

EVALUATION OF THE REDESIGNED QUESTIONNAIRE FOR THE CHILDREN IN CUSTODY CENSUS

Yukiko Ellis and Laurie Schwede, U.S. Bureau of the Census
Yukiko Ellis, U.S. Bureau of the Census, Washington, D.C. 20233

Key Words: *Questionnaire redesign,
Split-panel test, Response rate*

1. INTRODUCTION

Since 1994, the Census Bureau has conducted the Children in Custody Questionnaire Redesign Project for the Office of Juvenile Justice and Delinquency Prevention (OJJDP). Its overall objective is to develop a new questionnaire (or "form") that collects individual-level data on juveniles placed in facilities for offenses. The prior (or "old") form collected the aggregate-level data in the Census of Juvenile Detention, Correctional, and Shelter Facilities.

This paper reports findings from a split-panel test in the final phase of the three-phase project based on a survey of 480 sample facilities. The goals of the split-panel test are to: 1) compare the overall unit response rate of the new form to that of the old form, 2) identify problems with the new form and recommend improvements for the final instrument, and 3) identify the types of facilities that might benefit from assistance by OJJDP to improve the likelihood of success in the actual 1997 census.

2. BACKGROUND

The OJJDP requested the Center for Survey Methods Research (CSMR) at the Census Bureau to redesign the Census of Juvenile Detention, Correctional, and Shelter Facilities. This census, commonly known as the Children in Custody Census (CIC), has been conducted biennially since 1977 by the Governments Division at the Census Bureau for the OJJDP. The overall goal set by the OJJDP for the project was to design a census form that collected individual-level data to provide data users with more flexibility in data analysis than the aggregate format of the old form.

In Phase 1 of the project, CSMR staff conducted exploratory research with respondents at juvenile facilities (Schwede and Ott, 1995). In Phase 2 they developed and tested the first prototype of the roster questionnaire, "Census of Juveniles in Residential Placement" (Schwede and Moyer, 1996).

In the first part of Phase 3, the CSMR conducted cognitive interviews with the second version of the new roster questionnaire (Schwede and Gallagher, 1996). The questionnaire was further refined and the split-panel test was conducted with the prior unrevised form and the latest version of the new form (Ellis and Schwede,

1997).

The prior CIC form consists of thirteen sections on fourteen pages with small print. Topics include: facility type and characteristics, aggregate-level questions on juveniles by gender, age, race, Hispanic origin, types of offenses, and type of placement. Other topics on the old form which were not included on the redesigned roster form include: juveniles' length of stay; educational, treatment, and medical programs; expenditures; and number of juvenile deaths.

The new form consists of three sections on eleven pages with larger print. Every facility respondent fills in Sections I and III (five pages altogether) that ask for the facility-level information. Section II is filled in only by respondents whose facilities have any youth under age 21 placed there because of an adult offense or a non-adult offense. For each youth with an offense, the respondent answers eleven questions in Section II: ID, sex, date of birth, race, placement agent, placement agent level, most serious offense resulting in placement here, county and state of offense, adjudication status, admission date, and security level.

One main feature of the new form was the introduction of the expressions, young persons with "adult offenses" and "non-adult offenses." These expressions replaced the terms, respectively, young persons with "delinquent offenses" and "status offenses," which were observed in the earlier cognitive testings to be interpreted in varying manners by respondents.

One main function of Section I in the new form is for respondents to provide counts of juveniles in three groups defined by offense category. Our intention was to use this information in our evaluation of the roster in Section II as a built-in coverage test. We were concerned before the split-panel test that some respondents, especially those reporting for large facilities, might get tired of listing all juveniles in Section II and submit incomplete lists of juveniles.

3. METHODOLOGY

The universe population for the split-panel test consisted of juvenile facilities in existence in the U.S. in 1996 with single reporters (as opposed to central reporters who complete forms for more than one facility). The sampling frame was based on the last census of juvenile facilities conducted in 1995 with an update in 1996.

Stratified random sampling was used with juvenile facilities as sampling units. The variables used in stratification were public/private management and size of facilities.

Since the OJJDP's main interest in this phase of the project was to find out how well the new survey instrument worked, we allocated the sample of 480 facilities with a 4:1 ratio of the new questionnaire to the old questionnaire within each unique combination of management and facility size categories. The old questionnaire was to be used only in the comparison of unit response rates with the new questionnaire. Weight was defined as the inverse of the selection probability in one-stage stratified random sampling.

The analyses in this report are based on data from the unedited databases. However, for the analysis of the "most serious offense" variable in Section II, both edited and unedited data are used. Table 1 shows the number of facilities in the universe and in the sample by new/old form, public/private facility, and by facility size.

To examine the unit nonresponse pattern, we fitted logistic regression models to the facility-level data, using the variables, public/private facility, facility size, and form (old/new), as the explanatory variables. The final model with the best fit includes the main effects and two interaction terms, one between public/private type and form, the other between form and facility size. Based on this model, we computed the odds ratio of the unit nonresponse among facilities receiving old and new questionnaires.

In analysis of the item nonresponse for the individual-level data in Section II of the new questionnaire, we examined the "most serious offense" code separately from other item questions because of its importance in the questionnaire and its peculiar nature of nonresponse. Nonresponse for this variable could mean either that a respondent did not answer the

question or that the respondent coded the item differently from the way it was defined in the questionnaire. In the former case, the respondent was contacted again during the telephone non-response follow-up (NRFU) operation. In the latter case, the item was left blank in the unedited file but an attempt was made later to translate the respondent's code into a valid code for the edited file. We thus compared the code in the unedited and edited files to gain some insight on the quality of data we can expect from the census on this very important question item.

4. RESULTS

4.1 Unit Response Rate

Unit response rate was defined as the proportion of questionnaires mailed back to the Census Bureau. Table 2 summarizes the comparison of response rates between old and new questionnaires.

Table 2. Estimated Unit Response Rate (With standard error in parentheses)

	Old Q	New Q
Overall*	63.2% (4.9%)	81.0% (2.1%)
By Facility Size		
Small	72.7% (8.1%)	75.8% (3.6%)
Medium*	51.6% (7.5%)	82.9% (2.8%)
Large*	64.5% (8.0%)	84.7% (2.7%)
By Management		
Public	83.9% (6.7%)	85.3% (2.8%)
Private*	53.1% (6.6%)	78.9% (2.8%)

* The difference is statistically significant at 10% significance level.

Table 1. Number of Single Reporter Facilities in the Universe and Sample by New/Old Questionnaire, Public/Private Facility, and Facility Size

Facility Size (# of residents in the facility)	Universe		Sample						
	Public	Private	Public		Private		Total		Total
			Old Q	New Q	Old Q	New Q	Old Q	New Q	
Small (<15)	204	667	11	44	22	88	33	132	165
Medium (15-36)	187	286	9	41	22	88	31	129	160
Large (37+)	193	197	11	44	20	80	31	124	155
Total	584	1150	31	129	64	256	95	385	480

Table 3 presents the estimated odds ratios that were derived from the results of fitting the logistic regression model to the facility-level data, using the unit nonresponse as the response variable. Among medium private facilities, respondents with the old questionnaire were 5.7 times more likely not to respond to the census than those with the new questionnaire. The odds ratio was 4.3 among respondents of large private facilities.

Table 3. Estimated Odd Ratios and 90% Confidence Interval of Unit Nonresponse For Old Form Compared to New Form, Controlling for Public/Private Type and Facility Size

Facility Size	Public	Private
Small	0.5 (0.4, 0.7)	1.4 (1.2, 1.7)
Medium	1.9 (1.3, 2.6)	5.7 (4.3, 7.5)
Large	1.4 (1.0, 2.1)	4.3 (3.1, 6.0)

4.2 Item Nonresponse

Facility-Level Questions in Sections I and III

Of the 385 facilities receiving the new questionnaires, 312 respondents returned the questionnaires. These 312 respondents are the subjects of the analysis in this section.

Table 4 summarizes the item NR rates for critical items in Section I (denoted as Q.I henceforth) in the questionnaire. Q.I.8-12 have two question parts. Part 'a' asks about the presence or absence and part 'b' asks for the number of juveniles fitting the respective description if the answer to part 'a' is yes.

Table 4. Estimated Item Nonresponse Rate for Facility-Level Questions in Section I (Standard error is given in parentheses.)

Item No	Description	Item NR
Q.I.3	Type of facility	1.6% (0.5%)
Q.I.8a	Residents, 21+ years old	1.4% (0.7%)
Q.I.9a	Residents, <21 years old	2.0% (0.6%)
Q.I.10a	Any with adult offenses (yes/no)	3.9% (0.9%)
Q.I.11a	Any with non-adult offenses (yes/no)	7.0% (1.4%)
Q.I.12a	Any with no offenses (yes/no)	7.1% (1.4%)

When we examined the joint distribution of the four questions, Q.I.9a - 12a, we observed that, of the 312 respondents,

270 respondents (87%) answered all four questions;
42 respondents missed at least one question:

Among the 42,

10 missed Question 11a only;

11 missed Question 12a only;

7 missed Questions 10-12 altogether; and

4 missed Question 9 only.

Individual-Level Questions in Section II

Table 5 summarizes the item NR rates for person-level questions in Section II.

Table 5. Estimated Item Nonresponse Rate for Person-Level Questions in Sections II (Standard error is given in parentheses.)

Item No	Description	Item NR
Q.II.1	ID	0.0%
Q.II.2	Sex	0.1% (0.07%)
Q.II.3	Date of birth	2.0% (0.9%)
Q.II.4	Race	0.8% (0.4%)
Q.II.5	Placement agent	0.1% (0.06%)
Q.II.6	Placement level	0.9% (0.4%)
Q.II.7	Most serious offense	4.1% (1.9%)
Q.II.8	County of offense	1.6% (0.8%)
Q.II.8	State of offense	2.2% (0.9%)
Q.II.9	Adjudication Status	1.7% (0.8%)
Q.II.10	Admission Date	1.7% (0.8%)
Q.II.11	N of locked doors	6.4% (2.6%)

Table 6 summarizes the result of linking the facility-level data in Section I to the person-level data in Section II. The table presents an overall picture of how well respondents answered questions Q.I.9-12 as well as how well they understood whom to roster in Section II.

Consistency between the Aggregate Counts in Section with the Roster Count in Section II

One way we examined consistency between Sections I and II was to compare the sum of the aggregate counts of juveniles with adult offenses in Q.I.10b and of those with non-adult offenses in Q.I.11b with the roster count

Table 6. Result of Linking Data from Sections I and III with Data from Sections II

Stratum	# of allocated facilities (a)	# of responding facilities (b)	# of facilities with no youths or having only non-offenders (from Sect I) (c)	# of facilities erroneously not answering Section II (d)	# of facilities linked to Section II data (b)-(c)-(d) = (e)	Proportion of linked facilities among allocated facilities (e)/(a)% = (f)
Pub/Small	44	35	1	1	33	75%
Pub/Medium	41	36	1	0	35	85%
Pub/Large	44	39	0	1	38	86%
Priv/Small	88	65	16	3	46	52%
Priv/Medium	88	71	16	1	54	61%
Priv/Large	80	66	18	4	44	55%
Total	385	312	52^{1/}	10^{2/}	250^{3/}	65%

1/ Of the 52 facilities, 7 facilities had no juveniles under 21 and 45 facilities had no juveniles with offenses.

2/ The 10 facilities reported 72 juveniles with adult offenses and 103 juveniles with non-adult offenses in Section I.

3/ Of the 250 facilities, 2 facilities reported having no juveniles under 21 in Q.I.9a and 7 facilities had missing data in Q.I.9a. However, they all went on to answer Section II, adding 209 juveniles with adult offenses and 24 juveniles with non-adult offenses.

of juveniles with the most serious offense code of 10-99 in Q.II.7. That is, the total numbers of juveniles with adult and non-adult offenses in Section I should equal the number of juveniles listed in Section II. This option avoided the problem of our inability to incorporate juveniles with “unknown offense” (code 99 in Q.II.7) in our comparisons when we examined Q.I.10b and 11b separately. However, we still had a problem of how to handle juveniles with missing code in Q.II.7.

Table 7 summarizes the comparison of the sum of

Q.I.10b and 11b with the roster count in Section II. ‘Group 1: Q.I = Q.II’ indicates that the aggregate total of juveniles with offenses in Section I agreed with the roster count of juveniles with offenses in Section II. ‘Group 2: Q.I > Q.II’ indicates that the aggregate total from Section I was greater than the roster count in Section II. ‘Agree with missing code’ indicates that the aggregate total from Section I agreed with the roster count when we included juveniles with missing value in the roster count. ‘Group 3: Q.II > Q.I’ indicates that the

Table 7. Comparison of the Aggregate Total (Q.I.10b & 11b) in Section I With the Roster Count in Q.II.7: Estimated Proportion with the standard error in parentheses

Comparison	Overall	By Public / Private		By Facility Size		
		Public	Private	Small	Medium	Large
Group 1: Q.I = Q.II	60.0% (3.2%)	54.2% (4.2%)	64.4% (4.3%)	68.7% (5.4%)	62.7% (4.7%)	40.8% (4.7%)
Group 2: Q.I (aggregate total) > Q.II (roster count)						
Agree with missing code	4.7% (1.2%)	4.7% (2.0%)	4.7% (1.6%)	4.3% (1.9%)	2.9% (1.7%)	7.7% (2.7%)
Disagree	17.1% (2.6%)	13.1% (3.6%)	20.0% (3.0%)	15.8% (4.1%)	15.7% (3.5%)	21.0% (4.4%)
Group 3: Q.II (roster count) > Q.I (aggregate total)						
1.10b & 11b missed	11.1% (1.9%)	15.9% (2.3%)	7.5% (3.1%)	7.3% (3.5%)	11.3% (3.1%)	17.7% (2.9%)
Disagree	7.1% (1.5%)	12.1% (2.7%)	3.4% (1.8%)	3.9% (2.3%)	7.4% (2.4%)	12.8% (3.0%)
Total	100%	100%	100%	100%	100%	100%

roster count in Section II was greater than the aggregate total from Section I. The subgroup, '10b and 11b missed,' includes facilities which missed both questions, Q.I.10b and 11b.

5. DISCUSSION

In the split-panel test, the results on the performance of the new questionnaire in terms of overall response rates exceeded our expectations. The overall response rate for the old questionnaire was 63.2%, compared to 81.0% for the new questionnaire. The difference of 17.8% is statistically significant at the 10% significance level. Furthermore, the response rate improved significantly for private facilities from 53.1% with the old questionnaire to 78.9% with the new questionnaire. Based on these results, we conclude that the new individual-level data collection method will work in OJJDP's census of this juvenile facility universe.

The analysis of Section I showed that the questions with the highest item nonresponse rates were those on the presence of juveniles placed in the facility for non-adult offenses and for non-offense reasons (Q.I.11a and 12a). Differentiating non-adult (status) offenders and non-offenders has been a chronic problem (for respondents) that we have identified and tried to clarify in previous phases of this research. Part of the difficulty lies in different state laws, and other parts to the question on whether "a status offender is a juvenile who is going through juvenile services and declared one, or somebody who is just doing the behavior" as one respondent from a shelter summed up in a cognitive interview (Schwede and Ott, 1995).

The analysis of Section II revealed that the overall quality of data on the person-level items was good. The items with the highest nonresponse rates were the questions on "number of locked doors" (Q.II.11) and "most serious offense" (Q.II.7). The item nonresponse rates on these two questions were 6.4% and 4.1%, respectively. A closer look on Q.II.11 identified clustering of nonresponses within a small subsample of large private facilities, rather than widespread problems in many facilities. The reason for the clustering might include problems with question wording. For Q.II.7, the overall item nonresponse rate was reduced from 4.1% in the unedited file to 0.5% in the edited file after item NRFU and recoding of some respondents' codes.

One goal in designing the new questionnaire was to place built-in coverage tests in Section I to assess how well respondents would provide individual-level information in Section II. Our coverage tests consisted of comparing the consistency of Section I aggregate-level counts by offender status and individual-level counts of specific offenses in Section II: the total numbers of juveniles with adult and non-adult offenses

in Section I should equal the number of juveniles listed in Section II.

The findings indicated that 10 of the 260 facility respondents reporting juveniles with offenses in Section I failed to provide any individual-level data in Section II. These 10 respondents reported 175 juveniles in Section I who should have been rostered in Section I but were not. This finding that under 4% of the respondents failed to provide any individual-level data suggests that: 1) most respondents understood the directions as to when they should complete Section II; and 2) they were willing to provide individual-level information.

Further analysis revealed that Section I had a few problems as coverage tests. First, 42 out of 312 facility respondents who filled out Section I missed at least one aggregate question in Q.I.9a-12a in Section I. Half of the 42 respondents skipped either Q.I.11a on juveniles with non-adult offenses or Q.I.12a on juveniles with no offenses. On the other hand, there were 18 facility respondents who indicated they had no juveniles with adult-offenses (Q.I.10a) or non-adult offenses (Q.I.11a) in Section I, and yet, they went on to list more than 5000 juveniles with offenses in Section II. These patterns might indicate that some respondents had difficulty telling the three types of offense groups apart. For juveniles with adult offenses, a lot larger number of juveniles were rostered in Section II than were counted in the aggregate total in Section I possibly because respondents were more certain whom to include in the roster with the help of the list of specific offense codes provided in Section II. In contrast, for juveniles with non-adult offenses, a much lower number of juveniles were rostered in Section II than were counted in the aggregate total in Section I. One possible reason is that respondents might have thought they did not have to list them in Section II. For example, among the ten respondents who did not fill out Section II when they should have, five reported no juveniles with adult offenses but did report at least one juvenile with a non-adult offense in Section I. Another possible reason is that respondents may have loosely classified some juveniles as non-adult offenders in Section I who were later revealed to be non-offenders when assigned specific offense codes in Q.II.7.

The second problem with Section I was that 7 (out of 42) facility respondents missed Q.I.10-12 altogether while 4 missed Q.I.9 only. This might indicate that either they overlooked the importance of reporting the aggregate counts or they did not have these numbers available until they answered Section II. We observed this phenomenon most frequently among large public facilities, resulting in big discrepancies between the aggregate totals in Section I and their corresponding roster counts in Section II.

Although the new questionnaire is a big improvement over the prior questionnaire, there is still some room for improvement, especially on how we describe the three groups of juveniles by offense in Section I. Regardless of which set of terms we will be using in the actual census, we should try to simplify the respondents' task of grouping juveniles by offense category. One strategy is to present the list of offenses to respondents when we ask them for aggregate totals in Section I and not wait until respondents reach Section II as was done in the split-panel test. This list should have clear headings for each group of offenses. This recommendation is based on the observation that respondents, especially at large public facilities, seemed to know how to classify juveniles by offense better in Section II than in Section I probably because the list of offense codes was helpful to them.

The second recommendation is to add a built-in edit by adding a statement at the end of Section I, "The numbers you gave for question Q.I.10b, 11b, and 12b should add up to the number in question Q.I.9b." This edit might clarify some of the respondents' confusion about how the numbers in Q.I.9b-12b are related, and it has a great potential for reducing the number of item NRFU calls that the Census Bureau will have to make. Of course, the aggregate counts in Section I should be corroborated by the roster counts in Section II.

The third recommendation is to simplify question Q.I.10-12 by dropping first part ('a') of each of these questions. This would eliminate the pitfall of questions with confusing skip patterns.

Since it is not feasible to introduce the second and third recommendations in the 1997 CIC census, we recommend that these changes be incorporated into the pretest of the new Juvenile Facility Census. The juvenile facility census has the same universe as the 1997 CIC census and it has the same Section I questions. If the changes are found to improve data quality, we can incorporate the changes in the 1999 Census of Juveniles in Residential Placement.

6. CONCLUSION

The new questionnaire for the Children in Custody Census, in all likelihood, will be more successful than the old questionnaire in eliciting responses from juvenile facilities with single reporters. Small and private facilities will probably benefit most from the OJJDP's assistance to improve their unit nonresponse rates. The OJJDP might also want to concentrate on a few facilities, especially among large private facilities, to improve the data quality of person-level question items.

If the new questionnaire is used in the present form, we expect to find substantial discrepancies between the aggregate counts in Section I (where facility-level

questions are asked) and the roster counts in Section II (where person-level questions are asked) for juveniles with offenses. Section II in this survey reported a net total of 1345 more juveniles with offenses than Section I. One option is to eliminate the questions on the aggregate totals in Section I. However, we need Section I to define three groups of juveniles and identify eligible juveniles to be rostered in Section II. Another option is to provide respondents with as many tools as possible in Section I as in Section II to make it easier for them to provide accurate answers in Section I.

One such tool is a list of offenses with clear labels for respondents to use as a reference when they figure out the aggregate totals. Another device is to add an instruction at the end of Section I for respondents to check their aggregate totals.

We would hope that over time, as respondents gain experience with the new questionnaire, some of these problems will be resolved. We would also hope that some respondents might develop new record-keeping strategies to distinguish juveniles by both offense groups and specific offenses according to our codes in computer files. This would reduce the burden on them in completing our questionnaire and would be likely to yield more accurate and consistent data.

ACKNOWLEDGMENTS

The authors are grateful to Terry DeMaio, Carma Hogue, and Patricia Feindt for providing valuable comments and to Charlene Sebold and her staff for providing data.

REFERENCES

- Ellis, Y. and Schwede, L. (1997). "Evaluation of the Redesign Questionnaire for the Children in Custody Census." In print. Center for Survey Methods Research, U.S. Bureau of the Census.
- Schwede, L. and Gallagher, C. (1996). "Children in Custody Questionnaire Redesign Project: Results from Phase 3 Cognitive Testing of the Roster Questionnaire." Center for Survey Methods Research, U.S. Bureau of the Census.
- Schwede, L. and Moyer, L. (1996). "Children in Custody Questionnaire Redesign Project: Results from Phase 2 Questionnaire Development and Testing." Center for Survey Methods Research, U.S. Bureau of the Census.
- Schwede, L. and Ott, K. (1995). "Children in Custody Questionnaire Redesign Project: Results from Phase I." Center for Survey Methods Research, U.S. Bureau of the Census.