet may have been generated at the local, state, or federal level. Ideally, each entry shows:

(1) The date of action (month/day/year);
(2) The arresting agency, receiving jurisdiction, or correctional institution;
(3) The name of the subject; and
(4) Charge(s) arrested for or convicted of (including the number of counts per charge) and disposition.

An example RAP sheet, from California, is in Appendix E, Exhibit E.3.[3]

Each arrest count within the Street Month period was tallied on a coding form (see Appendix B, Exhibits B.3, B.4, and B.5) in the appropriate type-of-crime category. The crime categories on the coding form included all those in the two questionnaire arrest sections plus

---

[1] At the time he completed the questionnaire.

[2] Spanish version is in Appendix D, Exhibit D.4.

[3] The example was taken from an earlier study.

additional ones such as arson, homicide, and kidnapping. The number of additionally listed crime categories varied across states (not shown). The analyses described here use only the categories in the relevant questionnaire section.

The way information is listed on RAP sheets varies within and across states. An arrest charge is sometimes spelled out, other times it is abbreviated, and sometimes the Penal Code section number is used. The number of counts is not always listed. A single arrest charge may be entered several times. We suspect, also, that some arrests are not listed on RAP sheets because, for example, the charge was dropped or never filed or because the arresting agency does not report information to all jurisdictions maintaining RAP sheets for an individual. Differing local practices may cause some juvenile arrests to be omitted or unreported to other jurisdictions. Finally, in the case of parolees, it is sometimes the policy to record only the parole violation when an arrest has occurred on charges that will not be prosecuted.

A review of the example RAP sheet (Appendix Exhibit E.3) shows an instance of the kind of inconsistency encountered. The arrest incident on 1/23/63 (see * item in Exhibit E.3) shows an arrest for grand theft auto (we assume one count) and two counts of burglary. When we look at the disposition of this arrest, we see that two counts of burglary were dismissed, but an additional count of burglary and the one on grand theft auto resulted in a state prison term. There was no record of the third burglary count in the arrest description. It was possibly added by the District Attorney and therefore was not part of the police department's record which updated the RAP sheet arrests; or it was an earlier conviction reinstated because of a parole violation.

ARREST INCIDENT BIAS

Summary Bias

The Summary Bias Scores for arrest incidents are significantly positive in two states and negative (but not significantly different from zero) in the third.

These summary indexes, shown in Table III.1, reflect the difference between total number of reported and recorded arrests over the nine offense categories listed in the questionnaire.[4] For Michigan the .90 bias score reflects a 49 percent discrepancy between the survey and the record. The Texas score of .36 reflects an 18 percent discrepancy. The California score of -.36 is not statistically significant but reflects a -10 percent discrepancy.

Indirect evidence suggests that our Arrest Record data for Michigan and Texas are incomplete relative to the California data. Arrest omissions in records would cause the summary bias scores to be positive, which they are for Michigan and Texas.

We computed the ratio of recorded arrest incidents to recorded conviction categories over the nine crime categories covered by the arrest questions and adjusted for slight sample size differences. In California, there are about 3 recorded arrests per recorded conviction category; in Texas only 2, and in Michigan just 1.7. Assuming no differences in sentencing policies, we would expect these ratios to be the same in all three

Table III.1

PER PERSON SUMMARY BIAS SCORES FOR CONTINUOUS ARREST INCIDENTS BY STATE

State	Survey Sum	Record Sum	Summary Bias Score	Std. Error of Score	Summary Bias Rel. to Record	No. of Respondents
California	3.21	3.57	-.36	.23	-10%	335
Michigan	2.72	1.83	.90*	.21	+49%	342
Texas	2.40	2.04	.36*	.16	+18%	555

NOTES:
Excludes respondents for whom record information was unavailable.
Excludes respondents who did not answer (did not provide quantifiable responses) to any arrest incident item.
For a description of the Summary Item Bias Score, see Sec. I., equation 6.
*2 Standard Error Confidence Interval does not include zero.

[4] Missing item information was imputed at the group mean.

states. We don't think the aggregate conviction data are seriously biased because of the high summary agreement rates (Table II.1). So we suspect that the Michigan and Texas record arrest data are relatively incomplete.

State	Recorded Arrest to Conviction Ratio
California	3.1
Michigan	1.7
Texas	2.0

There are two parts to each arrest category report, whether the prisoner was arrested on the category charge and if yes, the number of arrests (or arrest counts) in the category.

We constructed a Summary Bias Score for the yes-no version of the arrest incident questions and found, in Table III.2, that two of three states have unbiased category reporting while the third, Michigan, exhibits a positive bias. This suggests that on the average, the bias problems are mainly in obtaining agreement about the frequency of arrest incidents rather than in whether arrests occurred. The remainder of the discussion, then, will treat the two kinds of arrest incident reporting separately.

Table III.2

PER PERSON SUMMARY BIAS SCORES FOR YES-NO VERSION OF ARREST INCIDENTS BY STATE

State	Survey Sum	Record Sum	Summary Bias Score	Std. Error of Score	Summary Bias Rel. to Record	No. of Respondents
California	1.98	2.04	-.05	.08	- 2%	335
Michigan	1.63	1.35	.28*	.08	+21%	344
Texas	1.38	1.44	-.05	.05	- 3%	557

NOTES:
Excludes respondents for whom record information was unavailable.
Excludes respondents who did not answer (did not provide quantifiable responses) to any arrest incident item.
For a description of the Summary Item Bias Score see Sec. I., equation 6.
*2 Standard Error Confidence Interval does not include zero.

Yes-No Item Bias

The individual yes-no item bias scores exhibit a mixture of positive and negative coefficients with no consistent pattern across the states. The individual state data are in Appendix Tables C.10, C.11, and C.12 and are summarized here in Table III.3. No item has statistically significant yes-no bias coefficients in all three states.

Confining our attention to items with consistent effects in two of the states,[5] we see negative coefficients on the Burglary and Robbery items (which are moderately serious crimes) and positive coefficients on the less serious categories of Auto Theft and Drug deals.

Item	State	Single Item Bias Score	Item	State	Single Item Bias Score
Burglary	California	-.09*	Auto Theft	Michigan	+.09*
	Texas	-.10*		Texas	+.07*
Robbery	California	-.08*	Drugs	California	+.09*
	Texas	-.06*		Michigan	+.07*

None of these effects is due to the kind of substitution misclassification found in the current conviction data. There are many possible confusion candidate pairs (oppositely signed Item Bias Scores in the same state). But the Item Bias Score correlation for none of the pairs is negative[6] so respondents (or records or coders) are not necessarily making consistent substitution mistakes (e.g., classifying Auto Theft arrests as Burglary arrests).

Exactly what is producing the yes-no item biases is not clear. They could be respondent errors, record errors, coding errors, or they could be due to linking the wrong official records to the questionnaires. However, we appear to have reached a limit on what our models and data can tell us in this area.

[5] Bias scores that have the same sign and are statistically significant in two states.

[6] Correlations not shown; coefficients less negative than -.10 were ignored.

Table III.3

SINGLE ITEM BIAS SCORES FOR YES-NO ARREST INCIDENT ITEMS BY STATE

Crime Category	California N	California Score	Michigan N	Michigan Score	Texas N	Texas Score
Burglary	320	-.09*	322	.04	536	-.10*
Robbery	317	-.08*	326	.02	516	-.06*
Assault	310	-.04	318	.04	514	.08*
Murder	310	.00	320	.01	508	-.01
Auto Theft	309	-.01	310	.09*	511	.07*
Theft	307	-.01	312	.02	508	-.08*
Forgery	307	.06*	312	-.01	513	.02
Fraud	301	.03*	308	.00	500	-.00
Drugs	307	.09*	319	.07*	513	.03

SOURCE: Appendix Tables C.10, C.11 and C.12.
NOTE: N = Number of Observations.
*2 Standard Error Confidence Interval does not include zero.

Continuous Variable Item Bias

Using the full range of data (the reported and recorded arrest frequencies) a similar mixture of positively and negatively signed item bias scores appears. (The results are in Appendix Tables C.13, C.14, and C.15, and summarized in Table III.4.) No consistent patterns of significant coefficients across all states exist although Auto Theft is significantly positive in two states (Michigan and Texas). Two items show no bias in all three states (Murder and Forgery). The remaining items show inconsistent patterns of bias coefficients across the states which is typical of "noisy" data that are unaffected by any single systematic source of distortion.

A beginning question is whether the observed biases are a general characteristic of the data or contributed by a small number of individuals making large errors (or having grossly inaccurate records). To test the sensitivity of the results to the presence of a few "outliers," we compare the Winsorized means to the untransformed item bias means in Table III.5 and find that, while the coefficients shift downward (in

Table III.4

SINGLE ITEM BIAS SCORES FOR CONTINUOUS ARREST INCIDENT ITEMS BY STATE

Crime Category	California N	California Score	Michigan N	Michigan Score	Texas N	Texas Score
Burglary	316	-.13	317	.22*	530	.19
Robbery	316	-.32*	324	.22	513	-.01
Assault	307	-.11	315	.08	512	.12*
Murder	307	.00	315	.02	507	-.01
Auto Theft	305	.03	307	.17*	507	.13*
Theft	303	.03	307	.11*	507	-.10*
Forgery	305	.02	311	-.05	512	-.01
Fraud	300	.05*	307	.00	499	.00
Drugs	305	.07	316	.13*	512	.05

SOURCE: Appendix Tables C.13, C.14, and C.15.
NOTE: N = Number of Observations.
*2 Standard Error Confidence Interval does not include zero.

a negative direction), the general pattern of results is preserved. The Winsorized means are computed by truncating the scores of the highest and lowest 1 percent of the score distribution to the next highest (or lowest) score. For example, if there were 200 observations arranged in order from most negative to most positive, we would take the two highest values (1 percent of 200 = 2) and reassign them the third highest value; we would also change the two most negative values to the third most negative value; then we would calculate the arithmetic average of the new distribution.[7]

The algebraically negative effect of Winsorizing suggests the presence of some large positive survey reports that are unconfirmed by the record and the absence of large positive record values unconfirmed by the questionnaire. The questionnaires and record item frequency distributions (not shown) are as might be expected from this result: The questionnaire arrest item frequency distributions generally have a longer

[7] See Chap. 10 of Carl Morris and John E. Rolph, *Introduction to Data Analysis and Statistical Inference*, The Rand Corporation, P-5819, June 1978. We do not report standard errors for the Winsorized means because they do not have a straightforward interpretation.

Table III.5

OUTLIER SENSITIVITY ANALYSIS FOR ARREST INCIDENT ITEMS:
UNTRANSFORMED AND WINSORIZED ITEM BIAS SCORES
BY CRIME CATEGORY AND STATE

Crime Category	State	ITEM BIAS SCORE N	Not Transformed	Winsorized
Burglary	California	316	-.13	-.16
	Michigan	317	.22*	.19
	Texas	530	.19	.07
Robbery	California	316	-.32	-.32
	Michigan	324	.22	.11
	Texas	513	-.01	-.08
Assault	California	307	-.11	-.11
	Michigan	315	.08	.07
	Texas	512	.12*	.11
Murder	California	307	.00	.00
	Michigan	315	.02	.01
	Texas	507	-.01	-.01
Veh.Theft	California	305	.03	.01
	Michigan	307	.17*	.18
	Texas	507	.13*	.11
Theft	California	303	.03	.02
	Michigan	307	.11*	.09
	Texas	507	-.10*	-.07
Forgery	California	305	.02	.02
	Michigan	311	-.05	-.04
	Texas	512	-.01	-.00
Fraud	California	300	.05*	.03
	Michigan	307	.00	.00
	Texas	499	.00	.02
Drugs	California	305	.07	.07
	Michigan	316	.13*	.12
	Texas	512	.05	.04

NOTE: N = Number of Observations.

*2 Standard Error Confidence Interval does not include zero. Standard errors for Winsorized distributions were not computed.

"right-tail" than the record frequency distributions. Record versions of arrest frequencies within a category seldom exceed seven instances per person while questionnaire values of ten, twenty, forty, and higher sometimes occur. While these high survey values are not confirmed by the record, we should not assume that they are necessarily untrue. In Texas and California, the record almost always mentions one or more arrests in the relevant category for the prisoner reporting large arrest numbers, but in Michigan, this kind of "confirmation" is less likely to take place. We suggest that some of the extreme values reported in the questionnaire may be true but, unfortunately, we cannot say how a future researcher might distinguish among true and false survey reports at the extremes.[8]

RELIABILITY OF ARREST INCIDENT DATA

The reliability analysis indicates that there is an important amount of response error variance in the arrest incident questionnaire data. The amount of unreliability and, hence, the degree to which estimates of association will be attenuated, is not clear from the data. The two methods of assessing reliability yield different results for Arrest Incident data. The upper and lower bound estimates of the MCAFs, produced by the Record Method, suggest that response reliability is low; the retest estimates indicate that the answers to questionnaire items have a moderate-to-high degree of reliability. This general pattern is found in all three state samples and is characteristic of just about all of the items. These patterns are also found for the other set of arrest items, Arrests for Crimes Done, discussed in Section IV. A discussion of the reasons for the different results is presented in Section V.

[8] We do suggest that additional research aimed at describing the completeness of state-level record data be undertaken in a representative sample of states. The procedures, suggested by Peter Greenwood, would begin with local level records and determine how much of the locally generated arrest and conviction data could be found in higher level (e.g., state or federal) records. To avoid an estimation bias, entries at state and federal levels should be traced back to local records also.

Table III.6

ARITHMETIC AVERAGES OF YES-NO VARIABLE
ARREST INCIDENT ITEMS
MCAF (RELIABILITY) ESTIMATES

	Record Method		Retest Method
State	Lower Bound*	Upper Bound**	$\sqrt{r}$ ***
California	.45	.67	.84
Michigan	.36	.60	.82
Texas	.43	.66	.76

*Average Correlation of Survey and Record.

**Square Root of Average Correlation of Survey and Record.

***Average of the Square Root of the Survey and Retest Correlation.

Table III.7

ARITHMETIC AVERAGES OF CONTINUOUS VARIABLE ARREST INCIDENT ITEMS
MCAF (RELIABILITY) ESTIMATES

	Record Method		Retest Method
State	Lower Bound*	Upper Bound**	$\sqrt{r}$ ***
California	.35	.59	.77
Michigan	.35	.59	.76
Texas	.37	.61	.70

*Average Correlation of Survey and Record.

**Square Root of Average Correlation of Survey and Record.

***Average of the Square Root of the Survey and Retest Correlation.

Average MCAF Estimates

The arithmetic averages of the MCAF estimates for Arrest Incident items (Tables III.6 and III.7) reflect the main reliability results: There is a low level of agreement between the questionnaire and recorded arrest incident data in all three states and for both the yes-no and the continuous version of the Arrest Incident variables. The agreement between original questionnaire and retest answers is higher, especially for the yes-no variables, but even here, important amounts of response unreliability are indicated.

The arithmetic average estimates are similar across the three state samples. Using the Record Method, the lower bound average MCAF estimates range from .35 in California and Michigan to .37 in Texas for the continuous variables and .36 (Michigan) to .45 (California) when the items are rescored dichotomously. The lower bound estimate assumes no errors in the record data, no survey/record matching errors, and no coding errors. The upper bound estimate allows for some error in the record data (or introduced by matching and coding). The continuous variable upper bound MCAF estimates range from .59 to .61 across states; for the yes-no variable the range is .60 to .67. The implication of these findings for subject matter correlation estimates is important: On the average the observed correlation of an arrest incident variable with another (perfectly measured) variable will be only 30 percent as large as the true population correlation. If the other variable is measured with error, the attenuation of the estimate will be greater.

MCAF estimates obtained by the retest method imply a much higher questionnaire reliability and, consequently, a much lower attenuation of correlation estimates. The arithmetic averages of retest MCAF estimates range from .70 to .77 across the states for the continuous variables (Table III.7), and from .76 to .84 for the yes-no variable (Table III.6).[9] The retest estimates, however, may be inflated due to a correlation of response errors across trials.

[9]Retest arithmetic averages were computed excluding items with no response variation (e.g., Fraud in California). If coefficients of .00 are imputed instead, the retest averages are decreased by about 13 percent each. Even with the imputation, the retest MCAF estimates are higher than the upper bound estimates from the record method.

Regardless of method, the MCAF estimates are potentially sensitive to the presence of a few extreme values. We have investigated the effect of removing questionnaire item responses greater than five (analyses not shown). Although the estimates do change somewhat, the direction of change is not uniformly toward higher MCAF reliability coefficients using either the record or retest method. A more direct demonstration of the weak effect of univariate "outliers" comes from comparing the yes-no variable MCAFs to the continuous variable MCAFs. The yes-no variable is a severe "truncation" of large responses because all responses greater than one are rescored as one. Tables III.6 and III.7 show that the average coefficients are only slightly higher for the yes-no variable than for the continuous variable, suggesting that the continuous variable coefficients are not especially sensitive to extreme values.

Item MCAF Estimates

We comment briefly here on the single item reliabilities which are listed in Appendix Tables C.16 through C.21. We retain the definition of a "problem item" as one with an upper-bound (or retest) MCAF estimate below .70.

Using the record method, we find that most continuously scored items are problems. Forgery is the only item never classified as a problem and Murder is not a problem in California and Texas. Arrests for these offenses are relatively few, but it is not clear that this has anything to do with their being reported and recorded reliably.

With the retest method, we find many fewer items appear to be reliability problems and no single item is a problem in all three states. Auto Theft is a problem on the Michigan continuous variable and on the Texas yes-no variable; otherwise no item is a problem in two of the three states.

The retest and record estimates produce radically different conclusions about the questionnaire reliabilities.[10] We suspect that the

[10]The reader interested in extreme uncertainty should look at the estimates for the continuous Robbery item in Michigan (Appendix Table C.20).

retest estimates overstate reliability because respondents may repeat their errors the second time the questionnaire is administered. Even if the retest estimates are discounted, the upper and lower bound record estimates are not always very close to each other. The problem remains that, although we observe disagreements between the questionnaires and the record, we do not have enough information to know how much of the disagreement is due to record error and what proportion originates in the questionnaire responses.

PREDICTORS OF BIAS AND TOTAL ERROR

Some of the variance in the Respondent Bias and Total Error variables is accounted for by our set of individual different predictors. The fitted equation values are shown in Appendix Tables C.22 and C.23.

Between 14 and 29 percent of the variance in the Michigan and Texas total error scores is "explained" by variables in the ability, memory, demographic, and street month difference sets. For California, however, neither equation has an adjusted R^2 greater than .02.

The variable with the most consistent effect is the number of skipped items in the Arrest Incident set. The positive coefficient indicates that respondents with more missing item responses also tend to have higher total error and (except California) higher average item bias scores. The finding probably indicates that skip errors and response errors are positively correlated for the arrest incident items. However, because the missing item scores must be imputed for the dependent variables,[11] the findings may only reflect an error in the method of imputing for skipped items.

In Texas, the number of prior felony convictions predicts both the Respondent Bias and the Total Error scores. The positive coefficient suggests that those with more prior convictions make more response errors. The effect is not present in the other state equations.

The implications of these results for analysis are twofold: First, some predictor variables such as race and education are unrelated to

[11]In effect, the missing item bias score is imputed as the mean bias for the person determined from answered items.

bias and total error (one exception). They can be used as predictors of arrest-related dependent variables with little danger of regression coefficients being artificially inflated by correlated response errors. Second, very little has been learned about characteristics of respondents with high error scores. The data, unfortunately, do not suggest a useful strategy of identifying respondents who are especially error-prone (and hence, introducing corrections for their errors).

IV. ARRESTS FOR CRIMES DONE

Respondents also reported the number of crimes they committed during the Street Months and were arrested for. These responses are compared to recorded arrests but, because survey and record definitions differ, we do not expect good agreement and we have difficulties interpreting the discrepancies. The major results are:

- o Positive Summary Bias for numbers of arrests (3 states)
- o Negative Summary Bias for Yes-No variable (3 states)
- o Patterns of statistically significant bias scores that are inconsistent across the state samples
- o No systematic misclassification of arrests among the categories
- o Low-to-moderate item reliabilities via the record method
- o High item reliabilities via the retest method
- o Some prediction of respondent error scores by ability, demographic and memory variables
- o Contrary to expectations, a pattern of results not much different from Arrest Incident Variables

QUESTIONNAIRE ITEMS

The questionnaire items covered here ask the respondent about the number of felony offenses he committed for which he was arrested. The felony offenses are those done during the Street Months Period (see Section I). Separate questions about ten offense categories were employed: Burglary, Business Robbery, Person Robbery, Murder, Assault, Theft, Vehicle Theft, Forgery, Fraud, and Drug Deals. The arrest question was not asked about murder commissions. Answers to the person and business robbery items were combined for this analysis (to conform to the record information about robbery which was not disaggregated).

For each offense, the respondent reported whether he committed any during the Street Month Period. If yes, he reported how many by

following the questionnaire logic. (Many respondents did not follow the logic correctly.) Next the respondent reported the number of committed offenses in the category that he was arrested for. The questionnaire arrest item acknowledged that the crime category and arrest charge category might be different. If the respondent said he did not commit the specific offense, but reported arrests, he was excluded from our analysis of that item. (See Appendix A, Exhibit A.5 for these questions.)

RECORDS

We used the same recorded arrest information for this analysis as we used for the Arrest Incident evaluation described in Section III.

DEFINITIONAL DISCREPANCIES

Section I discusses ways that the questionnaire Arrests for Crimes Done information may legitimately differ from record information:

1. The record arrest charge category and the questionnaire crime category may differ and both be correct.
2. Arrests for crimes not committed or crimes not listed on the questionnaire will not be reported in the questionnaire. These arrests may be in the coded record information, however.

In addition:

3. The questionnaire response may contain arrests that occurred after the end of the Street Month period if the offense leading to the arrest occurred within the street months. And the questionnaire responses may include arrests on misdemeanor charges that we did not code from official records.

All of these problems can result in underestimating reliability using the record method. The second and third problems may contribute to misestimating both the Item and Summary Bias Scores. (The second problem introduces a negative bias and the third introduces a positive bias.) The first problem causes Item Bias Scores to be

misestimated (the direction of the bias cannot be predicted); if both the record arrest charge and the questionnaire crime category are among the eight analysis categories, however, the Summary Bias Score will be unaffected.

BIAS IN ARRESTS FOR CRIMES DONE

Summary Bias

It is especially important to begin this analysis with an examination of the Summary Bias Scores. Because the questionnaires ask for arrests in the committed crime category rather than in the category of the actual arrest charge, we cannot expect an exact, category-by-category correspondence between the survey and the records, even if both sources are unbiased. If, however, both the committed crime and the arrest charge are any one of the eight offense categories in the questionnaire, the discrepancies will "average out" and the Summary Bias Score will be a purer indicator of any systematic survey and/or record distortions.

The Summary Bias Scores (Tables IV.1 and IV.2) tend to be positive for the continuous version of the variable (only the Michigan value is statistically significant) and negative for the yes-no version (the Michigan value is not statistically significant but the values for the other two states are). If the records are unbiased and definitional differences are ignored, this suggests that, on balance, the prisoners tend slightly toward denying committing and being arrested for the selected crimes, but if they do not deny, they tend to overstate how often they have been arrested for the offenses they committed.

The apparent exaggeration effect is especially strong in Michigan. Respondents reported 55 percent more arrests than were found in the records. Unfortunately, the source of the large bias is unknown. It could come from respondent exaggerations, record omissions, or from some of the definition differences discussed earlier. The Michigan bias in reporting yes-no arrest categories is small and not statistically significant.

Table IV.1

PER PERSON SUMMARY BIAS SCORES FOR CONTINUOUS VERSION OF ARRESTS-FOR-CRIMES-DONE ITEMS BY STATE

State	Survey Sum	Record Sum	Summary Bias Score	Std. Error of Score	Summary Bias Rel. to Record	No. of Respondents
California	3.66	3.32	.33	.29	+10%	344
Michigan	2.56	1.65	.90*	.30	+55%	360
Texas	2.02	1.90	.11	.14	+ 6%	585

NOTES:
Excludes respondents for whom record information was unavailable.
Excludes respondents who did not answer (did not provide quantifiable responses) to any Arrest-for-Crime-Done item.
For a description of The Summary Bias Score see Section I, equation 6.

*2 Standard Error Confidence Interval does not include zero.

Table IV.2

SUMMARY BIAS SCORES FOR YES-NO VERSION OF ARRESTS-FOR-CRIMES-DONE BY STATE

State	Survey Sum	Record Sum	Summary Bias Score	Std. Error of Score	Summary Bias Rel. to Record	No. of Respondents
California	1.48	1.86	-.38*	.08	-20%	344
Michigan	1.10	1.18	-.08	.06	- 7%	360
Texas	0.98	1.34	-.36*	.05	-29%	585

NOTES:
Excludes respondents for whom record information was unavailable.
Excludes respondents who did not answer (did not provide quantifiable responses) to any Arrest-for-Crimes-Done item.
For a description of The Summary Bias Score see Section I, equation 6.

*2 Standard Error Confidence Interval does not include zero.

The continuous Summary Bias scores are positive but not significant in California and Texas. The yes-no summaries, however, are significantly negative (-20 percent for California and -29 percent for Texas). So, in these states, the records contain a wider variety of felony arrest charges than the prisoners report but a smaller total frequency of arrests.

Single Item Biases

Again, no evidence of a systematic pattern of bias is found. No item has statistically significant bias scores of the same sign in all three states. Continuously scored items with coefficients not significantly different from zero in all three states are Robbery, Assault, and Forgery; Theft has significantly positive coefficients in California and Michigan. For yes-no scoring, only the Forgery item is unbiased in all three states; Robbery and Burglary have negative bias scores in two of the three states. The item-by-item data are shown in Appendix Tables C.24 through C.26 (continuous scoring) and Tables C.27 through C.29 (yes-no scoring). They are summarized in Tables IV.3 (continuous) and IV.4 (yes-no).

Because each arrest item potentially included arrests on different charges, we expected to find systematic misclassification errors, that is, an arrest systematically reported in one category and systematically recorded in another. As explained earlier, symptoms of systematic classification errors are a negative correlation of item bias scores that have oppositely signed item bias scores. Results (not shown) were not as expected; no statistically significant negative correlations were found when the items were scored dichotomously and only one was found for the continuous scoring case (California, Fraud and Forgery).

For these analyses we deleted cases who reported arrests for a crime category but said they committed no offenses in that category. In one state we formally examined the effect of these deletions on both the yes-no and continuous item bias scores. The state chosen was California because the serious crime items appear to have fewer questionnaire than record reports of arrests; hence, it was more

TABLE IV.3

SINGLE ITEM BIAS SCORES FOR CONTINUOUS ARRESTS-FOR-CRIMES-DONE ITEMS BY STATE

Crime Category	California N	California Score	Michigan N	Michigan Score	Texas N	Texas Score
Burglary	333	.07	352	.41*	566	.02
Robbery	326	.17	346	.12	566	.01
Assault	329	-.11	350	.03	571	.02
Theft	330	.29*	347	.34*	563	.04
Auto Theft	336	-.08*	351	.04	574	.06*
Forgery	330	-.03	356	-.04	562	.01
Fraud	338	.07*	353	.01	573	.02
Drugs	337	-.04	355	.02	574	-.04*

SOURCE: Appendix Tables C.24 C.25 & C.26

NOTES:

N = Number of Observations

*2 Standard Error Confidence Interval does not include zero.

TABLE IV.4

SINGLE ITEM BIAS SCORES FOR YES-NO ARRESTS-FOR-CRIMES-DONE ITEMS BY STATE

Crime Category	California N	California Score	Michigan N	Michigan Score	Texas N	Texas Score
Burglary	333	-.11*	352	.01	566	-.16*
Robbery	327	-.14*	346	-.04	566	-.09*
Assault	329	-.08*	350	-.04	571	.01
Theft	331	-.01	348	-.02	563	-.06*
Auto Theft	336	-.08*	351	.03	574	.00
Forgery	332	.02	356	-.02	563	-.00
Fraud	338	.02*	353	.01	573	-.00
Drugs	337	.03	355	.02	574	-.03*

SOURCE: Appendix Tables C.27 , C.28 , & C.29 .

NOTES:

N = Number of Observations

*2 Standard Error Confidence Interval does not include zero.

likely that an imputation bias might be operating. The pattern of bias results (not shown) was not affected importantly by whether the questionable cases were included. The signs of the item bias coefficients remained the same and the sizes did not change by more than one standard error. The number of affected cases averaged seven per item and ranged from two to sixteen over the eight items. Thus, the treatment of this kind of problem case does not appear to have important implications either for our evaluation analysis or for other analyses.

To see whether a few extreme cases determined the results, we Winsorized the full item bias score distributions and recalculated the means (Table IV.5). While truncating the extreme values sometimes has large effects on the means (generally moving them closer to zero), the general pattern of results remained unchanged. We conclude, therefore, that the observed item bias effects are not produced by a small number of atypical cases but are, instead, a more general characteristic of the whole set of observations.

RELIABILITY ESTIMATES

The uncertainty about whether questionnaire responses are reliable is as great for these arrest variables as for the Arrest Incident items discussed in Section III. Lower bound MCAF estimates are low (averaging in the high .20s or low .30s while the retest estimates are moderately high (averaging in the high .70s to low .90s). Record method estimates suggest that almost every item is a reliability problem while the retest method suggests few items have reliability problems. The reliability estimates are presented next.

Average MCAF Estimates

A summary of the MCAF estimates for Arrests for Crimes Done, obtained by unweighted arithmetic averaging,[1] is in Tables IV.6 and

[1]The average upper bound record estimate is obtained by taking the square root of the average lower bound estimate. Averages exclude items with no variation from both the numerator and denominator.

Table IV.5

OUTLIER SENSITIVITY ANALYSIS FOR ARRESTS-FOR-CRIMES-DONE ITEMS:
UNTRANSFORMED AND WINSORIZED ITEM BIAS SCORES
BY CRIME CATEGORY AND STATE

Crime Category	State	N	ITEM BIAS SCORE Not Transformed	Winsorized
Burglary	California	333	.07	-.04
	Michigan	352	.41*	.30
	Texas	566	.02	-.09
Robbery	California	326	.17	.09
	Michigan	346	.12	.07
	Texas	566	.01	-.04
Assault	California	329	-.11	-.10
	Michigan	350	.03	.01
	Texas	571	.02	.02
Theft	California	330	.29*	.21
	Michigan	347	.34*	.20
	Texas	563	.04	.05
Auto Theft	California	336	-.08*	-.09
	Michigan	351	.04	.06
	Texas	574	.06*	.03
Forgery	California	330	-.03	-.02
	Michigan	356	-.04	-.04
	Texas	562	.01	-.01
Fraud	California	338	.07*	.02
	Michigan	353	.01	.01
	Texas	573	.02	.02
Drugs	California	337	-.04	-.02
	Michigan	355	.02	.02
	Texas	574	-.04*	-.05

NOTES:
Symbols:
N = Number of Observations

*2 Standard Error Confidence Interval does not include zero.

and IV.7. Six estimates per state are shown; three for the continuous variables and three for the yes-no version of the variables. The first two numbers are the lower and upper bound estimates of the average Multiplicative Correlation Attenuation Factor obtained by comparing survey responses to official records. The third number in each set is the average retest estimate of the MCAF.

For any method of estimation, the results are very similar across the states: The lower bound coefficients are low in all states and the retest coefficients are moderately high.

With one exception,[2] MCAF averages are higher for the yes-no variables than the continuous variables. This suggests that the reliability problems get larger as either the questionnaire or record values exceed one. This can be caused by a few "outlier" cases who have very extreme values on the questionnaire or record (but not both) or can reflect an inequality of the coefficients of error correlated with true score ($b \neq b'$). If, for example, respondents exaggerate the number of true arrests above one or the records tend to omit more arrests if several are already recorded, this pattern of findings will be obtained.

The principal finding is that we have been unable to estimate average reliability within a range that is narrow enough to draw conclusions and make applied recommendations. The average MCAF coefficients range from .27 to .92 across states and estimation methods. We defer attempts at reconciliation until Section V.

Item MCAF Estimates

The two MCAF estimates yield very different conclusions about the single item reliabilities. The estimates and their implications are mentioned next, based on material in Appendix Tables C.30 through C.32 (continuous scoring) and C.33 through C.35 (yes-no scoring). A problem item is one whose upper bound (or retest) MCAF estimate is less than .70.

[2] California, Retest method average MCAFs.

Table IV.6

ARITHMETIC AVERAGES OF CONTINUOUS ARRESTS-FOR-CRIMES-DONE ITEMS MCAF (RELIABILITY) ESTIMATES

State	Record Method		Retest Method
	Lower Bound*	Upper Bound**	$\sqrt{r}$ ***
California	.34	.59	.92
Michigan	.27	.52	.77
Texas	.32	.57	.79

NOTES:
* Average Correlation of Survey and Record
**Square Root of Average Correlation of Survey and Record
***Average of the Square Root of the Survey and Retest Correlation

CONTINUED

1 OF 3

Table IV.7

ARITHMETIC AVERAGES OF YES-NO VARIABLE ARRESTS-FOR-CRIMES-DONE ITEMS MCAF (RELIABILITY) ESTIMATES

	Record Method		Retest Method
State	Lower Bound*	Upper Bound**	$\sqrt{r}$ ***
California	.43	.65	.84
Michigan	.34	.58	.82
Texas	.38	.62	.85

NOTES:
*Average Correlation of Survey and Record
** Square Root of Average Correlation of Survey and Record
***Average of the Square Root of the Survey and Retest Correlation

With the Record method, most items are problems in all states, regardless of whether the variable is scored over its entire range or just yes/no. The Robbery item, however, is a problem only in Michigan. The Forgery and Drug Arrest items are problems in only two or three of the six record evaluations (three states, two types of variable scoring). If these estimates are correct, analysts should be wary of using the low reliability arrest items in their analyses because a substantial amount of the measured variation between people is not true-score variation.

The retest method MCAFs indicate that no item is a problem in all three state samples, regardless of how the item is scored. Retest MCAFs are in the problem range for Theft (continuous variable scoring) in two states. Assault is a problem in the Michigan continuous variable and the California yes-no variable. If the retest estimates provide a true picture of item reliability, then analysts can proceed to use the nonproblem items in their empirical investigations, perhaps also adjusting their covariance matrices for the generally small amounts of non-true-score variation in the questionnaire measures.

The two methods of estimating MCAFs suggest very different things about item reliabilities and different conclusions about further analyses. We suggest future analytical approaches in Section V that may reduce the uncertainties. Before doing so, however, we investigate the possibility that some of the questionnaire/record inconsistencies in this set of arrest data can be predicted.

PREDICTORS OF BIAS AND TOTAL ERROR

We were not especially successful in predicting the bias and error scores using ability, memory, demographic, and street month difference variables. Results are in Appendix Tables C.36 and C.37.

As with Arrest Incidents, Michigan scores were most successfully predicted (highest R^2 = .14 for the Individual Difference bias) and Texas scores least predictable (R^2 = .02).

The Arrests for Crimes Done error scores are the only ones predicted by a demographic variable. The unstandardized regression

coefficients for age, shown in the accompanying text table, are negative. This suggests that older respondents make fewer errors than younger respondents (the effect is not statistically significant in California.[3]

Regression Coefficients, Age

Dependent Variable	California	Michigan	Texas
Respondent Bias Score	-.003	-.006*	-.004*
Total Error Score	-.006	-.008*	-.004*

*$p \leq .01$

The other variable with somewhat consistent effects is the number of felony convictions, a proxy for memory "set size." The coefficients shown in the text table support the interference hypothesis and suggest that higher error rates are associated with more true convictions (hence, more items in memory storage that must be processed to respond).

Regression Coefficients, Number Felony Convictions

Dependent Variable	California	Michigan	Texas
Respondent Bias Score	.05*	.03	.04*
Total Error	.03	.04*	.06*

*$p \leq .01$

Finally, the number of skipped items within the Arrests for Crimes Done section is positively associated with bias (two of three states) and total Error (Michigan only). The ambiguous interpretation possibilities are present here also. The two kinds of error (skipped questions and total error or respondent bias) could be positively associated, or an imputation method problem could be responsible (see previous discussion in Section III).

Race, a variable included in planned subject matter analyses, does not predict the error scores.

[3] Marcia Chaiken suggests the possibility that the age effect may be due to poor quality juvenile records for respondents who were juveniles during some part of the Street Months period.

V. DISCUSSION AND CONCLUSIONS

We began with two objectives: To test the widely held hypothesis that criminals will deny facts about their criminal careers on questionnaires and to describe the questionnaire measurement errors for future users of the data and for questionnaire designers in general.

THE DENIAL RESPONSE BIAS HYPOTHESIS

With a reasonably high amount of confidence we can conclude that the widely held hypothesis is not always true. For the sections of the questionnaire we examined, we seldom found that the convicted felons reported fewer arrests or convictions on the average than indicated in the official records. The results indicate either that prisoners exaggerated a little or that the coded official records did not contain all of the relevant information.

Readers can benefit from more detailed information about measurement errors in terms of reliability and validity, for groups of items, single items, and types of felons. We constructed bias scores to reflect invalidity and MCAF scores to reflect unreliability. We constructed a total error score and used it to search for demographic and other correlates of measurement errors. The evaluation analyses were conducted for each state sample separately, which provided an opportunity to replicate the evaluation results over several samples.

BIAS RESULTS

Summary Bias scores for groups of items tended to be close to zero or positive across sections of the questionnaire and across the three states. In terms of measurement errors, then, first moment statistics based on these item groups (such as means, proportions, or regression intercepts) will not be downwardly biased. We suspect that the positive summary bias scores for arrest incidents result from incomplete official records.

Bias for single items within these groups, however, exhibited a mixture of large and small, positively and negatively signed coefficients.

Departures from zero bias occurred much more often than expected by chance. The sizes and signs of the bias for any particular item generally were not consistent across the three states.

For only one group of items (current convictions) were we able to begin to understand this unexpected pattern of results: The conviction item biases are partly produced by systematic misclassification errors; e.g., a robbery conviction is sometimes entered incorrectly as a theft conviction either by the respondent, the official records, or by the record abstractor. The conviction is not denied; it is just entered into the wrong offense category. This partly explains why specific items appear biased more often than is expected by chance, but bias scores for the whole group of items are close to zero.

Unfortunately, much of the remaining item biases cannot be explained by systematic misclassifications. Our analyses indicate that outliers and some missing-data imputation procedures are not responsible for the observed bias results. Some remaining hypotheses have been offered but must await testing in new research. If there are systematic respondent biases, they probably involve an unwillingness to report committing crimes. This can result in underreporting categories of Arrests for Crimes Done. If the questionnaire causes bias, it probably does so in the Crimes-Done sections where the respondent must correctly execute a complex set of instructions even to reach the question about arrests. If the record check design is at fault, it is because we cannot quantify the effect of the different definitions used by respondents and records to report arrests and convictions. Finally, our analyses suggest the hypothesis that our coded version of the official records of arrests may be incomplete, causing continuously scored arrest items to be positively biased.

An item, however, seldom has the same amount and direction of bias in all three state samples. Further research is needed to look for the record, respondent, questionnaire, and definitional biases that operate differently across jurisdictions.

RELIABILITY RESULTS

We concluded that responses to the current conviction items are moderately-to-highly reliable, despite the confusion about where to

categorize specific conviction offenses. However, correlations of perfectly-measured variables with conviction items will, on the average, be only 80 percent as large as the true associations in the relevant population.

For arrest items, unfortunately, we have been unable to provide conclusive, useful reliability estimates. Our estimation problems stem from very high disagreement rates between questionnaire and record values, but very low test-retest disagreement rates. If we assume that the records are perfect, we conclude that questionnaire arrest responses are extremely unreliable (record method, lower bound estimates). If we assume that the questionnaires and records contain equal amounts of non-true score error, we conclude that both sources are moderately unreliable (record method, upper bound estimate). If we assume that the retest estimates provide the best information about response reliability (viz., that respondents don't make systematic mistakes and don't repeat random mistakes), we conclude that arrest responses have a moderate-to-high degree of reliability.

Without additional information, we have no formal way of choosing among these interpretations. Next we offer some comments that tend to support using the two bounds of the record method to estimate response reliability. Then, in the next subsection, we outline some additional analyses that can be carried out by subject matter experts to possibly reduce the uncertainty about reliability.

Could the presence of matching errors account for our low reliability estimates using the record method? The answer is yes, in theory, but probably no in practice. In theory, if we incorrectly linked questionnaires to records, we would observe low reliabilities by the record method and (possibly) high reliabilities by the retest method. (Retest reliability estimates are unaffected by record matching errors.) In practice, the matching errors would also cause low reliability estimates for the current conviction items and this did not happen. While we cannot rule out the possibility that match errors occurred,[1] we can say that they are not solely responsible for the pattern of low reliability estimates for the arrest items.

Could our arrest reliability results be due to records containing most of the error? The answer is theoretically yes and practically

[1] We were unable to build objective indicators of probable match errors.

uncertain. Of course, if the records contained most of the observed non-true-score error, this would result in low reliabilities estimated by the record method (the retest estimate is unaffected). One piece of evidence at least superficially refutes the hypothesis: The variances for questionnaire items are usually larger than the variances for the record items.[2] This would be a convincing refutation if we could assume that the record and survey response models contained only true score and random error terms, and that the true score variances were equal:

$$S = T + e_s; \qquad \text{var } S = \text{var } T + \text{var } e_s;$$

$$R = T + e_r; \qquad \text{var } T = \text{var } T + \text{var } e_r.$$

Then the measure (S or R) with the largest measured variance also has the largest error variance (unreliability). Under this model, then, because the survey items usually have larger measured variances than record items, the record could not contain more error variance (unreliability) than the survey. However, we do not think such simple models are appropriate. Recall, from Section I, that part of the unreliability can be due to response or record bias correlated with true scores (the term bT_{ij} or $b'T_{ij}$ in eqs. 1 and 2 and its subsequent expression in the reliability estimates). It is quite possible for the record b' parameter to have a negative value;[3] this would be consistent with high levels of record unreliability, low levels of total measured record variance and positive bias scores. Thus, the larger total measured variance, and positive bias scores. Thus, the sarily imply that the questionnaire responses contain more error.

Reliability estimates are potentially sensitive to outliers. We have presented the results without deleting outliers but have conducted many informal analyses that either excluded or truncated extreme

[2] For reasons of economy, we have not included these variances in the tables.

[3] A negative b' means that the number of arrests omitted from the record is positively related to the true number of times the person was arrested.

questionnaire and record values (several of these analyses are discussed in the text). Some estimates change as a result of treating outliers differently (not always toward higher reliability). The conclusions from the different estimates were almost always the same.

Retest reliability estimates are inflated if respondents repeat their initial errors. Is there direct evidence that respondents repeat their errors on the retest? The answer is no, but one analysis that would "rule out" the hypothesis failed to do so. We could rule out the "conditioning" hypothesis [$\text{cov}(e_1, e_2) \neq 0$] if we could show that responses actually got "better" on the retest compared to the original questionnaire. An index of "better" would be a higher correlation of the questionnaire and record responses. The data in Appendix Tables C.38 and C.39 indicate that record agreement does not consistly increase at retest. Therefore, we cannot rule out the possibility that response errors are repeated on the retest.

A final comment about the retest reliability estimates. A reviewer asked if a sample bias could explain why the retest reliability estimates were higher than the record method estimates. His hypothesis is that only cooperative, conscientious respondents would take the retest; hence this group would furnish higher quality data than the group as a whole. Our record method reliability estimates, of course, are based on the larger sample. We have compared primary questionnaire reliability (and bias) estimates for the retested and not retested groups. While small sample sizes limit the power of the significance tests, we failed to find any trend indicating that retest group responses were of generally higher quality. Thus, differences between the retest and record method reliability estimates are unlikely to be due to a sample bias.

LIMITATIONS ON THE MEASUREMENT EVALUATIONS

Some limitations on the validity and generalizability of our evaluation results should be mentioned explicitly:

1. The evaluations are of observation (measurement) errors only. We did not attempt to estimate sampling, record matching, form nonresponse, item nonresponse, refusal, or data reduction errors.

2. Our samples excluded respondents in jail because official record data were not obtained for them.

3. Record information was not available for all prisoners who filled out questionnaires. We have no reason to suspect that the excluded prisoners have response error characteristics that are different from the others, but have no empirical information to back up our assumption.

4. There was a substantial amount of item nonresponse for which we adopted various imputation strategies. Our imputations may have affected our findings and conclusions.

5. Our evaluations were limited to a small number of items within the larger questionnaire. We do not know whether our conclusions apply to the other questions.

6. Our conclusions about bias and reliability are only approximations because we are still unsure of the true structure of the survey and record errors. Our error models contain more parameters than is usual for this kind of measurement evaluation and our analyses (e.g., systematic misclassification) have gone even further. Nevertheless, substantial amounts of the error variation between states, items, and respondents is not accounted for, and the conclusions we can reach about arrest item reliabilities are very sensitive to assumptions. It appears that additional exploratory and hypothesis-testing research will be needed to describe the error structure of official records and of the questionnaire responses from convicted felons. With more complete descriptions, the kinds and amounts of errors that are present can be estimated and their effects on subject matter estimates and conclusions can be derived.

ADDITIONAL RELIABILITY ESTIMATION PROCEDURE

Next, we suggest an additional reliability estimation procedure for use after analysts develop models to predict arrest and conviction information. The method produces separate MCAF estimates for questionnaire and record data based on the relative amounts of measured variation that can be explained by a regression model.

Let $R^2(Q, X)$ be the proportion of measured variation in questionnaire item Q that can be explained by a vector of predictor variables, X, in a regression.

Let $R^2(V, X)$ be the proportion of explained variation in the corresponding record variable, V, using the same predictors.

Define $R^2(T, X)$ to be the proportion of true score variation explained by the predictors.

By direct extension of eq. 4 (Section I):

$$R^2(Q, X) = (1+b)^2(\text{var } T/\text{var } Q)R^2(T, X)$$
$$= MCAF_Q^2 R^2(T, X).$$

Similarly,

$$R^2(V, X) = MCAF_V^2 R^2(T, X).$$

Let A denote the ratio of the explained measured variances:

$$A = R^2(Q, X)/R^2(V, X) = MCAF_Q^2/MCAF_V^2.$$

Now,

$$\sqrt{A} = \sqrt{R^2(Q, X)} \;/\; \sqrt{R^2(V, X)}$$
$$= \sqrt{(1+b)^2(\text{var } T/\text{var } Q)} \;/\; \sqrt{(1+b')^2(\text{var } T/\text{var } V)}$$
$$= [(1+b)\sqrt{\text{var } V}] \;/\; [(1+b')\sqrt{\text{var } Q}.$$

Rewrite eq. 5, the correlation of survey and record values, as:

$$r(Q, V) = (1+b)(1+b')\text{var } T/(\sqrt{\text{var } Q}\,\sqrt{\text{var } V}).$$

Then,

$$\sqrt{A}\, r(Q, V) = [(1+b)^2 \text{var } T]\text{var } Q$$
$$= MCAF_Q^2.$$

Since,

$$A = MCAF_Q^2/MCAF_V^2,$$

It follows that,

$$MCAF_V = MCAF_S/\sqrt{A}.$$

These estimates assume that the questionnaire and record true score variances are equal (e.g., no definition differences) and that errors in the predictors are uncorrelated with errors in the dependent variables.[4]

Since the questionnaire contains two versions of each arrest question (Arrest Incidents and Arrests for Crimes Done), the equations over-identify the $MCAF_V$ estimate. The two estimates can be used to test the goodness of the entire set of assumptions underlying the additional reliability estimates. Alternatively, the extra degree of freedom might be used to estimate one of the implied unknowns such as the degree of difference in true score definitions between the record and one of the two arrest items. The logic can be extended to estimate the degree of retest conditioning although this estimate is of little practical importance here.

Use of the additional reliability estimation method depends on the existence of a good predictive model, one that produces an R^2 that is substantially above zero. We assume that analysts will be developing this kind of model in the future.

RECOMMENDATIONS TO ANALYSTS

Our main recommendation to those who analyze the Second Inmate Survey data is to use strategies that produce conclusions robust to the three kinds of measurement errors we have identified in this report: average biases, biases correlated with true scores, and random measurement errors. Three separate strategies are mentioned next. Probably all should be used when estimating relationships and when estimating a central tendency and its "mean squared error."

[4] A "worksheet" that can be filled in to make the estimate is in Appendix C.

Although average biases tend to be small (some exceptions), single item biases are sometimes very large. Our recommendation here involves taking advantage of the finding that item biases are "localized," that is, they do not occur with the same statistically significant size and sign across the three state samples. Specifically, we recommend that analysis models exclude terms that interact "state" with other variables and that the models be fitted separately with the data from each state. Findings that hold up across the states are unlikely to be artifacts of the average measurement biases.

A second measurement problem that affects almost every kind of statistical estimate is the possible existence of bias correlated with true score. It may arise because respondents exaggerate in proportion to the truth or because records become more incomplete as true values increase (or both). This bias is multiplicative (b x T). For estimates of linear relationships, our recommended approach is to turn it into an additive effect that is statistically less harmful; specifically transforming the observed values by taking the natural logarithm ($\log_e$) of each score (after adding a small, positive constant to take care of observed values of zero). If results do not change importantly when transformed scores are used, they are unlikely to arise because of correlated measurement biases. A different approach is to try nonparametric relationship estimation methods in addition to correlations. Nonparametric coefficients are less sensitive to a true multiplicative bias since they reflect only the order information among observations (e.g., relative ranks). Estimated relationships that "hold" under both estimation approaches are probably robust to the presence of correlated measurement biases.

Random measurement error is the third kind of non-true-score variation that may plague analyses. We recommend that single questionnaire items not be used alone in the analyses. Instead, consider combining items, thought to measure the same underlying construct (or "latent variable"), into scales. Simple scales based on addition or averaging will be more reliable[5] than single item measures when the single

[5] More reliable here means only that random error variance is a smaller proportion of the total measured variance.

item variances are approximately equal. Because scale construction is unlikely to remove biases correlated with true score the transformation and/or nonparametric approaches, recommended above, should be used with the scale scores.

A final recommendation addresses both kinds of unreliability (bias correlated with true score and random error variance). If there are both questionnaire and record versions of analysis variables available, estimate their MCAF reliabilities with the R^2 procedures outlined in this section and apply corrections for the estimated attenuation in the reported results. Do not assume that the records are perfect and merely rely on recorded values.

FUTURE SURVEYS OF CRIMINAL OFFENDERS

We close with some practical observations about the design and conduct of future surveys of offender populations.

Official record information is still needed to determine whether respondents furnish anything more than "random noise" when describing their arrests. We have questions about record accuracy so we suggest that additional record features be included in new research:

1. An investigation of the completeness of the arrest records that involves obtaining record data from local jurisdictions as well as from prison files. We are particularly interested in estimating the b' parameter values; that is, finding out if record omissions are correlated with the true number of arrests.

2. Independent double coding of a sample of records (two coders) to assess coder reliability. Errors introduced in coding can be inappropriately interpreted as response or record errors unless they can be estimated and corrections introduced for them.

3. Independent replication of the matching of questionnaires to records for a sample of cases to estimate the match errors. Match errors definitely need to be estimated in future record check studies.

4. Coding <u>all</u> recorded arrest and conviction information instead of just information of interest. This reduces the chance that coders will make incorrect decisions about what not to code, and thereby reduces a possible source of apparent record underreporting.

5. Coding more of the detailed information about arrests and convictions (e.g., statute, date, charge). This will allow analysts to try different ways of classifying events into broader categories before conducting analyses. It will enable methodologists to estimate any time-related memory decay response effects; it will allow a more empirically based resolution of disagreements between questionnaires and coders about Street Month periods; and it will help to resolve some definitional ambiguities concerning the difference between arrest types (e.g., occasion vs. count).

For future questionnaires we recommend investigating strategies that may increase reliability such as creating more redundancy among items and testing the effects of interviewer face-to-face administration.

Our current findings suggest that response error variance creates more of a threat to analytical inferences than average response biases. Well-trained interviewers have the potential for spotting and helping to correct mistakes made by respondents. Interviewers can be trained to execute complex questionnaire logic (skip patterns) much more accurately than respondents, and they can apply a set of uniform interpretations (of question terms, classification rules, time periods) across respondents. Personal interviews cost more than the self-administered questionnaires, and interviewers can introduce different kinds of errors into the data, especially if training is inadequate. We recognize these "trade-offs" between positive and negative benefits so we suggest a feasibility study rather than offering an unqualified recommendation.

Creating response redundancy, the classic method of reducing unreliability in test construction, may apply equally well to surveys. Response redundancy is introduced by asking similar questions that are hypothesized to measure the same underlying trait or characteristic. If the redundant item true-scores covary and the non-true-score components are uncorrelated across items, a combined item index is more reliable than any of the single items.[6]

---

[6]For a comprehensive discussion including how to estimate reliability gains from adding items, consult Lee J. Cronbach, "Coefficient Alpha and the Internal Structure of Tests," <u>Psychometrika</u>, Vol. 16, 1951, pp. 297-334.

If the questionnaire length is fixed, the designer must decide between measuring many different concepts (using single questions with potentially lower reliabilities) and measuring fewer concepts with more questions devoted to each (and presumably more reliable scale scores for each). Multiple item strategies have been developed and used for so-called "soft" concepts such as attitudes, mental abilities, and "permanent" income, but not for counts of well-defined events such as arrests, hospital stays, or pounds of body weight. Applying item redundancy principles to increase the reliability of arrest count and similar responses will not be easy. While we could rely on asking a question twice in exactly the same way (the retest method), we suspect this will be counterproductive because the response errors will be correlated. We must invent items that do not evoke similar response errors but, when combined, retain the desirable ratio scale properties of the straightforward, single item question.

Appendix A

EXCERPTS FROM QUESTIONNAIRE

This appendix contains reproductions of the questions in the survey questionnaire which were applicable to our analyses. Exhibits A.1 and A.2 show the set of questions and the calendar used to establish the Street Months. Exhibit A.3 gives the Current Conviction questions and Exhibit A.4 the Arrest Incidents questions. Exhibit A.5 shows the questions on Arrests for Crimes Done, and Exhibit A.6 the questions selected to generate the Predictor Variables. See Appendix D for the Spanish language version of these questions.

Exhibit A.1

STREET MONTHS QUESTIONS

1. The next section will be about the time before you were arrested for your present term. There is a blue card with a calendar on it. The instructions on this page tell you how to fill it out. Raise your hand if you have any trouble filling it out.

2. For the sentence you are now serving, in what year were you arrested? (If you were arrested several times for this sentence, use the earliest arrest.)

   Year Arrested: _____ *8 ''''/*

   Write that year where it says "Year Arrested" on the calendar.

3. In what month of that year was that arrest?

   Month Arrested: _____ *12 ''/*

   Write "arrested" on the calendar in that month (for the "Year Arrested" line.)

4. Now, draw a line through all the months after that month (to the end of the year).

5. You will not be asked about anything that happened in the months you drew the line through.

6. What was the year before you were arrested?

   Year Before Arrested: _____ *14 ''''/*

   Write that year on the calendar where it says "Year Before Arrested".

7. During all the months on the calendar before you were arrested (including both years) were you ever locked up for a month or more?

   NO ☐1 YES ☐2 ➧ Put X's in all the months when you were locked up. (If you can't remember exactly, think about the time of year it was and put X's in the number of months you were locked up around that time of year.) *18/*

8. Now look at the calendar. All the blank boxes (without X's or lines) are months when you were on the street before you were arrested.

9. Count all the blank boxes.
   How many months was that? ⟶ _____ Months *19 ''/*

10. You will be asked about these months and also about the month you marked "Arrested". To get the total of these months, add one month and write the total here. + 1 .
    ⟶ _____ Total Street Months *21 ''/*

11. Write this total number in the box on the calendar where it says "STREET MONTHS ON THE CALENDAR". You will need this number in answering the next questions.

12. Underneath the month marked "Arrested," write "Include this month."

    This will remind you to include this month in your answers.

23/

Exhibit A.2

CALENDAR FOR CALCULATING STREET MONTHS

INSTRUCTIONS FOR USING THIS CALENDAR ARE INCLUDED IN THE SURVEY.

	Winter			Spring			Summer			Fall		
	January	February	March	April	May	June	July	August	September	October	November	December
1977 YEAR BEFORE ARRESTED			X	X	X							
1978 YEAR ARRESTED	January	February	March	April	May	June	July	August	September	October	November	December
						X include this month						

STREET MONTHS ON THE CALENDAR ➧ 15

Exhibit A.3

CURRENT CONVICTION QUESTION

6. These questions are only about the crime(s) for which you are now serving a sentence. What charge(s) were you convicted of that you are serving time for now? *(Check all that apply.)*

☐ 1	Assault/ADW	43/
☐	Auto Theft/Vehicle Theft	44/
☐	Burglary	45/
☐	Drug Possession	46/
☐	Drug sales	47/
☐	Forgery/Bad check/Bad credit card	48/
☐	Fraud or Swindle	49/
☐	Kidnapping	50/
☐	Murder/Manslaughter	51/
☐	Possession or receiving stolen property	52/
☐	Rape	53/
☐	Robbery	54/
☐	Sex offense (other than rape)	55/
☐	Theft/Grand theft/Larceny	56/
☐	Weapons charge	57/
☐	Other, what? _____	58/

Exhibit A.4

ARREST INCIDENTS QUESTIONS

4. Look at the calendar. During the STREET MONTHS ON THE CALENDAR, how many times were you arrested for each of the following crimes? Count an arrest even if you did not actually do the crime you were arrested for. *(Check* NONE *if not arrested for that crime.)*

BURGLARY	_____	arrests	(OR)	NONE ☐ CO	12 ''/
ROBBERY OR ARMED ROBBERY	_____	arrests	(OR)	NONE ☐	14 ''/
ASSAULT, AGGRAVATED ASSAULT OR ASSAULT WITH A DEADLY WEAPON	_____	arrests	(OR)	NONE ☐	16 ''/
MURDER OR MANSLAUGHTER	_____	arrests	(OR)	NONE ☐	18 ''/
AUTO THEFT, MOTOR VEHICLE THEFT	_____	arrests	(OR)	NONE ☐	20 ''/
THEFT, GRAND THEFT, LARCENY OR GRAND LARCENY	_____	arrests	(OR)	NONE ☐	22 ''/
FORGERY, USE OF A STOLEN OR BAD CREDIT CARD OR BAD CHECK PASSING	_____	arrests	(OR)	NONE ☐	24 ''/
FRAUD	_____	arrests	(OR)	NONE ☐	26 ''/
SELLING DRUGS, POSSESSING DRUGS FOR SALE, OR TRANSPORTING DRUGS	_____	arrests	(OR)	NONE ☐	28 ''/

Exhibit A.5

ARRESTS FOR CRIMES DONE QUESTIONS

Burglary

1. During the STREET MONTHS ON THE CALENDAR did you do any burglaries? (Count any time that you broke into a house or a car or a business in order to take something.)

YES ☐ 1 NO ☐ 2 ➧ go on to page 18

5. How many of these burglaries were you arrested for? (Include all of the times you were arrested for doing a burglary even if you were charged with something else.)

_____ Arrests for burglaries

Business Robbery

1. During the STREET MONTHS ON THE CALENDAR did you rob any businesses? That is did you hold up a store, gas station, bank, taxi or other business?

YES ☐ 1 NO ☐ 2 ➧ go on to page 20

5. How many of these robberies were you arrested for? (Include all of the times you were arrested for robbing a business even if you were charged with something else.)

_____ Arrests for business robberies

Personal Robbery

1. During the STREET MONTHS ON THE CALENDAR did you rob any persons, do any muggings, street robberies, purse snatches, or hold-ups in someone's house or car? (Do not include any business robberies or hold-ups during a burglary that you already mentioned.)

YES ☐ 1 NO ☐ 2 ➧ go on to page 22

5. How many of these robberies were you arrested for? (Include all of the times you were arrested for robbing a person even if you were charged with something else.)

_____ Arrests for robbing people

Assault

1. Even if no one was hurt, during the STREET MONTHS ON THE CALENDAR did you assault someone, threaten someone with a weapon, shoot at someone, try to cut someone, or beat or strangle someone?

YES ☐ 1 NO ☐ 2 ➡ go on to page 26

4. How many times were you arrested when you assaulted, threatened, shot at, tried to cut, or beat or strangled someone?

_____ Arrests

Theft

1. During the STREET MONTHS ON THE CALENDAR did you do any theft or boosting? That is, did you steal from a till or cash register, shop lift, or pick pockets, or take something from someone without their knowledge? (Do not include car theft.)

YES ☐ 1 NO ☐ 2 ➡ go on to page 28

5. How many of these thefts were you arrested for? (Include all of the times you were arrested for doing a theft even if you were charged with something else.)

_____ Arrests for Thefts

Auto Theft

1. During the STREET MONTHS ON THE CALENDAR did you steal any cars, trucks or motorcycles?

YES ☐ 1 NO ☐ 2 ➡ go on to page 30

5. How many of these vehicle thefts were you arrested for? (Include all of the times you were arrested for stealing a vehicle, even if you were charged with something else.)

_____ Arrests for vehicle thefts

Forgery, Card

1. During the STREET MONTHS ON THE CALENDAR did you ever forge something, use a stolen or bad credit card, or pass a bad check?

YES ☐ 1 NO ☐ 2 ➡ go on to page 32

5. How many of these forgeries, bad checks or credit cards were you arrested for? (Include all of the times you were arrested for doing one of these things even if you were charged with something else.)

_____ Arrests

Fraud

1. During the STREET MONTHS ON THE CALENDAR did you do any frauds or swindles (illegal cons) of a person, business, or the government?

YES ☐ 1 NO ☐ 2 ➡ go on to page 34

5. How many of these frauds or swindles were you arrested for? (Include all of the times you were arrested for doing a fraud or swindle even if you were charged with something else.)

_____ Arrests for frauds or swindles

Drugs

1. During the STREET MONTHS ON THE CALENDAR did you ever deal in drugs? That is, did you make, sell, smuggle or move drugs?

YES ☐ 1 NO ☐ 2 ➡ go on to page 36

5. How many of these drug deals were you arrested for?

_____ Arrests for drugs

Exhibit A.6

SOURCE QUESTIONS FOR SELECTED PREDICTOR VARIABLES

X_1 Education

What is the highest grade you finished in school?

- ☐ 0 No schooling
- ☐ 1 6th grade or less
- ☐ 2 7th - 9th grade
- ☐ 3 10th - 11th grade
- ☐ 4 High school graduate
- ☐ 5 Some college
- ☐ 6 College graduate
- ☐ 7 Post graduate study

X_6 Time Since Arrested

For the sentence you are now serving, in what year were you arrested? (If you were arrested several times for this sentence, use the earliest arrest.)

Year Arrested: _____

X_7 Number of Felony Convictions

Altogether in your life, how many times have you been convicted of a felony?

- ☐ Never
- ☐ Once
- ☐ 2-3 times
- ☐ 4-6 times
- ☐ 7-10 times
- ☐ 11-15 times
- ☐ 16 or more times

X_{10} Age

How old were you on your last birthday?

_____ Years old

X_{11} Race

What is your race?

- ☐ 1 Asian
- ☐ 2 Black
- ☐ 3 Chicano/Latino
- ☐ 4 Indian/Native American
- ☐ 5 White
- ☐ 6 Other

Appendix B

EXCERPTS FROM RECORD CODING FORMS

Appendix B contains reproductions of the sections of the Official Records Coding Forms for each state which were used in these analyses. The first section shows the Current Conviction items and the second shows the Arrest items for each state.

Exhibit B.1

CALIFORNIA CURRENT CONVICTION CODING FORM*

OFFENSE TYPE	CAL. STATUTES	(1 = Yes)	
ASSAULT (664 = Attempt)	203-204, 216-222, 240-246 273a, 4500-4501.5	[ ]	(41)
AUTO THEFT	487, 489, 499b, 499d (V.C.) 10851	[ ]	(42)
BURGLARY (664 = Attempt)	459-464	[ ]	(43)
DRUG POSSESSION	(H & S) 11350, 57-58, 77	[ ]	(44)
DRUG SALES, ETC.	(H & S) 11351-4, 59-61, 63, 11378-80, 82-83	[ ]	(45)
FORGERY/NSF/CREDIT CARD	470-483, 484d-484i	[ ]	(47)
FRAUD	72, 424, 425 (IC) 556 et. seq.	[ ]	(48)
HOMICIDE	187 - 190, 192 - 193	[ ]	(50)
KIDNAP	207-210	[ ]	(51)
POSSESSION/RECEIVING STOLEN PROPERTY	496	[ ]	(52)
RAPE (664 = Attempt)	261, 264, 264.1, 264.5	[ ]	(53)
ROBBERY (664 = Attempt)	211-214	[ ]	(54)
SEX OFFENSE (NOT RAPE)	261.5, 266, 266a-i, 267, 273g, 281, 285, 288, 288.1 288a, 314, 647a	[ ]	(55)
THEFT/GRAND THEFT/LARCENY	484, 484b, c, 485-487 487a-487g, 488-490, 498-499a	[ ]	(56)
WEAPONS	4502, 4574, 12020-12034	[ ]	(57)

*Reproduction reduced 25% from original

Exhibit B.2

MICHIGAN CURRENT CONVICTION CODING FORM

OFFENSE TYPE	STATUTE #	TOTALS	
1. ASSAULT	750.81(a) - 750.89 750.91 - 750.92 750.397 - 750.397(a)	[ ][ ]	(58)
2. AUTO THEFT	750.413	[ ][ ]	(59)
3. BURGLARY	750.110 - 750.116	[ ][ ]	(60)
4. DRUG POSSESSION	335.153 - 335.154 335.61 335.67	[ ][ ]	(61)
5. DRUG SALE, ETC.	335.02 335.10 335.19 335.101 335.151 - 335.153	[ ][ ]	(62)
6. FORGERY/NSF/CREDIT CARD	750.131 750.248 - 750.266	[ ][ ]	(63)
7. FRAUD	750.271 - 750.300	[ ][ ]	(64)
8. KIDNAPPING	750.349 - 750.350	[ ][ ]	(65)
9. MURDER/MANSLAUGHTER	750.236 750.316 - 750.317 750.319 - 750.321 750.324 750.861	[ ][ ]	(66)
10. POSSESSION/RECEIVING STOLEN PROPERTY	750.335	[ ][ ]	(67)
11. RAPE	750.520	[ ][ ]	(68)
12. ROBBERY	750.529 - 750.531	[ ][ ]	(69)
13. SEX OFFENSE (NOT RAPE)	750.158 - 750.159	[ ][ ]	(70)
14. THEFT/GRAND THEFT/LARCENCY	750.356 - 750.367(b)	[ ][ ]	(71)
15. WEAPONS	750.223 - 750.235 750.237 752.862 752.863(a) 752.881 - 752.883	[ ][ ]	(72)

Exhibit B.3

CALIFORNIA ARRESTS CODING FORM

OFFENSE TYPE	CAL STATUTES	TALLY	TOTAL	
ASSAULT (664 = Attempt)	203-204, 216-222, 240-246, 273a, 4500-4501.5	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(41-42)
AUTO THEFT	487, 489, 499b, 499d, (VC) 10851	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(43-44)
BURGLARY (664 = Attempt)	459-464	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(45-46)
DRUG SALES, ETC.	(H & S) 11351-4, 59-61, 63, 11378-80, 82-83	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(47-48)
FORGERY/NSF/CREDIT CARD	470-483, 484d-484i	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(49-50)
FRAUD	72, 424, 425 (IC) 556 et. seq.	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(51-52)
HOMICIDE	187-190, 192-193	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(53-54)
ROBBERY (664 = Attempt)	211-214	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(61-62)
THEFT/GRAND THEFT/ LARCENY (664 = Attempt)	484, 484b, 484a 485-487, 487a-g, 488-490, 498-499a	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(63-64)

Exhibit B.4

MICHIGAN ARRESTS CODING FORM

50. ARREST TYPE	STATUTE #	TALLY	TOTALS	
1. ASSAULT Aggrevated asslt; Asslt, w/a deadly weapon; Felonious asslt., Asslt. to commit rape; robbery; and other felonies; Attempted homicide; Mayhem; Asslt. and battery; Battery.	750.81(a) - 750.89 750.91 - 750.92 750.397 - 750.397(a)	1 2 3 4 5 6 7 8 9 10	[ ][ ]	(38-39)
2. AUTO THEFT Unlawful use, taking, possession of motor vehicle.	750.413	1 2 3 4 5 6 7 8 9 10	[ ][ ]	(40-41)
3. BURGLARY Burglary; Possession of burglary tools - explosives; Breaking and entering; Entering w/o breaking.	750.110 - 750.116	1 2 3 4 5 6 7 8 9 10	[ ][ ]	(42-43)
4. DRUGS-SALE,ETC. Narcotics, Non-narcotics, Hallucinogens, Marihuana - Unlawful sale, distribution, and manufacturing; Distribution to minors.	335.02 335.10 335.19 335.101 335.151 - 335.153	1 2 3 4 5 6 7 8 9 10	[ ][ ]	(44-45)
5. FORGERY Forgery; Uttering & Publishing; checks with non-sufficient funds (NSF); Counterfeiting; Possession of counterfeit notes, tools, coins; Stolen credit card, use or possession.	750.131 750.248 - 750.266	1 2 3 4 5 6 7 8 9 10	[ ][ ]	(46-47)
6. FRAUD Fraudulent obtaining; selling; using; Gross fraud; Defrauding.	750.271 - 750.300	1 2 3 4 5 6 7 8 9 10	[ ][ ]	(48-49)
7. HOMICIDE Murder 1; Murder 2; Manslaughter; Negligent Manslaughter; Negligent Homicide (Weapon, Auto).	750.236 750.316 - 750.317 750.319 - 750.321 750.324 750.861	1 2 3 4 5 6 7 8 9 10	[ ][ ]	(50-51)
8. LARCENY Larceny from person; building; vehicle; other property; By false personation; Stealing; Conversion; Embezzlement; Grand theft; Grand larceny; Petty theft; Petty larceny.	750.356 - 750.367	1 2 3 4 5 6 7 8 9 10	[ ][ ]	(52-53)
9. ROBBERY Robbery, armed; Unarmed; Bank safe or vault robbery; Purse snatching	750.529 - 750.531	1 2 3 4 5 6 7 8 9 10	[ ][ ]	(54-55)

Exhibit B.5

TEXAS ARRESTS CODING FORM

17. OFFENSE TYPE	TEXAS STATUTES	TALLY	TOTAL	
01. ARSON	28.01 et.seq.	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(42-43)
02. ASSAULT	22.00 et.seq.	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(44-45)
03. AUTO THEFT	31.07	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(46-47)
04. BURGLARY	30.01 - 30.04	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(48-49)
05. DRUG SALES, ETC.	(C.S. 4476-15) 4.04-4.05, 4.07	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(50-51)
06. FORGERY/NSF/CREDIT CARD	32.21 - 32.48	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(52-53)
07. FRAUD	32.01 - 32.02	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(54-55)
08. HOMICIDE	19.01 et.seq.	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(56-57)
12. ROBBERY	29.01 et.seq.	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(64-65)
13. THEFT/GRAND THEFT/LARCENY	31.01 - 31.06	01 02 03 04 05 06 07 08 09 10 11 12 13 14 15	[ ][ ]	(66-67)

Appendix C

TECHNICAL TABLES

Table C.1

CHARACTERISTICS OF SAMPLES BY STATE

Variable	California N	California Value	Michigan N	Michigan Value	Texas N	Texas Value
Age						
mean years	354	28	427	26	662	27
Race						
% Black	123	35	270	64	345	52
% Chicano	63	18	12	3	65	10
% White	146	42	129	30	241	36
% Other	17	5	13	3	13	2
Education						
% completed 9th Grade or Less	58	16	108	25	202	30
% completed 10th-11th Grade	120	34	142	34	249	38
% completed High School	84	24	60	14	73	11
% completed over High School	92	26	114	27	138	21
Felonies Ever Convicted of						
mean number	352	3	428	3	654	3

Table C.2

ITEM BIAS SCORES FOR CURRENT CONVICTION ITEMS
CALIFORNIA PRISON SAMPLE

Conviction Category	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Assault	.20	.13	.07*	.02
Auto Theft	.07	.05	.02	.01
Burglary	.25	.27	-.02	.02
Drugs (Possession & Sale)	.15	.11	.03*	.01
Forgery, Card & Fraud	.05	.03	.02*	.01
Kidnap	.04	.03	.01	.01
Murder	.12	.10	.01	.01
Stolen Property	.08	.08	.00	.01
Rape	.05	.06	-.01	.01
Robbery	.40	.40	.00	.02
Other Sex	.02	.02	-.00	.01
Theft	.09	.06	.03*	.01
Weapons	.24	.32	-.08*	.02

NOTES:
N = 329

Symbols:
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.3

ITEM BIAS SCORES FOR CURRENT CONVICTION ITEMS
MICHIGAN PRISON SAMPLE

Conviction Category	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Assault	.13	.19	-.06*	.02
Auto Theft	.06	.03	.04*	.01
Burglary	.17	.20	-.04*	.02
Drugs (possession & sale)	.10	.10	-.01	.01
Forgery, Card & Fraud	.04	.05	-.01	.00
Kidnap	.02	.02	.00	.00
Murder	.10	.09	.01	.01
Stolen Property	.04	.06	-.03*	.01
Rape	.06	.10	-.03*	.01
Robbery	.33	.25	.08*	.02
Other Sex	.06	.01	.05*	.01
Theft	.08	.15	-.07*	.02
Weapons	.12	.15	-.03*	.02

NOTES:
N = 336

Symbols:
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.4

ITEM BIAS SCORES FOR CURRENT CONVICTION ITEMS
TEXAS PRISON SAMPLE

Conviction Category	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Assault	.10	.06	.04*	.01
Auto Theft	.09	.05	.04*	.01
Burglary	.39	.45	-.06*	.01
Drug (Possession & Sale)	.16	.13	.02*	.01
Forgery & Fraud	.09	.06	.03*	.01
Kidnap	.01	.00	.01*	.01
Murder	.04	.06	-.02*	.01
Stolen Property	.02	.00	.02*	.01
Rape	.04	.04	-.00	.00
Robbery	.22	.22	-.00	.01
Other Sex	.04	.03	.01	.01
Theft	.09	.13	-.04*	.01
Weapons	.04	.02	.02*	.01

NOTES:
N = 632

Symbols:
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C. 5

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR CURRENT CONVICTION ITEMS
CALIFORNIA PRISON SAMPLE

Conviction Category	Record Method		Retest Method
	Lower Bound*	Upper Bound**	$\sqrt{r}$ ***
Assault	.68	.82	.83
Auto Theft	.64	.80	.69
Burglary	.75	.86	.97
Drugs (Possession & Sale)	.72	.85	.90
Forgery, Card & Fraud	.66	.81	.84
Kidnap	.66	.81	1.00
Murder	.91	.95	1.00
Stolen Property	.55	.74	.70
Rape	.83	.91	1.00
Robbery	.85	.92	.96
Sex not Rape	.66	.81	--
Theft	.54	.73	.90
Weapons	.54	.74	.84

NOTES:
Survey, Record N = 329
Survey, Retest N = 53
*Correlation of the Survey & Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
--No Variation Observed

Table C. 6

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR CURRENT CONVICTION ITEMS
MICHIGAN PRISON SAMPLE

Conviction Category	Record Method		Retest Method
	Lower Bound*	Upper Bound**	$\sqrt{r}$ ***
Assault	.58	.76	1.00
Auto Theft	.48	.69	.89
Burglary	.73	.85	.94
Drugs (Possession & Sale)	.84	.91	--
Forgery, Card & Fraud	.93	.97	.90
Kidnap	.92	.96	.84
Murder	.92	.96	.80
Stolen Property	.55	.74	1.00
Rape	.59	.77	--
Robbery	.67	.82	.95
Sex not Rape	.14	.38	--
Theft	.55	.74	.92
Weapons	.68	.83	.83

NOTES:
Survey, Record N = 336
Survey Retest N = 42
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
--No Variation Observed

Table C.7

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR CURRENT CONVICTION ITEMS
TEXAS PRISON SAMPLE

Conviction Category	Record Method		Retest Method
	Lower Bound*	Upper Bound**	$\sqrt{r}$ ***
Assault	.68	.82	.98
Auto Theft	.50	.71	.97
Burglary	.81	.90	1.00
Drugs (Possession & Sale)	.82	.91	.97
Forgery, Card & Fraud	.76	.87	.96
Kidnap	.57	.76	.90
Murder	.75	.87	1.00
Stolen Property	.25	.50	.84
Rape	.73	.85	1.00
Robbery	.92	.96	.98
Sex not Rape	.63	.79	.93
Theft	.51	.71	.88
Weapons	.54	.74	.62

NOTES:
Survey, Record N = 632
Survey, Retest N = 122
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest

Table C.8A

REGRESSION COEFFICIENTS (β) FOR PREDICTOR VARIABLES
ON RESPONDENT BIAS SCORE FOR CURRENT
CONVICTION BY STATE

Variable	California β	California s.e.	Michigan β	Michigan s.e.	Texas β	Texas s.e.
Education (X_1)	.004	(.004)	-.006 *	(.003)	-.002	(.002)
Personal Administration (X_2)	.048	(.026)	-.019	(.025)	-.014	(.009)
Missed Current Conviction Items (X_3)	(Not included in these equations)					
Missed Arrest Incident Items (X_4)	.002	(.002)	-.002	(.001)	.0004	(.001)
Missed Arrest Done Items (X_5)	.008	(.005)	.008	(.005)	.002	(.004)
Time Since Arrest (X_6)	-.0001	(.0003)	-.0001	(.0001)	.0000	(.0001)
Number of Felony Convictions (X_7)	.004	(.004)	-.0005	(.004)	.001	(.003)
Record Street Months Later (X_8)	.041 *	(.013)	.012	(.010)	.017	(.010)
Survey Street Months Later (X_9)	-.009	(.033)	-.019	(.016)	-.011	(.013)
Age (X_{10})	-.001	(.001)	.0000	(.0005)	.0001	(.0003)
Race (X_{11})	.024 *	(.009)	.0006	(.007)	.006	(.006)
Constant	.093		.147		-.034	
Adjusted R^2	.062		.024		-.001	
F	3.059 *		1.784		.959	

NOTES:
California N= 312
Michigan N= 323
Texas N= 533

* $p \leq .01$

Ability Variables
X_1 = Education (reported in survey by respondent) (highest grade completed) 0 = No schooling, 1 = 6th grade or less, 2 = 7th-9th grade, 3 = 10th-11th grade, 4 = high school graduate, 5 = some college, 6 = college graduate, 7 = post graduate study
X_2 = Respondent required assistance with reading questionnaire; 0 = No, 1 = Yes
X_3 = Number of skipped items for Current Conviction Section
X_4 = Number of skipped items for Arrest Incident Section
X_5 = Number of skipped items for Arrests for Crimes Done Section
Memory Variables
X_6 = Elapsed time since end of street months. This variable was built using the year arrested response multiplied by 12 and added to the number of the month arrested. (Months coded January = 01....December = 12)
X_7 = Number of felony convictions in lifetime (reported in survey by respondent) 1 = Never, 2 = once, 3 = 2-3 times, 4 = 4-6 times, 5 = 7-10 times, 6 = 11-15 times, 7 = 16 or more times (Indicator of Interference)
Record Match Problem Potential
X_8 = Recorded street months ended later than survey version. 0 = No, 1 = Yes
X_9 = Survey reported street months ended later than record version. 0 = No, 1 = Yes
Demographic Variable
X_{10} = Respondent age (reported in survey by respondent)
X_{11} = Respondent race (dummy variable) 0 = nonwhite, 1 = white

Table C.8B

MEANS, STANDARD DEVIATIONS, AND HIGHEST PREDICTOR INTERCORRELATION FOR PREDICTOR VARIABLES ON CURRENT CONVICTION ITEMS RESPONDENT BIAS AND TOTAL ERROR SCORES BY STATE (SEE TABLE C.8A)

Variable	California			Michigan			Texas		
	mean	standard deviation	h.p.i.	mean	standard deviation	h.p.i.	mean	standard deviation	h.p.i.
Education (X_1)	3.59	1.18	X_2-.17	3.43	1.29	X_5-.21	3.21	1.31	X_2-.32
Personal Administration (X_2)	.03	.18	X_1-.17	.02	.14	X_7.16	.12	.33	X_1-.32
Missed Current Conviction Items (X_3)	(not included in these equations)								
Missed Arrest Incident Items (X_4)	.96	2.36	X_7.16	.96	2.46	X_{10}.14	1.03	2.57	X_5.21
Missed Arrest Done Items (X_5)	.36	.94	X_4.10	.22	.66	X_1-.21	.24	.75	X_4.21
Time Since Arrest (X_6)	924.35	14.86	X_8-.34	915.48	25.08	X_{10}-.27	923.20	22.78	X_{10}-.29
Number of Felony Convictions (X_7)	3.32	1.17	X_{11}.21	3.07	.96	X_{11}.17	3.04	1.04	X_{10}.27
Record Street Months Later (X_8)	.16	.37	X_6-.34	.15	.35	X_5.18	.09	.29	X_6-.16
Survey Street Months Later (X_9)	.02	.14	X_2.11	.05	.21	X_6.12	.05	.22	X_2.12
Age (X_{10})	27.36	6.35	X_7.12	26.09	7.13	X_6-.27	27.38	9.49	X_6-.29
Race (X_{11})	.42	.49	X_7.21	.32	.47	X_7.17	.36	.48	X_1.19

NOTES:

California N = 312
Michigan N = 323
Texas N = 533

SYMBOLS:

h.p.i. = highest predictor intercorrelation
** = coefficient could not be computed

Table C.9A

REGRESSION COEFFICIENTS (β) FOR PREDICTOR VARIABLES ON TOTAL ERROR SCORE FOR CURRENT CONVICTION BY STATE

Variable	California β	California s.e.	Michigan β	Michigan s.e.	Texas β	Texas s.e.
Education (X_1)	.0004	(.004)	.004	(.003)	-.003	(.003)
Personal Administration (X_2)	.059 *	(.024)	.008	(.031)	-.005	(.011)
Missed Current Conviction Items (X_3)	(Not included in these Equations)					
Missed Arrest Incident Items (X_4)	.002	(.002)	.001	(.002)	-.001	(.001)
Missed Arrest Done Items (X_5)	.006	(.004)	.015 *	(.007)	-.004	(.005)
Time Since Arrest (X_6)	-.0002	(.0003)	-.0001	(.0002)	-.0001	(.0002)
Number of Felony Convictions (X_7)	.003	(.004)	-.001	(.005)	.001	(.003)
Record Street Months Later (X_8)	.048 *	(.012)	.016	(.012)	.026 *	(.012)
Survey Street Months Later (X_9)	.032	(.031)	.015	(.020)	-.001	(.016)
Age (X_{10})	-.001	(.001)	-.0004	(.001)	-.0002	(.0004)
Race (X_{11})	-.003	(.009)	-.005	(.009)	-.001	(.007)
Constant	.234		.186		.177	
Adjusted R^2	.071		.002		-.004	
F	3.369 *		1.071		.797	

NOTES:
California N= 312
Michigan N= 323
Texas N= 533

* $p \leq .01$

Ability Variables

X_1 = Education (reported in survey by respondent) (highest grade completed) 0 = No schooling, 1 = 6th grade or less, 2 = 7th-9th grade, 3 = 10th-11th grade, 4 = high school graduate, 5 = some college, 6 = college graduate, 7 = post graduate study

X_2 = Respondent required assistance with reading questionnaire; 0 = No, 1 = Yes

X_3 = Number of skipped items for Current Conviction Section

X_4 = Number of skipped items for Arrest Incident Section

X_5 = Number of skipped items for Arrests for Crimes Done Section

Memory Variables

X_6 = Elapsed time since end of street months. This variable was built using the year arrested response multiplied by 12 and added to the number of the month arrested. (Months coded January = 01....December = 12)

X_7 = Number of felony convictions in lifetime (reported in survey by respondent) 1 = Never, 2 = once, 3 = 2-3 times, 4 = 4-6 times, 5 = 7-10 times, 6 = 11-15 times, 7 = 16 or more times (Indicator of Interference)

Record Match Problem Potential

X_8 = Recorded street months ended later than survey version. 0 = No, 1 = Yes

X_9 = Survey reported street months ended later than record version. 0 = No, 1 = Yes

Demographic Variable

X_{10} = Respondent age (reported in survey by respondent)

X_{11} = Respondent race (dummy variable) 0 = nonwhite, 1 = white

Table C.9B

MEANS AND STANDARD DEVIATIONS FOR PREDICTOR VARIABLES
ON CURRENT CONVICTION ITEMS
TOTAL ERROR SCORE BY STATE (SEE TABLE 9A)

Variable	California		Michigan		Texas	
	mean	standard deviation	mean	standard deviation	mean	standard deviation
Education (X_1)	3.59	1.18	3.43	1.29	3.21	1.31
Personal Administration (X_2)	.03	.18	.02	.14	.12	.33
Missed Current Conviction Items (X_3)	0.00	0.00	0.00	0.00	0.00	0.00
Missed Arrest Incident Items (X_4)	.96	2.36	.96	2.46	1.03	2.57
Missed Arrest Done Items (X_5)	.36	.94	.22	.66	.24	.75
Time Since Arrest (X_6)	924.35	14.86	915.48	25.08	923.20	22.78
Number of Felony Convictions (X_7)	3.32	1.17	3.07	.96	3.04	1.04
Record Street Months Later (X_8)	.16	.37	.15	.35	.09	.29
Survey Street Months Later (X_9)	.02	.14	.05	.21	.05	.22
Age (X_{10})	27.36	6.35	26.09	7.13	27.38	9.49
Race (X_{11})	.42	.49	.32	.47	.36	.48

NOTES:
California N = 312
Michigan N = 323
Texas N = 533

Table C.10

ITEM BIAS SCORES FOR YES-NO ARREST INCIDENT ITEMS
CALIFORNIA PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	320	.34	.44	-.09*	.03
Robbery	317	.42	.50	-.08*	.02
Assault	310	.32	.36	-.04	.03
Murder	310	.13	.13	.00	.02
Auto Theft	309	.18	.19	-.01	.02
Theft	307	.20	.20	-.01	.03
Forgery	307	.16	.10	.06*	.02
Fraud	301	.03	.00	.03*	.01
Drugs	307	.20	.11	.09*	.02

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.11

ITEM BIAS SCORES FOR YES-NO ARREST INCIDENT ITEMS
MICHIGAN PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	322	.28	.24	.04	.03
Robbery	326	.32	.31	.02	.03
Assault	318	.26	.22	.04	.03
Murder	320	.12	.11	.01	.02
Auto Theft	310	.17	.08	.09*	.02
Theft	312	.22	.20	.02	.03
Forgery	312	.06	.07	-.01	.01
Fraud	308	.01	.01	.00	.01
Drugs	319	.19	12	.07*	.02

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.12

ITEM BIAS SCORES FOR YES-NO ARREST INCIDENT ITEMS
TEXAS PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	536	.40	.49	-.10*	.02
Robbery	516	.19	.25	-.06*	.02
Assault	514	.19	.11	.08*	.02
Murder	508	.04	.05	-.01	.01
Auto Theft	511	.15	.08	.07*	.02
Theft	508	.13	.21	-.08*	.02
Forgery	513	.11	.09	.02	.01
Fraud	500	.02	.02	-.00	.01
Drugs	513	.16	.14	.03	.02

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.13

ITEM BIAS SCORES FOR CONTINUOUS ARREST INCIDENT ITEMS
CALIFORNIA PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	316	.81	.93	-.13	.14
Robbery	316	.62	.95	-.32*	.09
Assault	307	.44	.54	-.11	.06
Murder	307	.12	.12	.00	.02
Auto Theft	305	.27	.25	.03	.05
Theft	303	.29	.26	.03	.06
Forgery	305	.26	.24	.02	.05
Fraud	300	.05	.00	.05*	.02
Drugs	305	.35	.28	.07	.08

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.14

ITEM BIAS SCORES FOR CONTINUOUS ARREST INCIDENT ITEMS
MICHIGAN PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	317	.56	.33	.22*	.07
Robbery	324	.62	.40	.22	.13
Assault	315	.36	.28	.08	.04
Murder	315	.12	.10	.02	.02
Auto Theft	307	.30	.13	.17*	.05
Theft	307	.37	.26	.11*	.05
Forgery	311	.09	.14	-.05	.03
Fraud	307	.01	.01	.00	.01
Drugs	316	.30	.17	.13*	.05

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.15

ITEM BIAS SCORES FOR CONTINUOUS ARREST INCIDENT ITEMS
TEXAS PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	530	.93	.73	.19	.11
Robbery	513	.32	.33	-.01	.06
Assault	512	.25	.14	.12*	.03
Murder	507	.04	.05	-.01	.01
Auto Theft	507	.21	.08	.13*	.03
Theft	507	.24	.33	-.10*	.04
Forgery	512	.14	.15	-.01	.02
Fraud	499	.03	.03	.00	.02
Drugs	512	.24	.19	.05	.03

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.16

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR YES-NO ARREST INCIDENT ITEMS
CALIFORNIA PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	320	.48	.69	49	.84
Robbery	317	.62	.79	51	.87
Assault	310	.45	.67	51	.82
Murder	310	.58	.76	49	.91
Auto Theft	309	.49	.70	49	.84
Theft	307	.14	.38	47	.78
Forgery	307	.57	.76	48	.94
Fraud	301	--	--	46	--
Drugs	307	.28	.53	50	.71

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
--No Variation Observed

Table C.17

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR YES-NO ARREST INCIDENT ITEMS
MICHIGAN PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	322	.39	.63	41	.81
Robbery	326	.42	.65	43	.89
Assault	318	.28	.53	40	.67
Murder	320	.48	.69	40	.85
Auto Theft	310	.27	.52	39	.78
Theft	312	.26	.51	38	.65
Forgery	312	.52	.72	39	1.00
Fraud	308	.00†	.00	38	--
Drugs	319	.58	.76	38	.87

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
†Fraud Correlation of -.0080 Was Truncated to .00
--No Variation Observed

Table C.18

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR YES-NO ARREST INCIDENT ITEMS
TEXAS PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	536	.54	.74	119	.88
Robbery	516	.60	.77	111	.91
Assault	514	.36	.60	106	.85
Murder	508	.62	.79	104	--
Auto Theft	511	.35	.59	107	.68
Theft	508	.31	.56	108	.84
Forgery	513	.53	.73	108	.90
Fraud	500	.08	.28	101	.10
Drugs	513	.49	.70	108	.90

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of Survey and Record
***Square Root of the Correlation of the Survey and Retest
--No Variation Observed

Table C.19

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR CONTINUOUS ARREST INCIDENT ITEMS
CALIFORNIA PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	316	.27	.52	49	.29
Robbery	316	.37	.61	51	.87
Assault	307	.31	.56	50	.79
Murder	307	.59	.77	49	.91
Auto Theft	305	.33	.58	48	.68
Theft	303	.14	.38	46	.77
Forgery	305	.57	.76	47	.91
Fraud	300	--	--	46	--
Drugs	305	.23	.48	48	.76

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
--No Variation Observed

Table C.20

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR CONTINUOUS ARREST INCIDENT ITEMS
MICHIGAN PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	317	.42	.65	39	.98
Robbery	324	.08	.29	42	1.00
Assault	315	.32	.56	40	.88
Murder	315	.45	.67	38	.32
Auto Theft	307	.38	.62	38	.53
Theft	307	.43	.66	36	.52
Forgery	311	.77	.88	38	.99
Fraud	307	.00†	.00	38	--
Drugs	316	.32	.57	38	.87

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
†Fraud Correlation of -.0062 Was Truncated to .00
--No Variation Observed

Table C.21

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR CONTINUOUS ARREST INCIDENT ITEMS
TEXAS PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	530	.35	.59	118	.73
Robbery	513	.26	.51	109	.71
Assault	512	.28	.53	104	.88
Murder	507	.64	.80	104	--
Auto Theft	507	.41	.64	106	.72
Theft	507	.32	.57	108	.75
Forgery	512	.54	.74	108	.87
Fraud	499	.12	.34	101	.10
Drugs	512	.38	.62	107	.87

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
--No Variation Observed

Table C.22A

REGRESSION COEFFICIENTS (β) FOR PREDICTOR VARIABLES ON RESPONDENT
BIAS SCORE FOR ARREST INCIDENTS BY STATE

Variable	California β	California s.e.	Michigan β	Michigan s.e.	Texas β	Texas s.e.
Education (X_1)	.030	(.022)	-.020	(.022)	-.018	(.012)
Personal Administration (X_2)	.191	(.144)	.053	(.209)	.067	(.048)
Missed Current Conviction Items (X_3)	-.0001	(.009)	-.042 *	(.009)	.007	(.006)
Missed Arrest Incident Items (X_4)	.002	(.012)	.053 *	(.014)	.035 *	(.007)
Missed Arrest Done Items (X_5)	.050	(.029)	.032	(.044)	.019	(.022)
Time Since Arrest (X_6)	.001	(.002)	-.001	(.001)	.0000	(.001)
Number of Felony Convictions (X_7)	.021	(.023)	.053	(.030)	.038 *	(.015)
Record Street Months Later (X_8)	.047	(.071)	-.085	(.083)	-.055	(.052)
Survey Street Months Later (X_9)	-.250	(.162)	.003	(.129)	.075	(.069)
Age (X_{10})	-.004	(.004)	-.009 *	(.004)	-.003	(.002)
Race (X_{11})	-.063	(.051)	-.157 *	(.060)	.041	(.032)
Constant	-1.320		1.258		-.041	
Adjusted R^2	.006		.107		.054	
F	1.187		4.557 *		3.761 *	

NOTES:
California N= 318
Michigan N= 326
Texas N= 534

* $p \leq .01$

Ability Variables
X_1 = Education (reported in survey by respondent) (highest grade completed) 0 = No schooling, 1 = 6th grade or less, 2 = 7th-9th grade, 3 = 10th-11th grade, 4 = high school graduate, 5 = some college, 6 = college graduate, 7 = post graduate study
X_2 = Respondent required assistance with reading questionnaire; 0 = No, 1 = Yes
X_3 = Number of skipped items for Current Conviction Section
X_4 = Number of skipped items for Arrest Incident Section
X_5 = Number of skipped items for Arrests for Crimes Done Section
Memory Variables
X_6 = Elapsed time since end of street months. This variable was built using the year arrested response multiplied by 12 and added to the number of the month arrested. (Months coded January = 01....December = 12)
X_7 = Number of felony convictions in lifetime (reported in survey by respondent) 1 = Never, 2 = once, 3 = 2-3 times, 4 = 4-6 times, 5 = 7-10 times, 6 = 11-15 times, 7 = 16 or more times (Indicator of Interference)
Record Match Problem Potential
X_8 = Recorded street months ended later than survey version. 0 = No, 1 = Yes
X_9 = Survey reported street months ended later than record version. 0 = No, 1 = Yes
Demographic Variable
X_{10} = Respondent age (reported in survey by respondent)
X_{11} = Respondent race (dummy variable) 0 = nonwhite, 1 = white

Table C.22B

MEANS, STANDARD DEVIATIONS, AND HIGHEST PREDICTOR INTERCORRELATION FOR PREDICTOR VARIABLES ON ARREST INCIDENT ITEMS RESPONDENT BIAS AND TOTAL ERROR SCORES BY STATE (SEE TABLE C.22A)

Variable	California			Michigan			Texas		
	mean	standard deviation	h.p.i.	mean	standard deviation	h.p.i.	mean	standard deviation	h.p.i.
Education (X_1)	3.57	1.19	X_2-.16	3.40	1.31	X_5-.20	3.22	1.29	X_2-.31
Personal Administration (X_2)	.03	.17	X_1-.16	.02	.13	X_7.16	.13	.33	X_1-.31
Missed Current Conviction Items (X_3)	.61	2.97	X_9.17	.64	3.05	X_4.18	.51	2.53	X_1-.06
Missed Arrest Incident Items (X_4)	.77	2.04	X_7.20	.75	2.09	X_3.18	.71	2.06	X_5.19
Missed Arrest Done Items (X_5)	.34	.86	X_7.08	.21	.65	X_1-.20	.22	.69	X_4.19
Time Since Arrest (X_6)	924.63	14.54	X_8-.34	915.95	24.39	X_{10}-.22	923.99	21.44	X_{10}-.27
Number of Felony Convictions (X_7)	3.29	1.16	X_{11}.20	3.04	.96	X_{11}.19	3.02	1.04	X_{10}.33
Record Street Months Later (X_8)	.16	.37	X_6-.34	.14	.34	X_5.19	.09	.29	X_6-.16
Survey Street Months Later (X_9)	.03	.16	X_3.17	.05	.22	X_6.12	.05	.22	X_2.10
Age (X_{10})	27.46	6.56	X_7.11	25.94	6.90	X_6-.22	26.97	9.17	X_7.33
Race (X_{11})	.42	.49	X_7.20	.32	.47	X_7.19	.36	.48	X_2-.19

NOTES:

California N = 318
Michigan N = 326
Texas N = 534

SYMBOLS:

h.p.i. = highest predictor intercorrelation

Table C.23A

REGRESSION COEFFICIENTS (β) FOR PREDICTOR VARIABLES ON TOTAL ERROR SCORE FOR ARREST INCIDENTS BY STATE

Variable	California β	California s.e.	Michigan β	Michigan s.e.	Texas β	Texas s.e.
Education (X_1)	-.010	(.019)	.035*	(.017)	.004	(.011)
Personal Administration (X_2)	.128	(.127)	-.127	(.162)	.004	(.042)
Missed Current Conviction Items (X_3)	-.006	(.008)	.006	(.007)	.003	(.005)
Missed Arrest Incident Items (X_4)	.022*	(.011)	.117*	(.011)	.054*	(.006)
Missed Arrest Done Items (X_5)	.031	(.025)	-.021	(.034)	-.001	(.019)
Time Since Arrest (X_6)	.0002	(.002)	-.0002	(.001)	-.001	(.001)
Number of Felony Convictions (X_7)	.029	(.020)	.045	(.023)	.047*	(.013)
Record Street Months Later (X_8)	.029	(.062)	.201*	(.065)	.003	(.046)
Survey Street Months Later (X_9)	-.044	(.143)	-.007	(.100)	-.018	(.061)
Age (X_{10})	-.006	(.003)	-.004	(.003)	-.003	(.002)
Race (X_{11})	-.042	(.045)	.016	(.047)	-.017	(.028)
Constant	.308		.242		1.276	
Adjusted R^2	.016		.289		.138	
F	1.481		13.033 *		8.786 *	

NOTES:
California N= 318
Michigan N= 326
Texas N= 534

* p.≤.01

Ability Variables
X_1 = Education (reported in survey by respondent) (highest grade completed) 0 = No schooling, 1 = 6th grade or less, 2 = 7th-9th grade, 3 = 10th-11th grade, 4 = high school graduate, 5 = some college, 6 = college graduate, 7 = post graduate study
X_2 = Respondent required assistance with reading questionnaire; 0 = No, 1 = Yes
X_3 = Number of skipped items for Current Conviction Section
X_4 = Number of skipped items for Arrest Incident Section
X_5 = Number of skipped items for Arrests for Crimes Done Section
Memory Variables
X_6 = Elapsed time since end of street months. This variable was built using the year arrested response multiplied by 12 and added to the number of the month arrested. (Months coded January = 01....December = 12)
X_7 = Number of felony convictions in lifetime (reported in survey by respondent) 1 = Never, 2 = once, 3 = 2-3 times, 4 = 4-6 times, 5 = 7-10 times, 6 = 11-15 times, 7 = 16 or more times (Indicator of Interference)
Record Match Problem Potential
X_8 = Recorded street months ended later than survey version. 0 = No, 1 = Yes
X_9 = Survey reported street months ended later than record version. 0 = No, 1 = Yes
Demographic Variable
X_{10} = Respondent age (reported in survey by respondent)
X_{11} = Respondent race (dummy variable) 0 = nonwhite, 1 = white

Table C.23B

MEANS AND STANDARD DEVIATIONS FOR PREDICTOR VARIABLES
ON ARREST INCIDENTS ITEMS
TOTAL ERROR SCORES BY STATE (SEE TABLE C.23A)

Variable	California		Michigan		Texas	
	mean	standard deviation	mean	standard deviation	mean	standard deviation
Education (X_1)	3.57	1.19	3.40	1.31	3.22	1.29
Personal Administration (X_2)	.03	.17	.02	.13	.13	.33
Missed Current Conviction Items (X_3)	.61	2.97	.64	3.05	.51	2.53
Missed Arrest Incident Items (X_4)	.77	2.04	.75	2.09	.71	2.06
Missed Arrest Done Items (X_5)	.34	.86	.21	.65	.22	.69
Time Since Arrest (X_6)	924.63	14.54	915.95	24.39	923.99	21.44
Number of Felony Convictions (X_7)	3.29	1.16	3.04	.96	3.02	1.04
Record Street Months Later (X_8)	.16	.37	.14	.34	.09	.29
Survey Street Months Later (X_9)	.03	.16	.05	.22	.05	.22
Age (X_{10})	27.46	6.56	25.94	6.90	26.97	9.17
Race (X_{11})	.42	.49	.32	.47	.36	.48

NOTES:
California N = 318
Michigan N = 326
Texas N = 534

Table C.24

ITEM BIAS SCORES FOR CONTINUOUS ARRESTS-FOR-CRIMES-DONE ITEMS
CALIFORNIA PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	333	.93	.86	.07	.17
Robbery	326	1.11	.94	.17	.16
Assault	329	.42	.53	-.11	.06
Theft	330	.55	.26	.29*	.11
Auto Theft	336	.17	.25	-.08*	.04
Forgery	330	.19	.22	-.03	.04
Fraud	338	.07	.00	.07*	.03
Drugs	337	.22	.26	-.04	.06

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.25

ITEM BIAS SCORES FOR CONTINUOUS ARRESTS-FOR-CRIMES-DONE ITEMS
MICHIGAN PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	352	.73	.32	.41*	.14
Robbery	346	.49	.37	.12	.08
Assault	350	.29	.27	.03	.05
Theft	347	.58	.24	.34*	.15
Auto Theft	351	.17	.12	.04	.03
Forgery	356	.08	.13	-.04	.03
Fraud	353	.02	.01	.01	.01
Drugs	355	.19	.17	.02	.04

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.26

ITEM BIAS SCORES FOR CONTINUOUS ARRESTS-FOR-CRIMES-DONE ITEMS
TEXAS PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	566	.71	.69	.02	.10
Robbery	566	.30	.30	.01	.05
Assault	571	.16	.13	.02	.02
Theft	563	.36	.31	.04	.06
Auto Theft	574	.14	.09	.06*	.02
Forgery	562	.15	.14	.01	.03
Fraud	573	.04	.03	.02	.02
Drugs	574	.15	.19	-.04*	.02

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.27

ITEM BIAS SCORES FOR YES-NO ARRESTS-FOR-CRIMES-DONE ITEMS
CALIFORNIA PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	333	.31	.42	-.11*	.03
Robbery	327	.35	.49	-.14*	.02
Assault	329	.26	.34	-.08*	.03
Theft	331	.19	.20	-.01	.02
Auto Theft	336	.11	.19	-.08*	.02
Forgery	332	.11	.09	.02	.02
Fraud	338	.02	.00	.02*	.01
Drugs	337	.13	.10	.03	.02

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C. 28

ITEM BIAS SCORES FOR YES-NO ARRESTS-FOR-CRIMES-DONE ITEMS
MICHIGAN PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	352	.24	.22	.01	.02
Robbery	346	.25	.29	-.04	.03
Assault	350	.17	.21	-.04	.02
Theft	348	.17	.19	-.02	.02
Auto Theft	351	.10	.08	.03	.02
Forgery	356	.04	.06	-.02	.01
Fraud	353	.01	.01	.01	.01
Drugs	355	.13	.11	.02	.02

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.29

ITEM BIAS SCORES FOR YES-NO ARRESTS-FOR-CRIMES-DONE ITEMS
TEXAS PRISON SAMPLE

Crime Category	N	Survey Mean	Record Mean	"Bias" S-R	Standard Error of Difference
Burglary	566	.31	.47	-.16*	.02
Robbery	566	.13	.22	-.09*	.02
Assault	571	.12	.11	.01	.02
Theft	563	.15	.21	-.06*	.02
Auto Theft	574	.08	.08	.00	.01
Forgery	563	.08	.08	-.00	.01
Fraud	573	.02	.02	-.00	.01
Drugs	574	.10	.13	-.03*	.01

NOTES:

Symbols:
N Number of Observations
S-R Survey Mean Minus Record Mean

*2 Standard Error Confidence Interval of S-R does not include zero.

Table C.30

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR CONTINUOUS ARRESTS-FOR-CRIMES-DONE ITEMS
CALIFORNIA PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	333	.25	.50	57	.78
Robbery	326	.54	.73	57	1.00
Assault	329	.25	.50	53	.85
Theft	330	.15	.39	54	.99
Auto Theft	336	.36	.60	56	.94
Forgery	330	.63	.79	55	1.00
Fraud	338	--	--	58	1.00
Drugs	337	.22	.47	57	.79

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
--No Variation Observed

Table C.31

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR CONTINUOUS ARRESTS-FOR-CRIMES-DONE ITEMS
MICHIGAN PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	352	.28	.53	42	.98
Robbery	346	.16	.40	41	.97
Assault	350	.27	.52	44	.50
Theft	347	.08	.29	43	.45
Auto Theft	351	.38	.62	44	.59
Forgery	356	.60	.78	44	1.00
Fraud	353	.00†	.00	45	--
Drugs	355	.38	.62	45	.89

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
†Fraud Correlation of -.0077 Was Truncated to .00
--No Variation Observed

Table C.32

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR CONTINUOUS ARRESTS-FOR-CRIMES-DONE ITEMS
TEXAS PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	566	.21	.46	133	.81
Robbery	566	.51	.72	133	.79
Assault	571	.29	.54	134	.87
Theft	563	.15	.39	129	.53
Auto Theft	574	.38	.62	137	.88
Forgery	562	.41	.64	132	.92
Fraud	573	.05	.21	138	.58
Drugs	574	.57	.75	135	.91

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest

Table C.33

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR YES-NO VERSION ARRESTS-FOR-CRIMES-DONE ITEMS
CALIFORNIA PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	333	.42	.65	57	.82
Robbery	327	.60	.77	57	.94
Assault	329	.28	.52	53	.65
Theft	331	.34	.58	55	.71
Auto Theft	336	.42	.65	57	.84
Forgery	332	.50	.71	55	1.00
Fraud	338	--	--	58	1.00
Drugs	337	.44	.66	57	.79

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
--No Variation Observed

Table C.34

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR YES-NO VERSION ARRESTS-FOR-CRIMES-DONE ITEMS
MICHIGAN PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	352	.43	.66	42	.82
Robbery	346	.38	.62	41	.89
Assault	350	.29	.54	44	.74
Theft	348	.24	.49	43	.70
Auto Theft	351	.33	.57	44	.85
Forgery	356	.47	.68	44	1.00
Fraud	353	.00†	.00	45	--
Drugs	355	.57	.76	45	.74

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest
†Fraud Correlation of -.0081 Was Truncated to .00
--No Variation Observed

Table C.35

RELIABILITY ESTIMATES (MULTIPLICATIVE CORRELATION ATTENUATION FACTORS)
FOR YES-NO VERSION ARRESTS-FOR-CRIMES-DONE ITEMS
TEXAS PRISON SAMPLE

Crime Category	Record Method			Retest Method	
	N	Lower Bound*	Upper Bound**	N	$\sqrt{r}$ ***
Burglary	566	.49	.70	133	.81
Robbery	566	.55	.74	133	.91
Assault	571	.32	.57	134	.86
Theft	563	.21	45	29	.74
Auto Theft	574	.35	.59	137	.90
Forgery	563	.49	.70	133	.94
Fraud	573	.08	.28	138	.76
Drugs	574	.54	.74	135	.90

NOTES:
*Correlation of the Survey and Record
**Square Root of the Correlation of the Survey and Record
***Square Root of the Correlation of the Survey and Retest

Table C.36A

REGRESSION COEFFICIENTS (β) FOR PREDICTOR VARIABLES ON RESPONDENT
BIAS SCORES FOR ARRESTS-FOR-CRIMES-DONE BY STATE

Variable	California β	California s.e.	Michigan β	Michigan s.e.	Texas β	Texas s.e.
Education (X_1)	.008	(.020)	-.011	(.013)	-.020	(.010)
Personal Administration (X_2)	.195	(.136)	.165	(.127)	-.008	(.042)
Missed Current Conviction Items (X_3)	.006	(.008)	-.006	(.005)	.004	(.004)
Missed Arrest Incident Items (X_4)	-.007	(.010)	-.012	(.007)	.004	(.005)
Missed Arrest Done Items (X_5)	.069 *	(.025)	.153 *	(.027)	-.003	(.016)
Time Since Arrest (X_6)	.0001	(.002)	.0001	(.001)	-.0001	(.001)
Number of Felony Convictions (X_7)	.049 *	(.021)	.028	(.018)	.044 *	(.012)
Record Street Months Later (X_8)	.078	(.067)	.012	(.049)	.0004	(.044)
Survey Street Months Later (X_9)	-.099	(.153)	-.088	(.078)	.032	(.056)
Age (X_{10})	-.003	(.004)	-.006 *	(.002)	-.004 *	(.001)
Race (X_{11})	.009	(.048)	-.055	(.036)	.018	(.027)
Constant	-.372		.058		.097	
Adjusted R^2	.021		.136		.017	
F	1.633		5.809 *		1.876	

NOTES:
California N= 325
Michigan N= 338
Texas N= 562

* $p \leq .01$

Ability Variables
X_1 = Education (reported in survey by respondent) (highest grade completed) 0 = No schooling, 1 = 6th grade or less, 2 = 7th-9th grade, 3 = 10th-11th grade, 4 = high school graduate, 5 = some college, 6 = college graduate, 7 = post graduate study
X_2 = Respondent required assistance with reading questionnaire; 0 = No, 1 = Yes
X_3 = Number of skipped items for Current Conviction Section
X_4 = Number of skipped items for Arrest Incident Section
X_5 = Number of skipped items for Arrests for Crimes Done Section
Memory Variables
X_6 = Elapsed time since end of street months. This variable was built using the year arrested response multiplied by 12 and added to the number of the month arrested. (Months coded January = 01....December = 12)
X_7 = Number of felony convictions in lifetime (reported in survey by respondent) 1 = Never, 2 = once, 3 = 2-3 times, 4 = 4-6 times, 5 = 7-10 times, 6 = 11-15 times, 7 = 16 or more times (Indicator of Interference)
Record Match Problem Potential
X_8 = Recorded street months ended later than survey version. 0 = No, 1 = Yes
X_9 = Survey reported street months ended later than record version. 0 = No, 1 = Yes
Demographic Variable
X_{10} = Respondent age (reported in survey by respondent)
X_{11} = Respondent race (dummy variable) 0 = nonwhite, 1 = white

Table C.36B

MEANS, STANDARD DEVIATIONS, AND HIGHEST PREDICTOR INTERCORRELATION FOR PREDICTOR VARIABLES ON ARRESTS-FOR-CRIMES-DONE ITEMS, RESPONDENT BIAS AND TOTAL ERROR SCORES BY STATE (SEE TABLE C.36A)

Variable	California			Michigan			Texas		
	mean	standard deviation	h.p.i.	mean	standard deviation	h.p.i.	mean	standard deviation	h.p.i.
Education (X_1)	3.56	1.19	X_2-.16	3.41	1.31	X_5-.20	3.19	1.30	X_2-.31
Personal Administration (X_2)	.03	.17	X_1-.16	.02	.13	X_7.16	.12	.33	X_1-.31
Missed Current Conviction Items (X_3)	.60	2.94	X_9.17	.67	3.09	X_4.16	.67	2.88	X_4.15
Missed Arrest Incident Items (X_4)	.95	2.35	X_7.17	1.04	2.56	X_3.16	1.12	2.70	X_5.21
Missed Arrest Done Items (X_5)	.36	.93	X_4.09	.21	.64	X_1-.20	.25	.78	X_4.21
Time Since Arrest (X_6)	924.48	14.75	X_8-.33	915.60	24.74	X_{10}-.26	923.38	22.41	X_{10}-.28
Number of Felony Convictions (X_7)	3.29	1.17	X_{11}.21	3.05	.96	X_{11}.18	3.02	1.06	X_{10}.29
Record Street Months Later (X_8)	.16	.37	X_6-.33	.14	.35	X_5.18	.09	.29	X_6-.16
Survey Street Months Later (X_9)	.02	.16	X_3.17	.05	.21	X_6.12	.06	.23	X_{10}.14
Age (X_{10})	27.48	6.53	X_7.12	26.13	7.23	X_6-.26	27.37	9.53	X_7.29
Race (X_{11})	.42	.49	X_7.21	.31	.46	X_7.18	.36	.48	X_2-.19

NOTES:

California N = 325
Michigan N = 338
Texas N = 562

SYMBOLS:

h.p.i. = highest predictor intercorrelation

Table C.37A

REGRESSION COEFFICIENTS (β) FOR PREDICTOR VARIABLES ON TOTAL ERROR SCORE FOR ARRESTS-FOR-CRIMES-DONE BY STATE

Variable	California β	California s.e.	Michigan β	Michigan s.e.	Texas β	Texas s.e.
Education (X_1)	-.001	(.018)	.006	(.011)	-.004	(.009)
Personal Administration (X_2)	.113	(.121)	.024	(.110)	-.027	(.037)
Missed Current Conviction Items (X_3)	-.006	(.007)	.001	(.005)	-.001	(.004)
Missed Arrest Incident Items (X_4)	.002	(.009)	-.004	(.006)	.005	(.004)
Missed Arrest Done Items (X_5)	-.005	(.022)	.102 *	(.023)	.005	(.014)
Time Since Arrest (X_6)	-.002	(.001)	-.0004	(.001)	.0004	(.0005)
Number of Felony Convictions (X_7)	.033	(.019)	.045 *	(.016)	.060 *	(.011)
Record Street Months Later (X_8)	.055	(.059)	.068	(.042)	.077	(.039)
Survey Street Months Later (X_9)	.059	(.136)	.019	(.068)	-.003	(.050)
Age (X_{10})	-.006	(.003)	-.008 *	(.002)	-.004 *	(.001)
Race (X_{11})	.024	(.043)	-.020	(.031)	.007	(.024)
Constant	1.888		.699		-.187	
Adjusted R^2	.005		.110		.054	
F	1.140		4.773 *		3.896 *	

NOTES:
California N= 325
Michigan N= 338
Texas N= 562

* $p \leq .01$

Ability Variables
X_1 = Education (reported in survey by respondent) (highest grade completed) 0 = No schooling, 1 = 6th grade or less, 2 = 7th-9th grade, 3 = 10th-11th grade, 4 = high school graduate, 5 = some college, 6 = college graduate, 7 = post graduate study
X_2 = Respondent required assistance with reading questionnaire; 0 = No, 1 = Yes
X_3 = Number of skipped items for Current Conviction Section
X_4 = Number of skipped items for Arrest Incident Section
X_5 = Number of skipped items for Arrests for Crimes Done Section
Memory Variables
X_6 = Elapsed time since end of street months. This variable was built using the year arrested response multiplied by 12 and added to the number of the month arrested. (Months coded January = 01....December = 12)
X_7 = Number of felony convictions in lifetime (reported in survey by respondent) 1 = Never, 2 = once, 3 = 2-3 times, 4 = 4-6 times, 5 = 7-10 times, 6 = 11-15 times, 7 = 16 or more times (Indicator of Interference)
Record Match Problem Potential
X_8 = Recorded street months ended later than survey version. 0 = No, 1 = Yes
X_9 = Survey reported street months ended later than record version. 0 = No, 1 = Yes
Demographic Variable
X_{10} = Respondent age (reported in survey by respondent)
X_{11} = Respondent race (dummy variable) 0 = nonwhite, 1 = white

Table C.37B

MEANS AND STANDARD DEVIATIONS FOR PREDICTOR VARIABLES
ON ARRESTS-FOR-CRIMES-DONE ITEMS
TOTAL ERROR SCORE BY STATE (SEE TABLE C.37A)

Variable	California		Michigan		Texas	
	mean	standard deviation	mean	standard deviation	mean	standard deviation
Education (X_1)	3.56	1.19	3.41	1.31	3.19	1.30
Personal Administration (X_2)	.03	.17	.02	.13	.12	.33
Missed Current Conviction Items (X_3)	.60	2.94	.67	3.09	.67	2.88
Missed Arrest Incident Items (X_4)	.95	2.35	1.04	2.56	1.12	2.70
Missed Arrest Done Items (X_5)	.36	.93	.21	.64	.25	.78
Time Since Arrest (X_6)	924.48	14.75	915.60	24.74	923.38	22.41
Number of Felony Convictions (X_7)	3.29	1.17	3.05	.96	3.02	1.06
Record Street Months Later (X_8)	.16	.37	.14	.35	.09	.29
Survey Street Months Later (X_9)	.02	.16	.05	.21	.06	.23
Age (X_{10})	27.48	6.53	26.13	7.23	27.37	9.53
Race (X_{11})	.42	.49	.31	.46	.36	.48

NOTES:
California N = 325
Michigan N = 338
Texas N = 562

Table C.38

RELIABILITY ESTIMATES (LOWER BOUND MCAF)
FOR CONTINUOUS ARRESTS-FOR-CRIMES-DONE ITEMS BY STATE
RETEST SAMPLE

Crime Category	California			Michigan			Texas		
	N	S,R r	R,R r	N	S,R r	R,R r	N	S,R r	R,R r
Burglary	53	.30	.35	34	.25	.45	123	.47	.46
Robbery	53	.47	.46	34	.01	.08	124	.68	.63
Assault	50	.38	.37	36	.66	.61	123	.42	.30
Theft	50	-.03	-.04	35	.03	-.00	120	.04	.20
Auto Theft	53	.22	.03	36	.07	.13	126	.52	.75
Forgery	52	1.00	1.00	36	1.00	1.00	121	.56	.71
Fraud	54	--	--	37	--	--	127	-.02	-.02
Drugs	53	.57	.25	37	--	--	125	.43	.66

NOTES:

Symbols:
N = Number of Observations
S,R = Survey Record
R,R = Retest Record
r = Correlation Coefficient
-- = No Variation Observed

Table C.39

RELIABILITY ESTIMATES (LOWER BOUND MCAF)
FOR CONTINUOUS ARREST INCIDENT ITEMS BY STATE
RETEST SAMPLE

Crime Category	California N	California S,R r	California R,R r	Michigan N	Michigan S,R r	Michigan R,R r	Texas N	Texas S,R r	Texas R,R r
Burglary	46	.53	.51	32	.62	.09	108	.50	.52
Robbery	47	.42	.43	35	-.04	-.04	99	.40	.55
Assault	46	.40	.38	32	.67	.70	96	.48	.26
Homicide	46	.76	.76	31	-.07	-.05	95	--	.70
Auto Theft	44	.65	.67	32	-.07	.34	97	-.06	.35
Theft	43	.63	.32	31	-.20	.25	99	.59	.52
Forgery	43	.70	.56	32	1.00	1.00	98	.77	.64
Fraud	43	--	--	32	--	--	93	-.02	-.02
Drugs	44	.50	45	32	--	--	98	.45	.60

NOTES:

Symbols:
N = Number of Observations
S,R = Survey Record
R,R = Retest Record
r = Correlation Coefficient
-- = No Variation Observed

Appendix D

SELECTED SURVEY QUESTIONS: SPANISH LANGUAGE VERSION

This appendix contains reproductions of the questions, in the Spanish version of the questionnaire, which were used in our analyses. Exhibit D.1 and D.2 show the set of questions and the calendar used to establish the Street Months. Exhibit D.3 gives the Current Conviction questions and Exhibit D.4 the Arrest Incidents questions. Exhibit D.5 shows the questions on Arrests for Crimes Done, and Exhibit D.6 the questions selected to generate the Predictor Variables.

CONTINUED

2 OF 3

Exhibit D.1

STREET MONTHS QUESTIONS

1. La siguiente parte se refiere al período antes de que fue arrestado por la condena presente. Hay una tarjeta azul que contiene un calendario. Las instrucciones en esta página le explican como llenarlo. Levante la mano si encuentra problema en llenarlo.

2. ¿Por esta condena presente, en qué año fue arrestado? (Si fue arrestado varios veces por esta sentencia, ponga el arresto más corriente.)

   Año Arrestado: _____ *8 ,,,,/*

   Escriba ese año donde dice "Año Arrestado" en el calendario.

3. ¿En qué mes de ese año fue aquel arresto?

   Mes Arrestado: _____ *12 ''/*

   Escriba "arrestado" en el calendario en ese mes (por el "Año Arrestado").

4. Ahora, dibuje una linea por todos los meses después de ese mes (hasta el extremo del año).

5. No le preguntaremos nada acerca de lo qué le sucedió en los meses marcado con una linea.

6. ¿Cuál fue el año antes de que fue arrestado?

   Año Antes del Arresto: _____ *14 ''''/*

   Escriba ese año en el calendario donde dice "Año Antes del Arresto".

7. ¿Durante todo los meses en el calendario, antes de que fue arrestado (incluyendo ambos años), estuvo encarcelado un mes o más? *18/*

   NO ☐ 1 SI ☐ 2 ➧ Ponga X en todos los meses que estuvo encarcelado. (Si no puede recordar exactamente, piense qué temporada fue y ponga X en todos los meses que fue encarcelado durante esa temporada).

CARD 03

8. Ahora, véase el calendario. Todas las cajas blancas (sin X o lineas) son meses que estuvo afuera de la cárcel antes de que fue arrestado.

9. Cuente todas las cajas blancas. ¿Cuántos meses fueron? ⟶ _____ Meses *19 ''/*

10. Será cuestionado acerca de estos y también acerca del mes que marcó Vd. "Arrestado". Para obtener la suma de estos meses añada un mes y escriba el total aquí. ⟶ + 1 = _____ Todos Meses Fuera de la Cárcel *21 ''/*

11. Escriba este número total en la caja en el calendario donde dice "MESES EN EL CALENDARIO FUERA DE LA CARCEL". Necesitará este número para contestar las preguntas siguientes.

12. Bajo el mes marcado "Arrestado" escriba "Incluya este mes."

    Por escribir esto, le recordará incluir este mes en las respuestas. *23/*

Exhibit D.2

CALENDAR FOR CALCULATING STREET MONTHS

21542

LAS INSTRUCCIONES PARA EL USO DE ESTE CALENDARIO ESTAN INCLUIDAS EN EL ESTUDIO

	INVIERNO		PRIMAVERA			VERANO			OTOÑO			
1977 — AÑO ANTES DE SER ARRESTADO	enero	febrero	marzo	abril	mayo	junio	julio	agosto	septiembre	octubre	noviembre	diciembre
			X	X	X							
1978 — AÑO ARRESTADO	enero	febrero	marzo	abril	mayo	junio	julio	agosto	septiembre	octubre	noviembre	diciembre
						X						

Incluya este mes

MESES EN EL CALENDARIO FUERA DE LA CARCEL → 15

Exhibit D.3

CURRENT CONVICTION QUESTION

6. Estas preguntas se refieren, solamente a los crimenes (o crimen) por los que ahora esta cumpliendo esta condena. ¿De cuáles cargo(s) fue culpado, por los que le tienen encarcelado presentemente? *(Marque todo lo aplicable.)*

☐1	Asalto/AAM (Asalto con arma mortal)	43/
☐	Robo de automóvil/Robo de vehículo	44/
☐	Robo con escalamiento	45/
☐	Posesión de drogas	46/
☐	Venta de drogas	47/
☐	Falsificación/Cheques malos/Tarjeta de crédito mala	48/
☐	Fraude o estafa	49/
☐	Secuestro	50/
☐	Asesinato/Homicidio sin premeditación	51/
☐	Posesión o recibir propiedad robada	52/
☐	Rapto	53/
☐	Robo	54/
☐	Ofensa sexual (no siendo rapto)	55/
☐	Hurto/Hurto grande/Ratería	56/
☐	Delito de armas	57/
☐	Otro, ¿cuál? _____	58/

Exhibit D.4

ARREST INCIDENTS QUESTIONS

4. Véase el calendario. ¿Durante los MESES EN EL CALENDARIO FUERA DE LA CARCEL, cuántas veces fue arrestado por cada de estos crimenes? Cuente como un arresto aunque en actualidad, no cometió el crimen por lo que fue arrestado. *(Marque* NINGUN *si no fue arrestado por ese crimen.)*

ROBO CON ESCALAMIENTO	_____	arrestos	0	NINGUN ☐$_{00}$	12 ''/
ROBO O ROBO CON ARMA	_____	arrestos	0	NINGUN ☐	14 ''/
ASALTO, ASALTO AGRAVADO O ASALTO CON ARMA	_____	arrestos	0	NINGUN ☐	16 ''/
ASESINATO O HOMECIDIO SIN PREMEDITACION	_____	arrestos	0	NINGUN ☐	18 ''/
ROBO DE AUTOMOVIL, ROBO DE UN VEHICULO CON MOTOR	_____	arrestos	0	NINGUN ☐	20 ''/
HURTO, HURTO GRANDE, RATERIA, RATERIA GRANDE	_____	arrestos	0	NINGUN ☐	22 ''/
FALSIFICACION, USO DE TARJETA DE CREDITO MALA O ROBADA O USO DE CHEQUES MALOS	_____	arrestos	0	NINGUN ☐	24 ''/
FRAUDE	_____	arrestos	0	NINGUN ☐	26 ''/
VENDER DROGAS, POSEER DROGAS PARA LA VENTA O TRANSPORTAR DROGAS	_____	arrestos	0	NINGUN ☐	28 ''/

Exhibit D.5

ARRESTS FOR CRIMES DONE QUESTIONS

Burglary

1. ¿Durante los MESES EN EL CALENDARIO FUERA DE LA CARCEL cometió robos con escalamiento? (Cuente todas las veces que se metió a una casa o automóvil o comercio para robar.)

SI ☐$_1$ NO ☐$_2$ ➧ continue en la página 18

5. ¿Por cuántos de estos robos con escalamiento fue arrestado? (Incluya todos los arrestos por robo con escalamiento aunque fue cargado por otro crimen.)

_____ Arrestos por robos con escalamiento

Business Robbery

1. ¿Durante los MESES EN EL CALENDARIO FUERA DE LA CARCEL, robó algun negocio? Eso es decir, robó con arma una tienda, gasolinera, banco, taxi, u otro negocio?

SI ☐$_1$ NO ☐$_2$ ➧ continue en la página 20

5. ¿En cúantos de estos robos fue arrestado? (Incluyendo todas las veces que fue arrestado por robo de negocio aunque fue acusado por otra cosa.)

_____ Arrestos por robos de negocio

Personal Robbery

1. ¿Durante los MESES EN EL CALENDARIO FUERA DE LA CARCEL, robó a alguien, atacó, cometió robo de calle, arrebatimiento de bolsa o robó con arma en casa o automóvil? (No incluya robos de negocio o robos con arma durante un robo con escalamiento que ya mencionó.)

SI ☐$_1$ NO ☐$_2$ ➧ continue en la página 22

5. ¿En cuántos de estos robos fue arrestado? (Incluyendo todas las veces que fue arrestado por robo de persona, aunque fue acusado por otro crimen.)

_____ Arrestos por robos de persona

Assault

1. ¿Durante los MESES EN EL CALENDARIO FUERA DE LA CARCEL, aunque nadie fue lastimado, asaltó a alguien, amenazó a alguien con una arma, trató de balazear o apuñalar a alguien, o golpear o estrangular a alguien?

SI ☐1 NO ☐2 ➡ continue en la página 26

4. ¿Cuántas veces fue arrestado cuando asaltó, amenazó, disparó, trató de apuñalar, o golpear, o estranguló a alguien?

_____ Arrestado

Theft

1. ¿Durante los MESES EN EL CALENDARIO FUERA DE LA CARCEL, cometió un hurto o ratería de tiendas? Eso es decir, robo de una caja o registro, robo de tiendas, carterista o robar a alguien sin que supieran. (No incluya robo de automóvil.)

SI ☐1 NO ☐2 ➡ continue en la página 28

5. ¿En cuántos de estos hurtos fue arrestado? (Incluya todas las veces que fue arrestado por hurto, aunque fue cargado por otro crimen.)

_____ Arrestos por hurtos

Auto Theft

1. ¿Durante los MESES EN EL CALENDARIO FUERA DE LA CARCEL, cometió robos de automóvil, camión, motorcicleta?

SI ☐1 NO ☐2 ➡ continue en la página 30

5. ¿Por cuántos robos de vehiculos fue arrestado? (Incluya todas las veces que fue arrestado por robo de vehículo, aunque fue cargado por otra cosa.)

_____ Arrestos por robos de vehículo

Forgery, Card

1. ¿Durante los MESES EN EL CALENDARIO FUERA DE LA CARCEL, falsificó algo, usó una tarjeta de crédito mala o robada, o falsificó un cheque?

SI ☐1 NO ☐2 ➡ continue en la página 32

5. ¿Fue arrestado Vd. por cuántos de estos falsificaciones, tarjetas o cheques malos? (Incluya todas las veces que fue arrestado por una de estas cosas, aunque fue acusado por otra.)

_____ Arrestos

Fraud

1. ¿Durante los MESES EN EL CALENDARIO FUERA DE LA CARCEL, cometió fraude o estafa a persona, negocio, o gobierno?

SI ☐1 NO ☐2 ➡ continue en la página 34

5. ¿En cuántos de estos fraudes o estafas fue arrestado? (Incluyendo todas las veces que fue arrestado por fraude o estafa aunque fue acusado por otra cosa.)

_____ Arrestado por fraudes o estafas

Drugs

1. ¿Durante los MESES EN EL CALENDARIO FUERA DE LA CARCEL vendió drogas? Eso es decir, hizo, vendió, fue contrabandista o movió drogas?

SI ☐1 NO ☐2 ➡ continue en la página 36

5. ¿Cuántas veces fue arrestado por ventas de drogas?

_____ Arrestos por drogas

Exhibit D.6

SOURCE QUESTIONS FOR SELECTED PREDICTOR VARIABLES

X_1 Education

¿Cuál fué el grado o división más alto que completó en la escuela?

- ☐ 0 Ninguna eseñanza
- ☐ 1 Sexto grado o menos
- ☐ 2 7-9 grado
- ☐ 3 10-11 grado
- ☐ 4 Graduado de high school
- ☐ 5 Algo de colegio
- ☐ 6 Graduado de colegio
- ☐ 7 Estudio posgraduado

X_6 Time Since Arrest

¿Por esta condena presente, en qué año fue arrestado? (Si fue arrestado varios veces por esta sentencia, ponga el arresto más corriente.)

Año Arrestado: _____

Escriba ese año donde dice "Año Arrestado" en el calendario.

X_7 Number of Felony Convictions

¿En toda su vida, cuántas veces ha sido condenado de una felonía?

- ☐ Nunca
- ☐ Una vez
- ☐ 2-3 veces
- ☐ 4-6 veces
- ☐ 7-10 veces
- ☐ 11-15 veces
- ☐ 16 o más veces

X_{10} Age

¿Cuántos años cumplió el día de su último cumpleaños?

_____ Años

X_{11} Race

¿De qué raza es Vd.?

- ☐ 1 Asiático
- ☐ 2 Negro
- ☐ 3 Chicano/Latino
- ☐ 4 Indio/Nativo-Americano
- ☐ 5 Blanco
- ☐ 6 Otro

Appendix E

FIELDWORK FORMS AND EXAMPLE RAP SHEET

This appendix contains a reproduction of the notice we mailed to inmates selected to participate in the survey, and also a reproduction of the Agreement to Participate form. In addition, it contains an example of a California RAP Sheet.

Exhibit E.1

NOTICE MAILED TO SAMPLED INMATES

The Rand Corporation, a research company, will be doing a survey of men in this facility. This is part of a nationwide survey of men in prisons and jails. Your name has been randomly selected for the survey. You will be scheduled for a meeting where the researchers will explain the survey. You may then choose whether or not to take the survey. If you choose to take the survey, it will be given at that meeting. The survey will take about one hour and you will receive $5.00 to your account.

Exhibit E.2

DESCRIPTION OF SURVEY/AGREEMENT TO PARTICIPATE *

This form describes the Rand Jail/Prison Survey. It is also the form which you use to indicate that you agree to take the survey. If you agree to participate in the survey, print your name in the space on this form.

I agree to participate in a survey being conducted by The Rand Corporation. I understand that The Rand Corporation is a private, nonprofit corporation that does research on public policy issues. I understand further that the purpose of the survey is to collect information from men who are serving time in prisons and jails to find out our opinions and experiences with the criminal justice system, how we are treated when we are in jail or prison, and what are our opinions, past activities, and experiences in doing crime.

I understand that I will be given a booklet of questions to answer. The booklet has a number on it but I do not need to print my name on this booklet. I agree to print my name in the space provided on this form which has the same number as the booklet. My name may be retained for followup research but my name will be kept in a separate place from my answers.

I understand that The Rand Corporation will use the numbered sheet to combine my answers with information about my arrests, classification, and treatment by the criminal justice system. Researchers will collect this information from records kept by criminal justice agencies--such as police, courts, jails, and prisons.

I understand that Rand will use my answers to questions in the survey booklet and the information they collect from criminal justice agencies only for the purposes of research. Federal law requires that my answers and all of the other information collected by the researchers be kept strictly confidential. The law provides that copies of my answers are immune from legal process and cannot be admitted as evidence in any judicial or administrative proceeding without my written consent.* This means that unless I agree, no court, police department, jail or prison can get copies of my answers from the researchers. However, I understand that the law makes no mention of legislative proceedings and may not protect this information from a legislative subpoena.

I understand that my participation is completely voluntary. I do not have to participate in the survey and I do not have to give permission to The Rand Corporation to obtain information about my arrests, classification, and treatment by criminal justice agencies. By answering the questions in the survey I am agreeing to participate and to permit The Rand Corporation to obtain such information from criminal justice agencies. I can refuse to answer the questions either now or after I have seen the survey booklet. The only benefits to me from answering all the questions are that I will receive a payment of $5.00 and that I may later be asked to volunteer to participate in another survey, for which I will also be paid.

PLEASE PRINT YOUR NAME HERE _____ _____
FIRST LAST

11557 _____
Prison/Jail Number

*42 U.S. Code 3771(a) says:
"No officer or employee of the Federal Government, nor any recipient of assistance under the provisions of this chapter shall use or reveal any research or statistical information furnished under this chapter by any person and identifiable to any specific private person for any purpose other than the purpose for which it was obtained in accordance with this chapter. Copies of such information shall be immune from legal process, and shall not, without the consent of the person furnishing such information, be admitted as evidence or used for any purpose in any action, suit, or other judicial or administrative proceedings."

*reproduction is reduced 23%

Exhibit E.3

EXAMPLE CALIFORNIA RAP SHEET

STATE OF CALIFORNIA
DEPARTMENT OF JUSTICE
BUREAU OF IDENTIFICATION
P. O. BOX 13417, SACRAMENTO

The following CII record, NUMBER **IS FOR OFFICIAL USE ONLY**

PAGE 2

ARRESTED OR RECEIVED	DEPARTMENT AND NUMBER	NAME	CHARGE	DISPOSITION
9-26-61	PD SACRAMENTO, 52257		1.VIO.PROB. 2. 647F PC	9-28-61, 1.,FNF.6 MO CO JL; 2.,90 DS,CWGB 3 YRS
10-9-61	SO SACRAMENTO 25625		BURG	J/S,3 YRS PROB.
1-23-63	PD SACRAMENTO, 52257		GTA,BURG.2 CTS	3-6-63,BURG 2 CTS, DISM;CT 3 & GTA,ST. PR.
3-8-63	CALIF. DEPT. OF CORRECTIONS, A 76934		GT & BURG 2ND DEG(487 & 459 PC)	FROM: SACTO CO TERM:6 MOS-10 YRS & 6 MOS-15 YRS CC 5-17-65,PAROLED(10) 3-8-70,TENT.DISCH.
1-14-66	PD SAN FRANCISCO, 206079		1. 647F PC U/INFLU.ALC 2. 4143:B&P UNLAW.POSS. HYPO NEEDLE OR SYRINGE 3. E/R SACTO (211 PC/ROBB) ENR 4. DIR.CORR. (PER SEC 3056 PC/CUST OF PAROLED PRISONER)	1-14-66, 1.,30 DS SS 2-4-66, 2.,30 DS CJ CTS 2-13-66, 3.,DEL 1-14-66, 4.,DEL
2-12-66	PD SACRAMENTO, 52257		211 PC	3-23-66,COMM. TO ST. PRISON
2-14-66	SO SACRAMENTO, 25 625		211 PC/ROBB (CITY REMAND)	--

CONTINUED PAGE 3

ENTRIES INDICATED BY ASTERISK (*) ARE NOT VERIFIED BY FINGERPRINTS IN CII FILES.

Appendix F

BIAS IN THE REGRESSION COEFFICIENT IF THE DIFFERENCE-SCORE DEPENDENT VARIABLE IS REGRESSED ON ONE OF ITS COMPONENTS AND THE COMPONENT IS MEASURED WITH ERROR

The following explains why we did not use the number of reported or recorded arrests (or convictions) as predictors of the Individual Difference Bias Score. We want to know whether response error is a function of the number of events (arrests or convictions) to be recalled and we wish to use the recorded value as the predictor variable. The Individual Difference Bias Score is the dependent variable and is the difference between the number of reported and recorded events. The following shows that the estimated coefficient of the regression of the Difference Bias Score on the recorded value is negatively biased.

We wish to fit the linear function

$$\Delta = \alpha + \gamma T + \mu \qquad (1)$$

where

Δ = Bias (difference),

T = True number of events,

γ = parameter indicating the relationship of set size to bias,

α = constant term,

μ = stochastic equation error term.

Define the survey response, S, as a linear function of True score (T), bias correlated with true score (bT), and random error (e):

$$S = T + bT + e.$$

Define the record observation similarly:

$$R = T + b'T + e'.$$

Assume

$$\text{cov}\ (e,T) = \text{cov}\ (e',T) = \text{cov}\ (e,e') = 0.$$

These are the same structural models and assumptions used in this report. The uncorrelated biases (B, B´) are assumed to have no variance and are omitted.

The empirical bias score is defined as S - R and S - R = T(b - b´) + (e - e´).

The available data are used to estimate Equation (1) as follows:

$$S - R = a + gR + u$$

where

$$g = \text{cov}\ (S - R, R)/\text{var}\ R$$
$$= [(b - b')\ \text{var}\ T + b'(b - b')\ \text{var}\ T - \text{var}\ e']/\text{var}\ R. \qquad (2)$$

Now, assume that b = 0, and observe that $g \neq b \neq 0$

$$g = -[(b'+1)b'\ \text{var}\ T + \text{var}\ e']/\text{var}\ R.$$

The estimate is negatively biased.

g is an unbiased estimate of b if the record measure is perfect. Substituting (b´ = var e´ = 0) in eq 2:

$$g = (b\ \text{var}\ T)/\text{var}\ T = b.$$

Appendix G

WORKSHEET FOR ADDITIONAL MCAF ESTIMATES

Item Dependent Variables

Step	Entry	Survey $Q_{i=1,w=1}$ Record $V_{i=1}$	$Q_{i=1,w=2}$ $V_{i=1}$	Q_{i1} V_i	Q_{i2} V_i	ETC.
1	R^2 (Q_{iw}, X)					
2	R^2 (V_i, X)					
3	Step 1/ Step 2					
4	$\sqrt{\text{Step 3}}$					
5	r (Q_{iw}, Vi)					
6	Step 4 x Step 5					
7	$\sqrt{\text{Step 6}}$					
8	Step 6 ÷ Step 3					
9	$\sqrt{\text{Step 8}}$					

Proceed one column at a time:

STEP 1: Enter R^2 from the regression $Q_{iw} = a + BX + e$

Where Q is a questionnaire item, i indexes the subject matter of the item (burglary, robbery, etc.) and w indexes whether this is an arrest incident or arrest-for-crime-done item (w=1,2)
X is a vector of predictors that may change across i's (e.g., demographic characteristics, criminal history variables)
a is an intercept, B are the regression coefficients, and e is a stochastic error term.

STEP 2: Enter R^2 from the regression $V_i = a + BX + e$ where V_i is the record item corresponding to Q_i.

STEP 5: Enter the correlation of the questionnaire and record item scores.

NOTE: Estimates for Steps 1, 2, and 5 must be based on the same respondent samples. Seek statistical advice about missing data treatment options.

STEP 6: Is an estimate of $MCAF^2_{Qij}$

STEP 9: Is an estimate of $MCAF_{Vi}$

If the sampling distribution of $MCAF_{Vi}$ is known, it is possible to test the validity of the assumptions underlying the estimates by noting whether the two estimates of $MCAF_{Vi}$ are the same.

REFERENCES

Amsel, Z., W. Mandell, L. Matthias, C. Mason, and K. Hockerman, "Reliability and Validity of Self-Reported Illegal Activities and Drug Use Collected from Narcotic Addicts," International Journal of the Addictions, Vol.11, 1976, pp. 325-336.

Bale, R. N., "The Validity and Reliability of Self-Reported Data from Heroin Addicts: Mailed Questionnaires Compared with Face-to-Face Interviews," International Journal of the Addictions, Vol.14, 1979, pp. 993-1000.

Ball, J. C., "The Reliability and Validity of Interview Data Obtained from 59 Narcotic Drug Addicts," American Journal of Sociology, Vol.72, 1967, pp. 650-654.

Belson, W. A., "The Extent of Stealing by London Boys and Some of Its Origins," Advancement of Science, Vol.25, 1968, pp. 171-184.

Blackmore, J., "The Relationship Between Self-Reported Delinquency and Official Convictions Amongst Adolescent Boys," British Journal of Criminology, Vol.14, 1974, pp. 172-176.

Blalock, H. M., C. S. Wells, and L. F. Carter, "Statistical Estimation with Random Measurement Error," in E. F. Borgatta and G. W. Bohrnstedt (eds.), Sociological Methodology, 1970, Jossey-Bass, San Francisco, 1970.

Bohrnstedt, G. W., and T. M. Carter, "Robustness in Regression Analysis," in H. L. Carter (ed.), Sociological Methodology, 1971, Jossey-Bass, San Francisco, 1971.

Bridges, G. S., "Levels and Effects of Response Error in Self-Reports of Crime and Delinquency," Ph.D. Dissertation, University of Pennsylvania, 1979.

Cochran, W. G., "Errors of Measurement in Statistics," Technometrics, Vol.10, 1968, pp. 637-666.

Cochran, W. G., "Some Effects of Errors of Measurement on Multiple Correlations," Journal of the American Statistical Association, Vol.65, 1970, pp. 22-34.

Elliott, D. S., and H. L. Voss, Delinquency and Dropout, Lexington Books, Lexington, Massachusetts, 1974.

Erickson, M. L., and L. T. Empey, "Court Records, Undetected Delinquency, and Decision-Making," in D. R. Cressey and D. A. Ward (eds.), Delinquency, Crime, and Social Process, Harper and Row, New York, 1969.

Farrington, D. P., "Self-Reports of Deviant Behavior: Predictive and Stable?", Journal of Criminal Law and Criminology, Vol.64, 1973, pp. 99-110.

Gibson, H. B., S. Morrison, and D. J. West, "The Confession of Known Offenses in Response to a Self-Reported Delinquency Schedule," British Journal of Criminology, Vol.10, 1970, pp. 277-280.

Gould, L. C., "Who Defines Delinquency: A Comparison of Self-Reported and Officially-Reported Indices of Delinquency for Three Racial Groups," Journal of Social Problems, Vol.16, 1969, pp. 325-336.

Guilford, J. P., Psychometric Methods, McGraw-Hill, New York, 1954.

Guze, S. B., V. B. Tuason, M. A. Stewart, and B. Picken, "The Drinking History: A Comparison of Reports by Subjects and Their Relatives," Quarterly Journal of Studies on Alcohol, Vol.24, 1963, pp. 249-260.

Hackler, J. C., and M. Lautt, "Systematic Bias in Measuring Self-Reported Delinquency," Canadian Review of Sociology and Anthropology, Vol.6, 1969, pp. 92-106.

Hansen, Morris H., W. N. Hurwitz, and M. A. Bershad, "Measurement Errors in Censuses and Surveys," Bulletin of the International Statistical Institute, Vol.38, 1961, pp. 359-374.

Hardt, R. H., and S. Peterson-Hardt, "On Determining the Quality of the Delinquency Self-Report Method," Journal of Research in Crime and Delinquency, Vol.14, 1977, pp. 247-261.

Henry, Neil W., "Measurement Models for Continuous and Discrete Variables," in A. S. Goldberger and O. D. Duncan (eds.), Structural Equation Models in the Social Sciences, Academic Press, New York, 1973.

Hirschi, T., Causes of Delinquency, University of California Press, Berkeley, 1969.

Jessor, R., T. D. Graves, R. C. Hanson, and S. L. Jessor, "The Measurement of Drinking Behavior," Chapter 7 in Jessor et al., Society, Personality, and Deviant Behavior, Robert Krieger Publishing Company, Huntington, New York, 1975.

Joreskog, K. C., and A. S. Goldberger, "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable," Journal of the American Statistical Association, Vol.70, 1975, pp. 631-639.

Knupfer, Genevieve, "The Validity of Survey Data on Drinking Problems: A Comparison Between Respondents' Self Reports and Outside Sources of Information," in G. Knupfer and W. J. Clark (eds.), Drinking Patterns and Problems: A Decade of Studies in the San Francisco Area, Rutgers

Center of Alcohol Studies (forthcoming).

Kulik, J. A., K. B. Stein, and T. R. Sarbin, "Disclosure of Delinquent Behavior Under Conditions of Anonymity and Nonanonymity," Journal of Consulting and Clinical Psychology, Vol.32, 1968, pp. 506-509.

Locander, W., S. Sudman, and N. M. Bradburn, "An Investigation of Interview Method, Threat and Response Distortion," Journal of the American Statistical Association, Vol.71, 1976, pp. 269-275.

Marquis, K. H., The Methodology Used to Measure Health Care Consumption During the First Year of the Health Insurance Experiment, The Rand Corporation, R-2126-HEW, August 1977.

Marquis, K. H., Record Check Validity of Survey Responses: A Reassessment of Bias in Reports of Hospitalizations, The Rand Corporation, R-2319-HEW, May 1978.

Marquis, K. H., et al., Response Errors and Solution Options for Surveys of Sensitive Topics, The Rand Corporation, forthcoming in 1981.

Marquis, M. S., and K. H. Marquis, Survey Measurement Design and Evaluation Using Reliability Theory, The Rand Corporation, R-2288-HEW, June 1977.

Nye, F. I., and J. F. Short, Jr., "Scaling Delinquent Behavior," American Sociological Review, Vol.22, 1957, pp. 326-331.

Petersilia, J., "The Validity of Criminality Data Derived from Personal Interviews," The Rand Corporation, P-5890, July 1977.

Sobell, L. C., and M. B. Sobell, "Outpatient Alcoholics Give Valid Self-Reports," Journal of Nervous and Mental Disease, Vol.161, 1975, pp. 32-42.

Sobell, L. C. and M. B. Sobell, "Validity of Self-Reports in 3 Populations of Alcoholics," Journal of Consulting and Clinical Psychology, Vol.46, 1978, pp. 901-907.

Sobell, L. C., The Validity of Self-Reports: Toward a Predictive Model, Ph.D. Dissertation, University of California at Irvine, 1976.

Sobell, M. B., L. C. Sobell, and F. H. Samuels, "Validity of Self-Reports of Alcohol-Related Arrests by Alcoholics," Quarterly Journal of Studies on Alcohol, Vol.35, 1974, pp. 276-280.

Sobell, M. B., L. C. Sobell, and R. Vander Spek, "Relationship Among Clinical Judgment, Self Report, and Breath Analysis Measures of Intoxication in Alcoholics," Journal of Consulting and Clinical Psychology, Vol.47, 1979, pp. 204-206.

Stephens, R., "The Truthfulness of Addict Responses in Research Projects," International Journal of the Addictions, Vol.1, 1972, pp. 549-558.

Voss, H. L., "Ethnic Differentials in Delinquency in Honolulu," Journal of Criminal Law, Criminology, and Police Science, Vol.54, 1963, pp. 322-327.

Walsh, W. B., "Validity of Self-Report," Journal of Counseling Psychology, Vol.14, 1967, pp. 18-23.

Williams, J. R., and M. Gold, "From Delinquent Behavior to Official Delinquency," Social Problems, Vol.20, 1972, pp. 209-229.

Wyner, G., "Response Errors in Self-Reports of Arrest Histories," MPR Working Paper C-10, Mathematica Policy Research, Princeton, New Jersey, September 1976.

END