

# CONCEPTUAL AND METHODOLOGICAL ISSUES IN DEFINING THE BASIC STATISTICAL UNITS IN A FACILITY CENSUS

Laurie Schwede, U.S. Bureau of the Census<sup>1</sup>

Statistical Research Division, U.S. Census Bureau, Washington, D.C. 20233-9100

**Keywords:** Facility survey, Establishment survey, Business survey, Survey methods, Survey frame development

“What is a facility in a facility-level census?”, a respondent might ask. The answer to this question defining the target statistical unit in a census would seem straightforward and easy. According to the Random House College Dictionary (1980), a facility is defined as “something that is built, installed, or established to serve a specific function or perform a particular service.” This conjures up an image of one big building with one self-contained program, as shown by a simple rectangle in Diagram 1.

Diagram : Undifferentiated facility



1 Facility  
1 Program Type  
1 Building

Many facilities in the Census of Juvenile Detention, Correctional, and Shelter Facilities do fit this image. However, evidence from past edited questionnaires and from field interviews and observations in facilities indicate enormous variation in such factors as size, layout, programs and/or organizational complexity. In reality, it is neither easy nor straightforward to develop standardized definitions of “facility” and “facility type” to guide respondents to define and select the right reporting units.

The purpose of this paper is to present the results of research to improve the definitions and methods for delineating the basic statistical units in a census of residential juvenile facilities. The concepts defining the target statistical unit in this census--“facility” and “facility type”--will be discussed and the questions eliciting this information on the mailout form will be shown. Six factors affecting how respondents interpret these concepts and apply them to their own organizations to determine their reporting units will be identified. The interaction of these factors will be illustrated in general diagrams of facility layouts and arrangements. The implications of these interactions for the fit between the target statistical units intended in this census and the units respondents choose to report on will be discussed. The implications for frame coverage, data quality, data collection methodology, respondent burden and questionnaire revisions will be identified. Finally, the applicability of the findings to other group quarter and business surveys will be addressed.

## Background

This research to improve the delineation of target statistical units has been done as part of our multi-stage research to redesign the Census of Juvenile Detention, Correctional, and Shelter Facilities (also known as the Children in Custody, or CIC census) for the Office of Juvenile Justice and Delinquency Prevention in the Justice Department. This is a national, biennial census gathering data by mail on the characteristics of roughly 3,600 juvenile facilities and the young persons housed in them. Some respondents are single reporters, completing one self-administered form for one facility. Others are central reporters, completing forms for multiple facilities.

In this census, facility type has been determined primarily by the respondent’s answer to the facility type question on the census form. In 1994 when our research began, there was no definition of “facility” on the questionnaire. The question was worded as follows:

---

<sup>1</sup> The author thanks fellow project researchers Catherine Gallagher, Laurie Moyer, and Sharon Birch for their observations and insights and Lorraine Randall for creating the graphics in this paper. She thanks Sharon Birch, Laurie Moyer, Carma Hogue, Leslie Brownrigg, Joe Moone, Jenny Hess, Terry DeMaio, Elizabeth Martin, Nash Monsour, Linda Kehm, and Denise Smith for their comments. This paper reports the results of research and analysis undertaken by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. This report is released to inform interested parties of research and to encourage discussion.

**This facility is primarily a -**

Mark (X) the one box that best describes this facility.

- Detention center
- Shelter
- Reception or diagnostic center
- Training school
- Ranch, forestry camp, or farm
- Halfway house or group home

Just one choice was allowed; no "Other, specify" option was given. Due to problems both on previously edited forms and in exploratory interviews at 40 facilities, some of which were multi-purpose and very complex, we redesigned the question. New versions included more facility types, an "Other, specify" line, and a new mark-all-that-apply instruction. The most recent version is:

**What type of residential facility is the one listed on the front cover? If this is a multi-type facility, please mark (X) for those that apply.**

- Detention center
- Long-term secure facility
- Training school
- Runaway and homeless shelter
- Other type of shelter
- Reception or diagnostic center
- Halfway house
- Group home
- Boot camp
- Ranch, forestry camp, wilderness or marine program or farm
- Residential treatment center (with onsite mental health care)
- Residential treatment facility (without onsite mental health care)
- Substance abuse center
- Independent living
- Foster home
- Other, specify

Data for this paper come from three rounds of cognitive interviews and observational tours at 48 juvenile facilities in 17 states (documented in Schwede and Moyer 1996; Schwede and Gallagher 1997; Birch, Schwede, and Gallagher 1998). Project interviewers saw and heard how each respondent classified his/her facility during the cognitive interview and then unobtrusively observed the layout and organization during the tour. By comparing assessments of respondents and interviewers for each facility, we can examine the fit between the respondents' reporting units and types with the target units intended in this census.

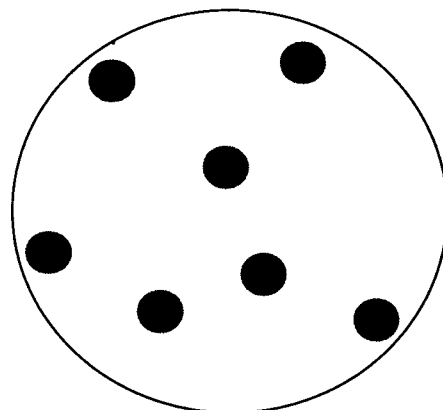
**Results**

In comparing the respondents' answers to the facility type question and our own observations, we identified six factors that appeared to affect 1) how respondents interpreted our request for facility-level information, and 2) how they applied our concepts of "facility" and "facility type" to decide which units to report on and how to classify them. These factors are

1. **The layout and geographical spread of the facility;**
2. **The extent and purpose of within-facility program differentiation;**
3. **The existence and types of linkages with other facilities and umbrella agencies;**
4. **The type of respondent (single or central reporter);**
5. **The specificity and accuracy of the pre-printed facility name and address on the form; and**
6. **The perceived response burden.**

The interaction of these factors can be illustrated in a series of simple diagrams that show the organization and layout of facilities we visited.

Diagram 2: Undifferentiated Facility Campus

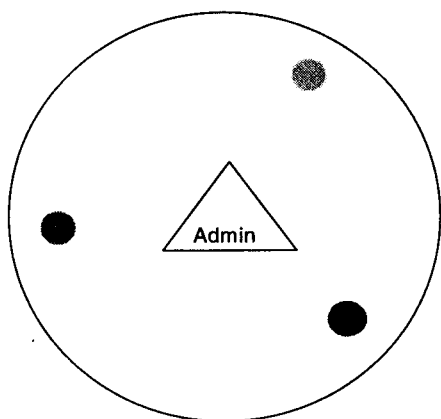


- 1 Facility
- 1 Program Type
- 7 Buildings
- 1 Campus

Diagram 1 (on the previous page) showed a rectangle symbolizing a simple, undifferentiated facility with one self-contained program in one building wholly used by the facility. Diagram 2 is functionally the same, but with the program spread over more than one campus building. In this and following diagrams, the large circle indicates the perimeter of a campus. The small circles embedded within the large circle represent residential buildings on the campus that are part of the same facility. Circles shaded the same denote units that are basically the same.

The boundaries for these undifferentiated facilities in Diagrams 1 and 2 are clear; these respondents are most likely to identify the correct target units.

Diagram 3: Internally Differentiated Facility



- 1 Facility
- 3 Subprograms providing separate, specialized treatment
- 4 Buildings
- 1 Campus

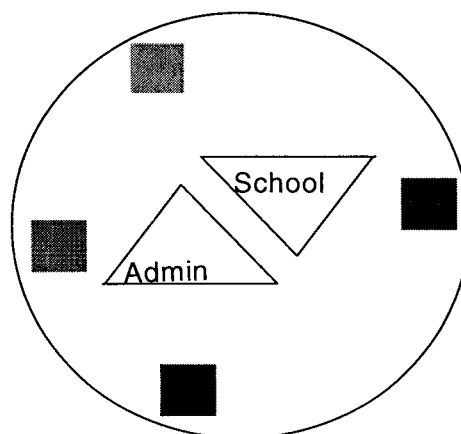
Diagram 3 shows an internally differentiated facility on one campus, with three residential buildings that are each relatively self-contained and demarcated in different shades. We found a number of facilities differentiated like this on the basis of such diverse factors as gender, age, security level, or sub-facility program type. A facility could be considered differentiated according to one factor, but undifferentiated on another. For example, a facility with three programs—a group home, a shelter, and a farm—is differentiated according to program, but undifferentiated in terms of the same low security level.

To maintain comparability across facilities, we picked one standard criterion to determine the existence and numbers of subdivisions in facilities in this census. This criterion is the existence of separate living units to provide specialized programs corresponding to mini-facility type categories. We coined the phrase “functional unit” to refer to these simple, within-facility programs. In Diagram 3, each functional unit is shown as being in a separate building, but this is not always the case: functional units may consist of more than one building designated for a special program. As the most basic, homogeneous unit, it is the lowest common denominator in differentiated facilities comparable to the undifferentiated facilities shown in Diagrams 1 and 2. The “functional unit” thus represents a possible second-level target statistical unit for this census. We are testing whether “functional units” can be identified by respondents in a mailout survey.

Many respondents in internally differentiated facilities seemed torn between their desire to identify their

subdivisions on the questionnaire and their resistance to the added work of completing forms for each subunit rather than one overall facility form. Clearly some respondents would choose to report only at the facility level, which could lead to data distortions. For example, in an internally differentiated facility with both a medium security detention center and a low security group home, how does a respondent provide one overall answer to security questions for the facility as a whole? Either the medium or low security answering pattern alone distorts the reporting of security levels for one of the functional unit subpopulations. The completion of separate forms for the detention center and the group home allows these different security levels within the facility to be recorded accurately, improving comparability with other facilities.

Diagram 4: Co-located Undifferentiated Facilities



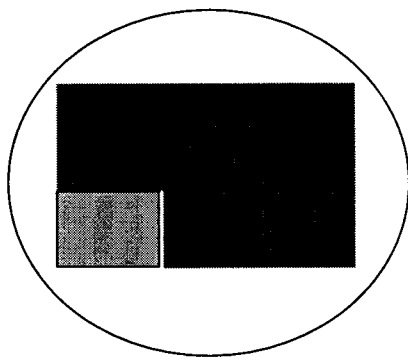
- 4 Facilities
- Dependent Linkages: sharing of resources
- 1 program in each (may be same or different)
- 6 Buildings
- 1 Campus

Diagram 4 shows one campus with several buildings depicted by *squares*, each of which is a separate *facility* of about equal status, rather than a subdivision within a facility, as in Diagram 3. Sometimes different facilities on a shared campus are self-contained and autonomous, sharing no resources. Often, however, the co-located facilities are interdependent. They may be linked through a common agency and/or through the sharing of some resources, such as administration and school buildings and staff, as shown. Representing these linkages on forms designed for stand-alone, self-contained, undifferentiated facilities is very difficult for respondents. For example, how should a central reporter completing forms for these four co-located facilities allocate the number of shared staff among the four facility forms? He/she could count all staff on just one form, count all staff on all four forms,

or try to divide the number of staff equally among the four forms. Depending on which option was chosen, the staff could be seriously overcounted or undercounted. Questions on the types of linkages among facilities and subprograms need to be developed so survey staff can identify co-located facilities at risk of these reporting errors and adjust to maintain reporting consistency.

To distinguish the separate co-located facilities in Diagram 4 from the internally differentiated facility in Diagram 3, it is necessary to identify differences in the purposes and populations of co-located units. Questions on the nature and extent of linkages, such as shared administration, ownership, staff, buildings, and budgets are also useful in this determination.

Diagram 5: Co-located Dominant/Dependent Facilities

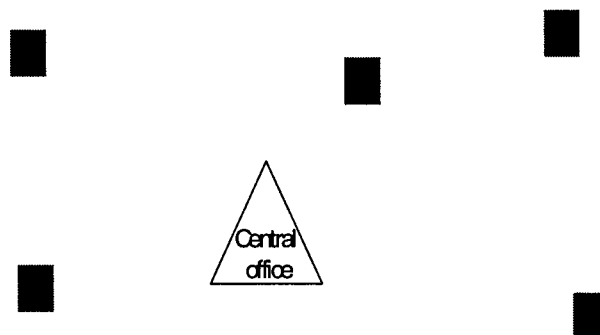


2 Facilities

Dependency Linkage: One dominant over the other through lease, subcontract, or other arrangement  
 1 program in each (may be same or different)  
 1 Building here  
 1 Campus (if there is more than one building)

Diagram 5 with rectangles represents 2 or more *facilities* co-located in a building or on a campus with linkages of dominance/dependency. The large, dominant facility may own the entire campus and contract or lease part of it to another agency for some other program (e.g., a state-level training school that donated the use of one wing to a county-run detention center). In the dominant facilities, some respondents switched reporting units as they worked through the questionnaire, giving population counts for the large facility only, but reporting capacity for both the small and large facilities. This gives a misleading picture of crowding. Such inconsistencies in the unit reported on *within* the dominant facility's questionnaire may not be detectable during edits. *Cross-facility* duplications can occur when respondents in both facilities are sent forms and both report the young persons in the small one. When the existence of these internal, dependent facilities is not recorded on the frame, the facilities and juveniles in them may be missed altogether, producing undercounts.

Diagram 6: Satellite Network of Undifferentiated Facilities



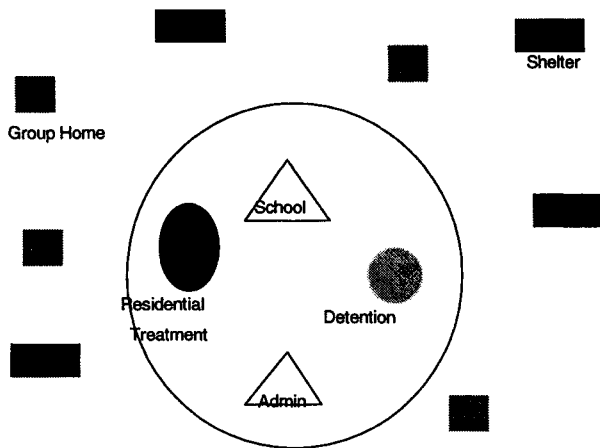
5 Facilities

Umbrella Linkage: umbrella agency network  
 1 Program (Same program type in all facilities)  
 6 buildings, including nonresidential office  
 Geographically dispersed  
 Central reporter or single reporter

Diagram 6 introduces the new configuration of small, undifferentiated facilities of the same type, such as group homes, that are linked in an umbrella agency network and dispersed around a city, region, or state. Often a central reporter at the agency's main office completes forms for all dispersed facilities. However, at least one agency-level central reporter aggregated data from all 9 dispersed undifferentiated facilities in his network onto just one summary questionnaire, rather than completing 9 forms. He attached a note giving the names and addresses of the facilities, inviting survey staff to divide his total population onto 9 forms if they wished to do so! This respondent deliberately changed the reporting unit from the facility to the agency to save himself the time and burden of completing 9 questionnaires. His aggregated data on the summary questionnaire give the false appearance of one large, centralized facility, rather than a network of 9 small, dispersed facilities. Using one summary form would also produce an undercount of group homes on the frame and in the database.

Hence one factor influencing the production of summary questionnaires is the respondent's perception and rejection of added burden. Another factor leading to the production of summary forms is the lack of frame information on the names and addresses of specific facilities within these dispersed agency networks. Some respondents surprised us by saying the names and addresses printed on their questionnaires from the frame were, in fact, the names and mailing addresses of their *umbrella agencies*, not those of specific facilities. If just the overall agency name and mailing address are printed on the form, a respondent would reasonably conclude one summary form for the whole network is what is wanted.

Diagram 7: Differentiated Network of Facilities



1 internally differentiated facility and  
 8 undifferentiated facilities  
 Linkages: Umbrella agency network  
 4 different programs  
 12 Buildings  
 One campus, with two dispersed networks  
 Central reporter or single reporter

Diagram 7 shows an amalgam of patterns already discussed: one campus with 2 internally differentiated programs as well as geographically dispersed networks of 2 types of undifferentiated satellite facilities. All units carry the same umbrella agency name and mailing address and are administered from a campus office. A central reporter is likely to report for all agency units, but may share the task with several reporters.

Internally and geographically complex arrangements like this are common for private agencies offering a continuum of services to juveniles. Some of these central reporters marked many facility types to show the range of agency programs, but then balked at filling out separate forms. A Diagram 7 complex agency respondent might complete: one overall summary form, one summary form for each of the four programs, or ten facility questionnaires. Clearly, ten forms would be more accurate, but summary forms would be less burdensome. Some agency respondents would submit just summary forms or nothing. Summary forms produce undercounts and classification errors of facilities and their populations.

Summary, Recommendations, and Applicability to Other Non-Household Surveys

In this paper I have illustrated a variety of facility and agency arrangements and identified problems that occur as respondents try to apply our target statistical unit

concepts of “facility” and “facility type” to identify their appropriate reporting units.

These examples show that it is neither easy nor straightforward to decide on the appropriate reporting units for complex facilities and agencies, especially on a mailout form where the respondent is basically on his/her own to decide what is intended from written instructions. Clearly, the facility type question alone is not sufficient for identifying the right target and reporting units in the complex facilities and agencies we observed in the field.

Several changes are needed. First, a definition of “facility” appropriate to this census should be included in the next version of the questionnaire:

“A juvenile residential facility is a place where young persons who have committed offenses may be housed overnight. A facility has living/sleeping units, such as one or more wings, floors, dorms, barracks or cottages on one campus or in one building. Any buildings with living/sleeping units that are not on the same campus should be considered separate facilities and should be recorded on separate questionnaires in this census.”

The second recommendation is to operationalize the six important factors identified in this paper into questions to define the appropriate units more accurately. Simple questions on facility layout, spread, the existence of subdivided programs and linkages have been added to the mailout questionnaire. The patterns of answers to these new questions and the facility type question will flag complex facilities and agencies in need of followup. However, in some cases, the final determination of the correct target units will be too complex and subtle to be done reliably with the data from the mail mode only.

The third recommendation is to add a new second-stage telephone followup for the flagged complex facilities. Expert survey staff would interview respondents to learn about the configurations of their facilities and agencies, decide on the appropriate statistical units for which data are available, and negotiate agreements with respondents on the units to be reported on. This process of working with respondents in organizations to identify the correct reporting and collection units and to set up reporting arrangements is called “*profiling*” (Colledge 1995).

The advantages in more clearly defining the target statistical unit of “facility” and adding a new statistical unit of “functional unit” include 1) greater frame accuracy, 2) coverage improvements, 3) better data quality, and 4) better comparability across facilities. However, the tradeoffs include 1) initial cost increases for the sponsor, 2) initial increases in workloads for survey staff to improve the accuracy of the frame, 3) increased burden for some respondents, and 4) possible increases

in nonresponse for some complex facilities.

This research was done in a census of residential juvenile facilities. The conceptual and methodological problems in defining the basic statistical units identified here are encountered in organizational surveys of businesses, farms, institutions, and non-institutional group quarters as well. As Cox and Chinnappa point out in *"The Unique Features of Business Surveys"* (1995:15):

The composition and complexity of the basic units [being surveyed], their distribution in the target population, and their volatility over time create difficult problems for all steps of survey design and implementation."

At the beginning of the paper I quoted the dictionary definition of a "facility" that a respondent might consult if he/she found no definition of "facility" on the questionnaire, as was the case with the original questionnaire in this census. At the end, I developed a definition of "facility" to use in the redesigned CIC census for the Justice Department. Is this new definition of "facility" consistent with definitions for the target statistical units used in other organizational surveys? The term, "facility," is not commonly used in other non-household data collections. In the decennial census the basic statistical unit for "special places" including residential facilities is the "group quarter." The "group quarter" may be defined as that part of *one* building other than a usual house, apartment or mobile home, in which persons live or stay for a particular purpose in one of the types of group quarters on the Census Bureau's "1998 Alphabetical Group Quarters Code List." The definition of "facility" given here differs from the "group quarter" in not being restricted to just one building, but rather including all residential buildings associated with the facility on one campus. The group quarters code list for juvenile institutions classifies facilities on the basis of length of stay and type of juvenile, rather than the facility types used in the redesigned CIC census, shown earlier.

In economic surveys, the basic statistical units are "establishment" and "enterprise." The North American Industry Classification system (NAICS) categorizes all establishments in terms of the activities in which they are primarily engaged. In this system, the establishment is

"generally a single physical location where business is conducted or where services or industrial operations are performed"... "The establishment as a statistical unit is defined as the smallest operating entity for which records provide information on the cost of resources (material, labor, and capital) employed to produce the unit of output." (Executive Office of the President 1997: 16).

The enterprise, or company, may be comprised of more than one establishment "Facility," as defined here, is consistent with "establishment" in terms of the services offered at one location, without specifying the number of buildings, and allowing for the specification of subdivisions within these units. However, the NAICS system distinguishes just three categories of residential facilities used in this juvenile facility census—temporary shelters, residential mental health and substance abuse facilities, and other residential care facilities—and does not have codes to separate juvenile from adult facilities.

Thus the target statistical unit of "facility" does overlap with "group quarter" and "establishment" in some circumstances, but varies in others. In definition, it is closer to "establishment" than to "group quarter." In terms of classification criteria, "facility" is again closer to the activities breakdown in the NAICS coding scheme, but provides a more detailed list of facility types that are specifically limited to facilities that house juveniles.

The six factors identified here—layout and geographical spread, extent and purpose of internal differentiation, existence and types of linkages with other entities, type of reporter, completeness of frame information on relevant subdivisions, and perceived response burden—and their interactions, as shown in the diagrams, are also applicable to surveys of group quarters and establishments.

#### REFERENCES CITED:

- Birch, S., Schwede, L., and Gallagher, C. 1998. "Juvenile Residential Facility Census Questionnaire Redesign Project: Results from Phase 2 Cognitive Interview Testing." Washington DC, Census Bureau.
- Colledge, M. "Frames and Business Registers: An Overview," in *Business Survey Methods*, B. Cox, D. Binder, B. Chinnappa, P. Kott, eds. New York, John Wiley & Sons, Inc.
- Cox, B. and Chinnappa, B. 1995. "Unique Features of Business Surveys," in B. Cox, D. Binder, B. Chinnappa, P. Kott, eds. *Business Survey Methods*, New York, John Wiley & Sons, Inc.
- Executive Office of the President. 1997. North American Industry Classification System Manual, United States, 1997. Lanham, Maryland, Bernan Press.
- Schwede, L. and Gallagher, C. 1997. "Children in Custody Questionnaire Redesign Project: Results from Phase 3 Cognitive Testing of the Roster Questionnaire." Washington D.C., SRD, Census Bureau.
- Schwede, L. and Moyer, L. 1996. "Children in Custody Questionnaire Redesign Project: Results from Phase 2 Questionnaire Development and Testing." Washington D.C., SRD, Census Bureau.