

B. Collection of Information Employing Statistical Methods

This submission requests clearance for the 2008 National Postsecondary Student Aid Study (NPSAS:08), including a field test and the full-scale study. The sampling design for the full-scale NPSAS:08 study is presented in appendix E. The purpose of the NPSAS:08 field test is to fully test all procedures, methods, and systems of the study in a realistic operational environment prior to implementing them in the full-scale study. Specific plans for such field test activities are provided below.

1. Respondent Universe

a. Institution Universe

To be eligible for the NPSAS:08 field test, institutions are required during the 2006–07 academic year to:

- offer an educational program designed for persons who have completed secondary education; and
- offer at least one academic, occupational, or vocational program of study lasting at least 3 months or 300 clock hours; and
- offer courses that are open to more than the employees or members of the company or group (e.g., union) that administers the institution; and
- have a signed Title IV participation agreement with the U.S. Department of Education; and
- be located in the 50 states, the District of Columbia, or Puerto Rico; and
- be other than a U.S. Service Academy.

Institutions providing only avocational, recreational, or remedial courses or only in-house courses for their own employees are excluded. U.S. Service Academies are excluded because of their unique funding/tuition base.

b. Student Universe

The students eligible for inclusion in the sample for the NPSAS:08 field test are those who were enrolled in a NPSAS-eligible institution in any term or course of instruction at any time from July 1, 2006 through April 30, 2007 and who were:

- enrolled in either (a) an academic program; (b) at least one course for credit that could be applied toward fulfilling the requirements for an academic degree; or (c) an occupational or vocational program that required at least 3 months or 300 clock hours of instruction to receive a degree, certificate, or other formal award; and
- not currently enrolled in high school; and
- not enrolled solely in a GED or other high school completion program.

2. Statistical Methodology

a. Institution Sample

The institution sampling frame for the NPSAS:08 field test will be constructed from the 2005–06 Integrated Postsecondary Education Data System (IPEDS) header, Institutional Characteristics (IC), Fall Enrollment, and Completions files.

Three hundred institutions will be selected for the field test sample. We expect to obtain an overall eligibility rate of 98 percent and an overall participation (response) rate¹ of about 84 percent of institutions (based on the NPSAS:04 full-scale study). The eligibility and response rates will likely vary by institutional strata. Based on these expected rates, approximately 244 institutions will provide lists for selection of sample students. The estimated institution sample sizes and sample yield by the 22 institutional strata (described below) for the field test are presented in table 7.

The 300 field test sample institutions will be selected purposively from the complement of the institutions selected for the full-scale study. This ensures no institution will be in both the field test and full-scale samples without affecting the representativeness of the full-scale sample. Note that five strata have a field test sample size of zero because all institutions in these strata will be included in the full-scale sample. The full-scale study design is presented in appendix E.

¹ The institution response rate of 84 percent assumes that institutional participation will not be mandatory.

Table 7. NPSAS:08 expected field test estimated institution sample sizes and yield

Institutional stratum	Frame count ¹	Number sampled ²	Number eligible	List respondents
Total	6,610	300	294	244
Public				
Less-than-2-year	245	4	3	3
2-year	1,165	8	8	7
Total less-than-4-year	1,410	12	11	9
Bachelor's high education	20	4	4	3
Bachelor's low education	75	14	14	12
Master's high education	53	13	13	11
Master's low education	209	73	73	62
Total 4-year non-doctorate-granting	357	104	104	89
Doctorate-granting high education	27	0	0	0
Doctorate-granting low education	107	0	0	0
First-professional-granting high education	32	0	0	0
First-professional-granting low education	123	0	0	0
Total 4-year doctorate-granting	289	0	0	0
Private not-for-profit				
Less-than-2-year	99	2	2	2
2-year	222	2	2	2
Total less-than-4-year	321	4	4	4
Bachelor's high education	93	22	21	17
Bachelor's low education	370	51	49	40
Master's high education	110	31	30	24
Master's low education	437	31	30	24
Total 4-year non-doctorate-granting	1,010	135	130	107
Doctorate-granting high education	37	6	6	5
Doctorate-granting low education	146	16	16	12
First-professional-granting high education	82	0	0	0
First-professional-granting low education	324	11	11	9
Total 4-year doctorate-granting	589	33	33	26
Private for-profit				
Less-than-2-year	1,387	4	4	3
2-year or more	1,247	8	8	7
Total private for-profit	2,634	12	12	10

¹ Institution counts based on IPEDS:2003–04 header file.

² The field test sample size for a stratum will become zero if all institutions in the stratum are selected for the full-scale sample.

NOTE: Detail may not sum to totals because of rounding. NPSAS:08 = 2008 National Postsecondary Student Aid Study.

The nine sectors traditionally used for NPSAS analyses will be further broken down to form the same 22 strata used in NPSAS:2000 (the last NPSAS to spawn a B&B study) in order to break down 4-year institutions by degree type and percentage of students receiving education degrees, which is an important domain for the B&B longitudinal study. The strata are as follows:

1. public less-than-2-year;
2. public 2-year;
3. public 4-year non-doctorate-granting bachelor's high education;
4. public 4-year non-doctorate-granting bachelor's low education;
5. public 4-year non-doctorate-granting master's high education;
6. public 4-year non-doctorate-granting master's low education;
7. public 4-year doctorate-granting high education;
8. public 4-year doctorate-granting low education;
9. public 4-year first-professional-granting high education;
10. public 4-year first-professional-granting low education;
11. private not-for-profit less-than-2-year;
12. private not-for-profit 2-year;
13. private not-for-profit 4-year non-doctorate-granting bachelor's high education;
14. private not-for-profit 4-year non-doctorate-granting bachelor's low education;
15. private not-for-profit 4-year non-doctorate-granting master's high education;
16. private not-for-profit 4-year non-doctorate-granting master's low education;
17. private not-for-profit 4-year doctorate-granting high education;
18. private not-for-profit 4-year doctorate-granting low education;
19. private not-for-profit 4-year first-professional-granting high education;
20. private not-for-profit 4-year first-professional-granting low education;
21. private for-profit less-than-2-year; and
22. private for-profit 2-year or more.

b. Student Sample

The student sample sizes for the field test will be set to approximate the distribution planned for the full-scale study presented in appendix E. As shown in table 8, the field test is designed to sample approximately 3,000 students, including 2,089 baccalaureate recipients; 811 other undergraduate students; and 100 graduate and first-professional students. Based on past experience, we expect to obtain 92 percent eligibility rates and 70 percent student interview response rates, overall and within each sector. We also plan to employ a variable-based (rather than source-based) definition of study respondent, similar to that used in NPSAS:04 with revisions as deemed necessary by NCES. We expect the study response rate to be about 90 percent. We expect approximately 2,478 student survey respondents, including 1,746

baccalaureate recipients; 649 other undergraduate students; and 83 graduate and first-professional students.

Consistent with the procedures implemented in NPSAS:04, the field test student sample will be drawn from the first 150 of the approximately 244 participating institutions as the student enrollment lists are submitted. The remaining institutions will be sampled, but the sample will not be released unless needed. Sufficient numbers of students can be sampled from the first 150 enrollment lists to ensure proper testing of systems and procedures. However, the purpose of limiting the number of institutions from which the student sample is drawn is to ensure that the sample size for each institution is sufficient for CADE to be properly tested. There will be approximately 16.5 responding students per institution. In this way, a more accurate assessment of institutional burden (required for preparing and submitting student record data) can be made. Sampling from the institutions submitted after the first 150 (for which the sample won't be released) will allow us to refine procedures related to sampling from atypical lists and identifying potential baccalaureate recipients.

There will be seven student sampling strata:

1. potential baccalaureate recipients who are business majors;
2. potential baccalaureate recipients who are not business majors;
3. other undergraduate students;
4. masters students;
5. doctoral students;
6. other graduate students; and
7. first-professional students.

As was done in NPSAS:2000 and NPSAS:04, certain student types (potential baccalaureate recipients, other undergraduates, masters students, doctoral students, other graduate students, and first-professional students) will be sampled at different rates to control the sample allocation. Differential sampling rates facilitate obtaining the target sample sizes necessary to meet analytic objectives for defined domain estimates in the full-scale study.

To ensure a large enough sample for the Baccalaureate and Beyond (B&B) follow-up field test study, the base year sample includes a large percentage of potential baccalaureate recipients (see table 8). The NPSAS sampling rates for students identified by institutions as potential baccalaureates will be adjusted to yield the appropriate sample sizes after accounting for the percentage of students identified by institutions as potential baccalaureate recipients who actually receive a baccalaureate degree during the study year (about 87 percent, based on NPSAS:2000 data).²

Creating Student Sampling Frames. Several alternatives for the types of student enrollment lists that can be provided by the sample institutions are available. Our first preference

² In NPSAS:2000, the baccalaureate recipients were identified by separate lists usually sent close to the end of the spring term, so the 87 percent estimate may need to be adjusted downwards to help determine the appropriate field test sampling rates.

is to obtain an unduplicated list of all students enrolled in the specified time frame. However, lists by term of enrollment and/or by type of student (e.g., baccalaureate recipient, undergraduate, graduate, and first-professional) will be accepted. The student ID numbers can be used to easily unduplicate electronic files. If an institution has difficulty meeting these requirements, we will be flexible and select the student sample from whatever type of list(s) that the institution can provide, so long as it appears to accurately reflect enrollment during the specified terms of instruction. If necessary, we are even prepared to provide institutions with specifications to allow them to select their own sample.

In prior NPSAS studies that spun off a B&B cohort, lists of potential baccalaureate recipients were collected with the student list of all enrolled undergraduates and graduates/first professionals. However, these baccalaureate lists often were not received until late in the spring or in the summer, after baccalaureate recipients could be positively identified. To help facilitate earlier receipt of lists, we will request that the enrollment lists for 4-year institutions include an indicator of class level for undergraduates (1st year, 2nd year, 3rd year, 4th year, or 5th year). From NPSAS:2000, we estimate that about 55 percent of the 4th and 5th year students will be baccalaureate recipients during the NPSAS year, and about 7 percent of 3rd year students will also be baccalaureate recipients. To increase the likelihood of correctly identifying baccalaureate recipients, we will also request that the enrollment lists for 4-year institutions include an indicator (B&B flag) of students who have received or are expected to receive a baccalaureate degree during the NPSAS year (yes, no, don't know). We will instruct institutions to make this identification before spring graduation so as not to hold up the lists because of this requirement. These two indicators will be used instead of requesting a baccalaureate recipient list, and we plan to oversample 4th and 5th year undergraduates (seniors) and students with a B&B flag of "yes" to ensure obtaining sufficient yield of baccalaureate recipients for the B&B longitudinal study. We expect that most institutions will be able to provide undergraduate year for their students and a B&B flag. We will use whichever indicator seems to give the more accurate count of baccalaureates when compared to IPEDS. Our full-scale procedures will be revised based on field test results.

Since a disproportionately large proportion of baccalaureate recipients are business majors, we will also request major field of study and Classification of Instructional Programs (CIP) code on the lists to allow us to undersample business majors. A similar procedure was used effectively in NPSAS:2000 (the last NPSAS to include a B&B cohort). We expect that most institutions can and will provide the CIP codes.

The following data items will be requested for all NPSAS-eligible students enrolled at each sample institution:

- name;
- Date of birth (DOB)
- Social Security number (SSN);
- student ID number (if different from SSN);
- student level (undergraduate, masters, doctoral, other graduate, first-professional);
and

- locating information (local and permanent street address and phone number and school and home e-mail address).

As part of initial sampling activities, we will ask participating institutions to provide SSN and DOB for all students on their enrollment list³. We recognize the sensitivity of the requested information, and appreciate the argument that it should only be obtained for sample members. However, collecting this information for all enrolled students is critical to the success of the study for several reasons:

- It is possible that some minors will be included in the study population, so we will need to collect DOB to identify minors and obtain parental consent prior to data collection.
- The NPSAS:08 study includes a special analytic focus on a new federal grant (the National SMART grant) and SSN is needed to identify and over-sample recipients of this new grant.
- Having SSN will ensure the accuracy of the sample, because it is used as the unique student identification number by most institutions. We need to ensure that we get the right data records when collecting data from institutions for sampled students. It will also be used to unduplicate the sample for students who attend multiple institutions.
- Making one initial data request of institutions will minimize the burden required for participation (rather than obtaining one set of information for all enrolled students, and then later obtaining a set of information for sampled students.)
- An issue related to institutional burden is institutional participation. It is very likely that some institutions will respond to the first request, but not to the second. Refusal to provide SSNs after the sample members are selected will contribute dramatically to student-level nonresponse, because it will increase the rate of unlocatable students (see the following bullet).
- Obtaining SSN early will allow us to initiate locating and file matching procedures early enough to ensure that data collection can be completed within the allotted schedule. The data collection schedule would be significantly and negatively impacted if locating activities could not begin at the earliest stages of institutional contact.
- As part of a federally mandated study, NPSAS data are critical for informing policy and legislation, and are needed by Congress in a timely fashion. Thus, the data collection schedule is also critical. We must be able to identify the sample, locate students, and finish data collection and data processing quickly. This will not be possible within the allotted time frame if we are unable to initiate locating activities for sampled students once the sample has been selected.

The following section describes our planned procedures to securely obtain, store, and discard sensitive information collected for sampling purposes.

³ For institutions unwilling to provide SSN or location data for all students on enrollment lists, we will request SSN or locating data only for sample students immediately after the sample is selected.

Obtaining student enrollment lists. The student sample will be selected from the lists provided by the sampled institutions. To ensure the secure transmission of sensitive information, we will provide the following options to institutions: **1)** upload encrypted student enrollment list files to the project’s secure website using a login id and “strong” password provided by RTI, or **2)** provide an appropriately encrypted list file via e-mail (RTI will provide guidelines on encryption and creating “strong” passwords).

In past administrations of this study, hard copy lists were accepted via Fed-Ex or fax. We will no longer offer these options. We expect that a very few institutions will ask to provide a hard copy list (In NPSAS:04 full-scale study, 30 institutions submitted a hard-copy list – mostly via FedEx). In such cases, we will encourage one of the secure electronic methods of transmission. If that is not possible, we will accept a faxed list (but not a Fed-Ex list.) Although fax equipment and software does facilitate rapid transmission of information, this same equipment and software opens up the possibility that information could be misdirected or intercepted by individuals to whom access is not intended or authorized. To safeguard against this, as much as is practical, RTI protocol will only allow for lists to be faxed to a fax machine housed in a locked room and only if schools cannot use one of the other options. To ensure the fax transmission is sent to the appropriate destination, we will require a test run with non-sensitive data prior to submitting the actual list to eliminate errors in transmission from misdialing. RTI will provide schools with a FAX cover page that includes a confidentiality statement to use when transmitting individually identifiable information⁴. We will ensure that the hard copy list is shredded immediately after the sample has been selected. Immediately after the student sample is selected, RTI will ensure that the SSNs for non-selected students will be securely discarded (see description below).

Storage of enrollment files. For electronic lists, SSNs will be deleted from student records for all unsampled students immediately upon completion of sample selection.

- Encrypted electronic files sent via e-mail to a secure e-mail folder will only be accessible to a few staff members on the sampling team. These files will then be copied to a project folder that is only accessible to these same staff members. Access to this project folder will be set so that only those who have authorized access will be able to see the included files. The folder will not even be visible to those without access. After being copied, the files will be deleted from the e-mail folder. After selecting the sample of students for each school, the original file containing all students with SSNs will be immediately deleted. While in use, files will be stored on the network that is backed up regularly to avoid the need to recontact the institution to provide the list again should a loss occur. RTI’s information technology service (ITS) will use standard procedures for backing up data, so the backup files will exist for three months.
- Files uploaded to the secure NPSAS website will be copied from the NCES server to the same project folder mentioned above. After being moved, the files will be immediately deleted from the NCES server. After selecting the sample of students for each school, the original file containing all students with SSNs will be immediately deleted. As above, it is necessary for the files to be stored on the project share so that they can be backed up by ITS in case any problems occur that cause us to lose data.

⁴ These procedures are consistent with those endorsed by HIPAA. See <http://www.hipaadvisory.com/action/faxfacts.htm>

ITS will use their standard procedures for backing up data, so the backup files will exist for three months.

- Paper lists will be kept in one locked file cabinet in Regent Place. Only NPSAS sampling staff will have access to the file cabinet. The paper lists will be shredded immediately after the sample is selected, keyed, and QC'ed. The keying will be done by the same sampling staff who select the sample.

Table 8. NPSAS:08 expected field test student sample sizes and yield

Institutional sector	Sample students				Eligible students				Study respondents				Responding students per responding institution
	Total	Baccalaureates	Other undergraduate students	Graduate/first-professional students	Total	Baccalaureates	Other undergraduate students	Graduate/first-professional students	Total	Baccalaureates	Other undergraduate students	Graduate/first-professional students	
Total	3,000	2,089	811	100	2,761	1,933	735	93	2,478	1,746	649	83	16.5
Public less-than-2-year	40	0	40	0	31	0	31	0	25	0	25	0	12.6
Public 2-year	80	0	80	0	69	0	69	0	54	0	54	0	13.4
Public 4-year non-doctoral	1,040	716	277	47	969	667	258	44	850	585	226	38	16.3
Public 4-year doctoral	0	0	0	0	0	0	0	0	0	0	0	0	0.0
Private not-for-profit less-than-4-year	40	0	40	0	35	0	35	0	31	0	31	0	15.7
Private not-for-profit 4-year non-doctoral	1,350	1,123	210	17	1,244	1,035	193	16	1,144	952	178	14	17.1
Private not-for-profit 4-year doctoral	330	200	100	30	308	186	93	28	278	168	84	25	16.3
Private for-profit less-than-2-year	40	0	40	0	33	0	33	0	30	0	30	0	15.2
Private for-profit 2-year or more	80	50	24	6	72	45	22	5	66	41	20	5	16.4

Selection of Sample Students. The unduplicated number of students listed by each institution will be compared against the sampling frame (constructed from the 2005–06 IPEDS header, IPEDS-IC, Fall Enrollment, and Completions files) as a quality assurance check. Range checks required for acceptance of the student lists will be developed in consultation with NCES (e.g., agreement within 25 percent if the IPEDS counts are actual, rather than imputed, counts).

The student sampling procedures implemented in the field test will be as comparable as possible to those planned for the full-scale study, even though simpler procedures would suffice for the field test alone. For example, students will be sampled at fixed rates based on student sampling stratum and institutional stratum in the full-scale study, so students will be selected at fixed rates defined by institutional and student strata in the field test also. Sample yield will be monitored and the sampling rates will be adjusted, if necessary. This is to achieve the required field test sample sizes, just as they will be in the full-scale study.

Students will be sampled on a flow basis as student lists are received. Stratified systematic sampling procedures will be used. Lists will be unduplicated by student ID number prior to sample selection.

The final student sampling frames will be preserved as documentation of the sample selection task—on CD for electronic lists and in confidential permanent files for paper lists. More importantly, a master sample file, which will also serve as the student-level control file for the Receipt Control System (RCS), will be created and updated as each student sample is selected. The master sample file will contain at least the following data elements (in addition to receipt control variables):

- study ID number for the sample student (e.g., NPSASID);
- institution ID number (e.g., IPEDS UNITID);
- institution sampling stratum;
- student sampling stratum; and
- student selection probability.

Some of these data elements may not be necessary for the field test (e.g., student selection probability), but they will be generated so that the software developed for the field test will not need to be modified to produce this sample file for the full-scale study. Individually identifying information (e.g., student name and ID number at the sample institution) will be maintained separately to preserve confidentiality, and the study ID number will provide the linkage between these files.

In addition to selecting the student sample from enrollment lists received by participating institutions, a subsample of about 10 percent of interview respondents will be randomly selected to be re-interviewed⁵ to enable analysis of the reliability of items in the field test instrument. The Case Management System (CMS) will be programmed to randomly select this subsample. The subsampling rate will be set so that all re-interview cases are identified during the first half of field test data collection and all re-interviews can be completed at approximately the same time

⁵ Reinterviews will be conducted approximately 3 to 4 weeks after the initial interview, and will contain a subset of items (either new items or those that have been difficult to administer in the past). Reinterviews will be conducted in the same administration mode as the initial interview.

as the regular interviews, while allowing sufficient time to elapse between each initial interview and the re-interview.

3. Methods for Maximizing Response Rates

Response rates in the NPSAS:08 field test and full-scale study are a function of success in two basic activities: identifying and locating the sample members involved, then contacting them and gaining their cooperation. Two classes of respondents are involved: institutions, and students (undergraduate, graduate, and first-professionals) who were enrolled in those institutions.

a. Institution Contacting

The success of NPSAS:08 is closely tied to the active participation of selected institutions. Because institution contacting is the first stage of the study, upon which all other stages depend, obtaining the cooperation of as many institutions as possible is critical. The consent and cooperation of an institution's chief administrator is essential and helps to encourage the timely completion of the institutional tasks. Most chief administrators are aware of NPSAS and recognize the study's importance to postsecondary education. For those administrators who may believe that the study is overly burdensome, the first contact provides an opportunity to have a senior staff member address their concerns. At institutions newly selected for participation in NPSAS:08, the chief administrator contact provides an invaluable opportunity to establish rapport.

Proven Procedures. NPSAS:08 procedures will be developed from those used successfully in NPSAS:04. Initial institution contact information will be obtained from the IPEDS-IC file and used to telephone each institution (to verify data of record—e.g., the institution's name, address, and telephone number and the name and address of the chief administrator). Verification calls will begin in September 2006 and last approximately one week. Materials will be mailed to chief administrators in late September 2006, with follow-up calls continuing through early November. This schedule follows the model implemented in 2004 that established contact with the coordinator prior to the holiday season. The descriptive materials sent to chief administrators will be clear, concise, and informative about the purpose of the study and the nature of subsequent requests. The package of materials sent to chief administrators will contain:

- an introductory letter from the NCES Commissioner on U.S. Department of Education letterhead;
- a pamphlet describing NPSAS:08, including a study summary, outline of the data collection procedures, the project schedule, and details regarding the protection of respondent privacy and study confidentiality procedures; and
- a form confirming the institution's willingness to participate in the study, identifies an Institution Coordinator, and requests contact information for the chief administrator and the institution coordinator.

Follow-up calls to secure field test participation and name a study coordinator occur after allowing adequate time for materials to reach the chief administrators. Identified coordinators will receive a package containing duplicates of materials sent to the chief administrators plus

materials clearly explaining the coordinator's critical role in gaining access, consideration, and participation from staff within their institution. Also provided will be checklists clearly describing the steps of the data collection process and anticipated levels of effort.

Experienced staff from RTI's Call Center Services (CCS) carry out these contacts and are assigned a set of institutions that is their responsibility throughout the process. This allows RTI staff members to establish rapport with the institution staff and provides a reliable point of contact at RTI. Staff members are thoroughly trained in basic financial aid concepts and in the purposes and requirements of the study, which helps them establish credibility with the institution staff.

Endorsements. In previous NPSAS studies, the specific endorsement of relevant associations was extremely useful in persuading institutions to cooperate. Endorsements from 26 professional associations have been secured for NPSAS:08. These associations are listed in appendix F. In addition to providing general study endorsement, the National Association of Student Aid Administrators (NASFAA) promotes the study at its national and regional meetings and through the association's publications.

Minimizing burden. As in prior NPSAS studies, different options for providing enrollment lists and for extracting/recording the data requested for sampled students are offered. The coordinator is invited to select the methodology of greatest convenience to the institution. The optional strategies for obtaining the data are discussed later in this section. With regard to student record abstractions, "preloading" a customized list of financial aid awards into the computer assisted data entry (CADE) for each institution reduces the amount of data entry required for the institution and more closely tailors CADE to award names likely to be found in students' financial aid records. During institution contacting, the names of up to four of the most commonly awarded institution grants and scholarships are identified to assist in this process. Data on institution attributes such as institution level and control, highest level of offering, and other attributes are verified and updated as well.

b. Institutional Data Collection Training

Institution Coordinator Training. The purpose of an effective plan for training institution coordinators is two-fold: to make certain that survey procedures are understood and followed, and to motivate the coordinators. The project relies on these procedures to assure institutional data are recorded accurately and completely. Because institution coordinators are a critical element in this process, communicating instructions about their survey tasks clearly is essential.

Institution coordinators will be trained during the course of telephone contacts by call center staff. Written materials will be provided to coordinators explaining each phase of the project (enrollment list acquisition, student sampling, institution data abstraction, etc.) as well as their role in each.

Training of institution coordinators is geared toward the method of data collection selected by the institution. All institution coordinators will be informed about the purposes of NPSAS, provided with descriptions of their survey tasks, and assured of our commitment to maintaining the confidentiality of institution, student, and parent data. The CADE system is a World Wide Web application; and the CADE website, accessible only with an ID and password, provides institution coordinators with instructions for all phases of study participation. Copies of

all written materials, as well as answers to frequently asked questions, are available on the website.

In addition to the training activities described above, we will establish an exhibit booth at NASFAA's national conference in July of 2007. Attending this conference allows project management to meet staff from institutions who have previously participated, and answer any questions regarding the study, CADE, or institution burden. Because the date of this conference coincides with the completion of field test institution data collection activities, we may also be able to solicit feedback from financial aid administrators of field test institutions.

Field Data Collector Training. RTI will develop the training plan and training materials for the field-CADE data collectors and make arrangements for the training. One training session will be held, conducted by staff members who will be responsible for management of the institutional records data collection and who are experienced in conducting data collection from educational institutions. The training is designed to ensure that the data collectors are fully prepared to identify problems that may be encountered in working with schools and school records and to apply solutions that will result in the collection of consistently high quality data by all field staff. The training will include:

- a thorough explanation of the background, purpose, and design of the survey;
- an overview of the NPSAS institutional records data collection activity and its importance to the success of the study;
- a description of the role of the NPSAS data collector and his/her responsibility for obtaining complete and accurate data;
- an explanation of the role of the institution coordinator and how the data collector will interact with him/her;
- a full explanation of confidentiality and privacy regulations that apply to the data collector, including signing of nondisclosure affidavits;
- procedures for obtaining financial aid data from sample schools that must be visited;
- use of the CADE module and field case management system to collect, manage, and transmit data;
- completion and review of sample exercises simulating the various situations that will be encountered collecting student financial aid data from the various types of institutions included in the sample; and
- communication and reporting procedures.

The NPSAS Field Data Collector Manual will fully address each of the training topics and will describe all field data collection procedures in detail. The manual will be designed to serve as both a training manual and a reference manual for use during actual data collection. Training will emphasize active participation of the trainees and provide extensive opportunities for them to deal with procedures and the Information Management System (IMS). A major goal is preparing trainees to interact appropriately with the variety of school staff and different types of financial aid administration and record-keeping systems they will encounter at the NPSAS:08 sample institutions.

c. Collection of Student Data from Institutional Records

The highest priority goal for NPSAS:08 reflects its student aid focus. Institutions and federal financial aid databases are the best source for these data. Historically, institutional records have been a major source of student financial aid, enrollment, and locating data for NPSAS. As part of the institution contacting, institution coordinators will be asked to select a method of data collection—self-CADE (CADE completed by the institution via data entry through a secured website), field-CADE (CADE with the assistance of field data collectors), or data-CADE (submission of an electronic data file via a secured website). For the field test, we have assumed, based on our previous NPSAS experience, that 21 percent of eligible institutions will submit data-CADE—with 13 percent requiring a field data collector and the remaining 66 percent performing the abstraction themselves.

Prior to data collection, student records are matched to the U.S. Department of Education Central Processing System (CPS)—which contains data on federal financial aid applications—for locating purposes and to reduce the burden on the institutions for the student record abstractions. The vast majority of the federal aid applicants (about 95 percent) will match successfully to the CPS prior to CADE data collection, so we will ask the institution to provide the student's last name and Social Security number for the small number of federal aid applicants who did not match to the CPS prior to CADE. We will collect these two pieces of information in CADE and then submit the new names and Social Security numbers to CPS for file matching after CADE data collection has ended. Any new data obtained for the additional students will be delivered on the Electronic Code Book (ECB) with the data obtained prior to CADE. Under either scenario, we will have reduced the level of effort at the institution and thereby reduced the CADE cycle time.

Self-CADE via the Internet. Goals for NPSAS:08 CADE include reducing the data collection burden on NPSAS institutions (thereby reducing project costs by reducing the need for field data collectors), expediting data delivery, improving data quality, and ultimately ensuring the long-term success of NPSAS. NPSAS:2000 demonstrated the viability of a web-based approach to CADE data collection, and NPSAS:04 saw increased use of data-CADE, particularly by institutional systems. We propose to use a self-CADE instrument nearly identical to that used in NPSAS:04.

We had success with the self-CADE instrument in NPSAS:04 and believe more institutions are becoming accustomed to web applications, which will result in significant data collection schedule efficiencies. Under self-CADE, the NPSAS schedule will further benefit from the fact that multiple offices within the institution can enter data into CADE simultaneously, as successfully demonstrated in NPSAS:2000 and NPSAS:04.

Because the open Internet is not conducive to transmitting confidential data, any internet-based data collection effort necessarily raises the question of security. However, we intend to incorporate the latest technology systems into our web-CADE⁶ application to ensure strict adherence to NCES confidentiality guidelines. Our web server will include a Secure-Sockets Layer (SSL) Certificate, resulting in encrypted data transmission over the Internet. The SSL technology is most commonly deployed and recognizable in electronic commerce applications that alert users when they are entering a secure server environment, thereby protecting credit card numbers and other private information. Also, all of the data entry modules on this site are

⁶ To be used for both self-CADE and data-CADE submissions.

password protected, requiring the user to log in to the site before accessing confidential data. The system automatically logs the user out after 20 minutes of inactivity. This safeguard prevents an unauthorized user from browsing through the site. Additionally, we will stay attuned to technological advances to ensure the NPSAS:08 data are completely secure.

Data-CADE. Our CADE experience in NPSAS:2000 and NPSAS:04 confirmed that some coordinators prefer submitting files containing the institution data, rather than performing data entry into CADE. Allowing the institutions to submit CADE data in the form of a data file (via upload to the project's secure website) provides a more convenient mechanism by which institutions can provide data electronically (without performing data entry). Detailed specifications will be provided to the institutions that request this method. We will contact the institution to discuss thoroughly the content of the file and to clarify the exact specification requirements. To mitigate the costs of RTI programmers processing files in various formats, we will request that institutions providing CADE data files use the .CSV format.

Security for the CADE data files will be the same as that described above for self-CADE. File transmission via the web site will be protected by industry standard Secure Socket Layer (SSL) encryption technology.

Field-CADE. Field data collectors will conduct data abstractions at institutions not choosing self-administered CADE. The data collectors will arrange their visit to the institution with the coordinator and, once there, will abstract data from student records and key the data into CADE software using an RTI-provided laptop computer. The field-CADE data collection system will be identical to the self-CADE instrument but will run in local mode on the laptop, enabling the field data collector to enter the data without needing access to a data line at the institution.

Field data collectors will use a CADE procedures checklist to help them conduct discussions with the coordinator and perform all necessary tasks. The data collector will be provided with electronic files containing CADE preload information for all sampled students. When records abstraction is completed, the data collector will transmit a completed CADE file to RTI.

Data security will be of primary importance during field-CADE data collection. The following steps will be taken to ensure the protection of confidential information in the field.

Field laptops will be encrypted using a whole disk encryption software package, Pointsec. Pointsec encrypts the entire disk sector by sector, including the system files, temp files, and even deleted files. Boot protection authenticates users before the computer is booted, this prevents the operating system from being subverted by unauthorized persons.

- Field laptops will be configured so that during the startup a warning screen will appear, stating that the computer is the property of RTI and that criminal penalties apply to any unauthorized persons accessing the data on the laptop. The user must acknowledge this warning screen before startup will complete. Each laptop will have affixed a printed version of the same warning with a toll-free number to call if the laptop is found. Laptops will be configured to require a login and password at startup, and the case management system software will require an additional login and password before displaying the first menu. Field staff are instructed never to write down the passwords anywhere.

- To reduce the risk of intrusion should a laptop be obtained by an unauthorized person, communications software on field laptops will be configured to connect to RTI's network for data transfer (described in the paragraph below). The SQL server database used for data transfer will contain only case assignment and status data, including name and locating information; survey response data will be retrieved from the laptops and stored in a restricted project share. Completed cases' data files will be removed from the laptop during transmission after the data have been verified as being received at RTI.
- Data being sent to and from field laptops are stored in a domain of the RTI network that is behind the RTI firewall but allows access, with appropriate credentials, to users accessing RTI resources while physically outside the private domain (the innermost security login level accessible only by internal RTI staff). The particular file share in which the ingoing and outgoing data are housed is protected by NT security which allows access to the data only by RTI system administrators, field system programmers, and the controlled programs that are invoked when field interviewers' laptops connect via direct dialup to RTI's modems and communicate with the Integrated Field Management System (IFMS).

CADE Quality Control. As part of our quality control procedures, we will emphasize to CADE data abstractors the importance of collecting information for items. Items will not only have edit-checks applied to them during the CADE abstraction, they will also be analyzed by CADE when abstraction for a student is complete for a given section of the instrument. This CADE feature indicates which key items are missing or out of range and will provide both field data collectors and institution staff with an indication of the overall quality of their abstraction efforts.

As data are collected at institutions, either by field data collectors or institution staff, they will ultimately reside on the Integrated Management System (IMS). In the case of self-CADE institutions, the data will already be resident on the RTI web-server and will be copied directly into a special CADE subdirectory of the IMS. Web-based CADE will also allow improved quality control over the CADE process, as RTI central staff will be able to monitor data quality for participating schools closely and on a regular basis. When CADE institutions call for technical or substantive support, we will be able to query the institution's data and communicate much more effectively regarding any problems.

In the case of field-CADE institutions, the CASES files will be transmitted electronically from their modem-equipped laptop computers to the same location. From this subdirectory, automated quality control software, running nightly, will read the data files that arrived that day and produce quality control reports. These reports will summarize the completeness of the institution data and make comparisons to all other participating institutions, as well as to similar (i.e., same Level and Control) institutions.

d. Student Locating

Student interviews and student institutional record abstraction will occur simultaneously so that schedule requirements are met. To achieve the desired response rate, we propose a tracing approach that consists of up to four steps designed to yield the maximum number of locates with

the least expense. During the field test, we will evaluate the effectiveness of these procedures for the full-scale study effort. The steps of our tracing plan include the following elements.

- **Tracing prior to the start of data collection.** Our advance tracing operation will involve batch database searches and interactive database searches.
- **Lead letter mailings to sample members.** A personalized letter (signed by an NCES official), study leaflet, and information sheet will be mailed to all sample members to initiate data collection. This letter will include a toll-free 800 number, study website address, and study ID and password, and will request that sample members call to schedule an appointment to complete the interview by telephone, or complete the self-administered interview. One week after the lead letter mailing, a thank you/reminder postcard will be sent to sample members.
- **Intermediate tracing (during CATI but before intensive tracing).** Cases are processed in batches through Accurant for address and telephone updates. All new information is loaded into our CATI system for attempts to contact the sample members. Cases for which no new information is returned are forwarded to Call Center Services (CCS) tracing services.
- **Intensive tracing.** The goal of intensive tracing is to obtain a telephone number the sample member can be reached at by a CATI interviewer in a cost-effective manner. Tracing procedures may include (1) checking Directory Assistance for telephone listings at various addresses; (2) using criss-cross directories to obtain the names and telephone numbers of neighbors and calling them; (3) calling persons with the same unusual surname in small towns or rural areas to see if they are related to or know the sample member; and (4) contacting the current or last known residential sources such as neighbors, landlords, and current residents at the last known address. Other more intensive tracing activities could include (1) database checks for sample members, parents, and other contact persons, (2) credit database and insurance database searches, (3) drivers' license searches through the appropriate state departments of motor vehicles, and (4) calls to alumni offices and associations.

e. Student Data Collection: Self-Administered Web and CATI

Training Procedures. Training programs for those involved in survey data collection are critical quality control elements. Training for the help desk operators who answer questions for the self-administered web-based student interview and CATI telephone interviewers will be conducted by a training team with extensive experience. We will establish thorough selection criteria for help desk operators and telephone interviewers to ensure that only highly capable persons—those with exceptional computer, problem-solving, and communication skills—are selected to serve on the project and will contribute to the quality of the NPSAS data.

Contractor staff with extensive experience in training interviewers will prepare the NPSAS:08 Student Survey Telephone Interviewer Manual, which will provide detailed coverage of the background and purpose of NPSAS, sample design, questionnaire, and procedures for the CATI interview. This manual will be used in training and as a reference during interviewing. (Interview-specific information will be available to interviewers in the Call Center in the form of question-by-question specifications providing explanations of the purpose of each question and any definitions or other details needed to aid the interviewers in obtaining accurate data.) Along

with manual preparation, training staff will prepare training exercises, mock interviews (specially constructed to highlight the potential of definitional and response problems), and other training aids.

A comprehensive training guide will also be prepared for use by trainers to standardize training and to ensure that all topics are covered thoroughly. Among the topics to be covered at the telephone interviewer training will be:

- the background purposes and design of the survey;
- confidentiality concerns and procedures (interviewers will take an oath and sign an affidavit agreeing to uphold the procedures);
- importance of locating/contacting sample members and procedures for using the IMS/CATI locating and tracing module;
- special practice with online coding systems used to standardize sample member responses to certain items (e.g., institution names);
- review, discussion, and practice of techniques for explaining the study, answering questions asked by sample members, explaining the respondent's role, and obtaining cooperation;
- extensive practice in applying tracing and locating procedures;
- demonstration interviews by the trainers;
- round-robin (interactive mock interviews for each section of each questionnaire, followed by review of the question-by-question specifications for each section);
- completion of classroom exercises;
- practice interviews with trainees using the web/CATI instrument to interview each other while being observed by trainers, followed by discussion of the practice results; and
- explanation of quality control procedures, administrative procedures, and performance standards.

Telephone survey unit supervisors will be given project-specific training in advance of interviewer training and will assist in monitoring interviewer performance during the training.

Student Interviews (web/CATI). Student interviews will be conducted using a single web-based survey instrument for both self-administered and CATI data collection. The data collection activities will be accomplished through the Case Management System (CMS), which is equipped with the following capabilities:

- on-line access to locating information and histories of locating efforts for each case;
- state-of-the-art questionnaire administration module with full "front-end cleaning" capabilities (i.e., editing as information is obtained from respondents);
- sample management module for tracking case progress and status; and
- automated scheduling module which delivers cases to interviewers and incorporates the following features:

- *Automatic delivery of appointment and call-back cases at specified times.* This reduces the need for tracking appointments and helps ensure the interviewer is punctual. The scheduler automatically calculates the delivery time of the case in reference to the appropriate time zone.
- *Sorting of non-appointment cases according to parameters and priorities set by project staff.* For instance, priorities may be set to give first preference to cases within certain sub-samples or geographic areas; cases may be sorted to establish priorities between cases of differing status. Furthermore, the historic pattern of calling outcomes may be used to set priorities (e.g., cases with more than a certain number of unsuccessful attempts during a given time of day may be passed over until the next time period). These parameters ensure that cases are delivered to interviewers in a consistent manner according to specified project priorities.
- *Restriction on allowable interviewers.* Groups of cases (or individual cases) may be designated for delivery to specific interviewers or groups of interviewers. This feature is most commonly used in filtering refusal cases, locating problems, or foreign language cases to specific interviewers with specialized skills.
- *Complete records of calls and tracking of all previous outcomes.* The scheduler tracks all outcomes for each case, labeling each with type, date, and time. These are easily accessed by the interviewer upon entering the individual case, along with interviewer notes, thereby eliminating the need for a paper record of calls of any kind.
- *Flagging of problem cases for supervisor action or supervisor review.* For example, refusal cases may be routed to supervisors for decisions about whether and when a refusal letter should be mailed, or whether another interviewer should be assigned.
- *Complete reporting capabilities.* These include default reports on the aggregate status of cases and custom report generation capabilities.

The integration of these capabilities reduces the number of discrete stages required in data collection and data preparation activities and increases capabilities for immediate error reconciliation, which results in better data quality and reduced cost. Overall, the scheduler provides a highly efficient case assignment and delivery function by reducing supervisory and clerical time, improving execution on the part of interviewers and supervisors by automatically monitoring appointments and call-backs, and reducing variation in implementing survey priorities and objectives.

In addition to the management aspect of data collection, the survey instrument is another component designed to maximize efficiency and yield high-quality data. Below are some of the basic questionnaire administration features of the web-based instrument:

- Based on responses to previous questions, the respondent or interviewer is automatically routed to the next appropriate question, according to predesignated skip patterns.

- The web-based interview automatically inserts “text substitutions” or “text fills” where alternate wording is appropriate depending on the characteristics of the respondent or his/her responses to previous questions.
- The web-based interview can incorporate or preload data about the individual respondent from outside sources (e.g., previous interviews, sample frame files, etc.). Such data are often used to drive skip patterns or define text substitutions. In some cases, the information is presented to the respondent for verification or to reconcile inconsistencies.
- With the web/CATI instrument, numerous question-specific probes may be incorporated to explore unusual responses for reconciliation with the respondent, to probe “don’t know” responses as a way of reducing item non-response, or to clarify inconsistencies across questions.
- Coding of multi-level variables. An innovative improvement to previous NPSAS data collections, the web-based instrument uses an assisted coding mechanism to code text strings provided by respondents. Drawing from a database of potential codes, the assisted coder derives a list of options from which the interviewer or respondent can choose an appropriate code (or codes if it is a multi-level variable with general, specific, and/or detail components) corresponding to the text string.
- Iterations. When identical sets of questions will be repeated for an unidentified number of entities, such as children, jobs, schools, and so on, the system allows respondents to cycle through these questions as often as is needed.

In addition to the functional capabilities of the CMS and web instrument described above, our efforts to achieve the desired response rate will include using established procedures proven effective in other large-scale studies we have completed. These include:

- Providing multiple response modes, including self-administered and interviewer-administered options.
- Offering incentives to encourage response (see incentive structure described below).
- Prompting calls initiated prior to the start of data collection to remind sample members about the study and the importance of their participation.
- Assigning experienced CATI data collectors who have proven their ability to contact and obtain cooperation from a high proportion of sample members.
- Training the interviewers thoroughly on study objectives, study population characteristics, and approaches that will help gain cooperation from sample members.
- Providing the interviewing staff with a comprehensive set of questions and answers that will provide encouraging responses to questions that sample members may ask.
- Maintaining a high level of monitoring and direct supervision so that interviewers who are experiencing low cooperation rates are identified quickly and corrective action is taken.
- Making every reasonable effort to obtain an interview at the initial contact, but allowing respondent flexibility in scheduling appointments to be interviewed.

- Providing hesitant respondents with a toll-free number to use to telephone RTI and discuss the study with the project director or other senior project staff.
- Thoroughly reviewing all refusal cases and making special conversion efforts whenever feasible (see next section).

Refusal Aversion and Conversion. Recognizing and avoiding refusals is important to maximize the response rate. We will emphasize this and other topics related to obtaining cooperation during data collector training. Supervisors will monitor interviewers intensely during the early days of data collection and provide retraining as necessary. In addition, the supervisors will review daily interviewer production reports produced by the CATI system to identify and retrain any data collectors who are producing unacceptable numbers of refusals or other problems.

After encountering a refusal, the data collector enters comments into the CMS record. These comments include all pertinent data regarding the refusal situation, including any unusual circumstances and any reasons given by the sample member for refusing. Supervisors will review these comments to determine what action to take with each refusal. No refusal or partial interview will be coded as final without supervisory review and approval. In completing the review, the supervisor will consider all available information about the case and will initiate appropriate action.

If a follow-up is clearly inappropriate (e.g., there are extenuating circumstances, such as illness or the sample member firmly requested that no further contact be made), the case will be coded as final and will not be recontacted. If the case appears to be a “soft” refusal, follow-up will be assigned to an interviewer other than the one who received the initial refusal. The case will be assigned to a member of a special refusal conversion team made up of interviewers who have proven especially adept at converting refusals.

Refusal conversion efforts will be delayed for at least one week to give the respondent some time after the initial refusal. Attempts at refusal conversion will not be made with individuals who become verbally aggressive or who threaten to take legal or other action. Refusal conversion efforts will not be conducted to a degree that would constitute harassment. We will respect a sample member’s right to decide not to participate and will not impinge this right by carrying conversion efforts beyond the bounds of propriety.

Incentives to Convert Refusals, Difficult and Unable-to-Locate Respondents. As described in the justification section (section A), we have proposed to offer incentive payments to nonresponding members of the sample population. We believe there will be three groups of nonrespondents: persons refusing to participate during early response or production interviewing, persons who have proven difficult to interview (i.e., those who repeatedly break appointments with an interviewer), and those who cannot be located or contacted by telephone. Our approach to maximizing the response of these persons—and thereby limiting potential nonresponse bias— involves an incentive payment to reimburse the respondent for time and expenses. The NPSAS:08 field test will be used to conduct an experiment to determine whether a \$10 prepaid incentive followed by \$20 upon survey completion yields higher response rates than the promise of a \$30 incentive. Additional detail about planned field test experiments is provided in section B.4.

Additional Quality Control. In addition to the quality control features inherent in the web-based interview (described in section 3), we will use data collector monitoring as a major quality control measure. Supervisory staff from RTI's Call Center Services (CCS) will monitor the performance of the NPSAS:08 data collectors throughout the data collection period to ensure they are following all data collection procedures and meeting all interviewing standards. In addition, members of the project management staff will monitor a substantial number of interviews. In all cases, students will be informed of the fact that the interview may be monitored by supervisory staff.

"Silent" monitoring equipment is used so that neither the data collector nor respondent is aware when an interview is being monitored. This equipment will allow the monitor to listen to the interview and simultaneously see the data entry on a computer screen. The monitoring system allows ready access to any of the work stations in use at any time. The monitoring equipment also enables any of the project managers and client staff at RTI or NCES to dial in and monitor interviews from any location. In the past, we have used this capability to allow the analysts to monitor interviews in progress; as a result, they have been able to provide valuable feedback on specific substantive issues and have gained exposure to qualitative information that has helped their interpretation of the quantitative analyses.

Our standard practice is to monitor 10 percent of the interviewing done by each data collector to ensure that all procedures are implemented as intended and that the procedures are effective, and to observe the utility of the questionnaire items. Any observations that might be useful in subsequent evaluation will be recorded and all such observations will be forwarded to project management staff. Staff monitors will be required to have extensive training and experience in telephone interviewing as well as supervisory experience.

4. Tests of Procedures and Methods

The following sections will briefly discuss four areas of data collection believed to affect overall study response. These areas are: presentation of notification materials, reminder prompting, early response incentive offers, and nonresponse conversion incentives. This section will also introduce our plans to conduct experiments to evaluate the effectiveness of these areas during the 2008 National Postsecondary Student Aid Study (NPSAS:08) field test.

a. Notification Materials

Much research about survey response has focused on the impact of procedures and materials used in contacting sample members, including the number of contacts, the timing of contacts, and the presentation of materials (Heberlein and Baumgartner 1978). Some studies have suggested that, in addition to the content and timing of study materials, the packaging or presentation of information sent to sample members is also very important to increasing survey response (Dillman 2001).

In particular, the method of mail delivery has been found to be an important factor. For instance, Abreu and Winters (1999) found that Priority Mail® was effective when used as a method to increase response rates among non-responding cases. Similarly, Moore and An (2001) found the use of Priority Mail in a prenotification mailing and a reminder mailing to be most effective in their mail questionnaire survey (2001). The reason is obvious: content is ineffective if the envelope is ignored, assumed to be junk mail. Couper, et al. found that mail is usually sorted by only one person in half of all house holds. Furthermore, 60 percent of people discard

mail without opening it (Couper, Mathiowetz, and Singer 2005), and therefore, it is imperative that researchers maximize the chances of their mailings being read. Increasing the look of legitimacy can go a long way to ensuring that the mail is opened by the intended recipient, thereby increasing the likelihood of survey response.

While no formal experiments have been conducted to examine the impact of study materials on response rates for the National Center for Education Statistics (NCES) studies, our past experience has suggested that the look of study materials is important. For instance, study materials have been sent to sample members on official letterhead, followed by e-mail and postcard reminders, and telephone prompting. As supplements to the traditional mailings, we have also used high visibility fliers as participation reminders, sent through regular mail and FedEx. The Beginning Postsecondary Students (BPS:04/06) field test used FedEx to deliver materials to sample members who were identified as nonrespondents towards the end of data collection and saw positive response after this mailing.

Consistent with current industry practice, the NCES postsecondary studies have implemented a strategy of multiple contacts through various means: postal mail, e-mail, and telephone contacts. When possible, advanced notification mailings, which have been associated with increased response rates (Leeuw, Hox, Korendijk, and Lensvelt-Mulders 2006; Curtin, Presser, and Singer 2005; Goldstein and Jennings 2002), have been sent prior to the actual survey invitation. However, the NPSAS study is unique in that data collection begins so soon after the student sample is selected that there is not enough time for an “advanced notification” mailing. This scheduling limitation makes it even more important that the first contact with students attracts attention.

We believe that using the United States Post Office Priority Mail delivery system will signal the importance of the information contained in the package, increasing the likelihood that the materials will be read, and in turn, the likelihood of survey response. We propose to test the effectiveness of Priority Mail versus regular mail for the initial mailing to sample members (which includes the introduction to the study and the invitation to participate).

Prior to the start of data collection, the field test sample will be randomly assigned to two groups: one group will receive the initial study materials via Priority Mail and the other group will receive the same materials via regular mail, as has been done in the past. The initial mailing will contain important information about the study, including the study brochure and information needed to log into the study website to complete the interview. A toll-free telephone number will be provided so sample members can contact the study’s Help Desk for assistance, and can also complete a telephone interview if desired. Finally, the sample member will be informed of the details of the incentive offer and the expiration date of the early response period.⁷ Results will be measured by comparing the response rates at the end of the early response period for these two groups to determine whether response was greater for those who received the Priority Mail.

⁷ The “early response period” is defined as the first 3 weeks after the data collection notification is sent to a sample member. Sample members are notified of the study and asked to participate by completing a web-based self-administered survey. Help Desk staff are available during this time period to assist sample members and complete a telephone interview if desired.

Null Hypothesis to be Tested⁸

There will be no difference in response rates at the end of the early response period for those who receive the study materials and survey invitation via Priority Mail or regular mail.

Estimated Cell Sizes

Control group		Treatment group		Detectable difference with 95 percent confidence
Definition	Sample size	Definition	Sample size	
Regular mail	1,179	Priority Mail®	1,179	3.2

b. Prompting

Research has shown that additional contacts with sample members increase the likelihood of participation (Moore and Dillman, 1980). Prompting calls, which are outbound calls made by project staff to sample members reminding them to participate, are likely effective because they provide another reminder about a study, and also give interview staff an additional opportunity to provide the information needed to participate. Prompting calls also give an early indication of the quality of locating information for a case.

In the field test for BPS:04/06, an experiment was conducted to evaluate the effectiveness of telephone prompting calls made half-way through the early response period. Response rates at the end of the early response period were significantly higher among sample members who received the prompting calls. (21.5 percent vs. 10.4 percent, respectively; $z = 5.52$, $p < 0.05$).

Furthermore, it was discovered that the lower response rate typically observed among base year nonrespondents when compared with base year respondents (BPS:96/98, BPS:96/01) was not apparent among the cases who received the prompting treatment. This suggests that prompting reduced the likelihood of nonresponse to the follow-up interview among base year nonrespondents.

However, it is not yet known whether prompting will be equally beneficial for all NPSAS sample members, or if the effect of prompting on response rates is different for sub-groups. To prompt all sample members in a study as large as NPSAS could have potentially serious cost implications. While there will not be enough cases in the field test sample to support a comparison of the effects of prompting across sub-groups, response patterns to the previous full-scale administration (NPSAS:04) will be examined to determine whether certain groups would be more appropriate for prompting (e.g., those who responded via the web, but not during the early response period, sample members who were classified as refusals, or those with high call counts, etc.) If it is determined that prompting is effective in increasing response rates, but would be cost-prohibitive to implement for the entire sample, then information from the NPSAS:04 full-scale study will be used to identify sub-groups of students for whom prompting is expected to be most effective.

We propose to test the effectiveness of prompting calls during the early response period in the NPSAS:08 field test. Prior to data collection, the field test sample will be randomly assigned to two groups: one will receive telephone prompting calls reminding them to log in to

⁸ See section 5 for more detail about the field test sample, assumptions, and experimental design plans.

the study website and complete an interview and the other group will receive no prompting calls. For those in the treatment group, prompting calls will occur approximately 2 weeks into the 4-week early response period. Response rates at the end of the early response period for the two groups will be compared to determine whether the prompting calls are significantly associated with higher response rates during the early response period.

Null Hypothesis to be Tested

There will be no difference in response rates at the end of the early response period for those who receive prompting calls during the early response period and those who do not.

Estimated Cell Sizes

Control group		Treatment group		Detectable difference with 95 percent confidence
Definition	Sample size	Definition	Sample size	
No prompting	1,110	Early prompting	1,110	3.3

NOTE: Five percent of the sample members will respond prior to the prompting phase during the early response period (e.g., within the first 2 weeks of data collection, before any outbound prompting calls are made).

c. Early Response Incentives

The offer of monetary incentives has been studied extensively in survey research. Incentive amounts have been shown to have a strong, positive linear relationship to response rates (Yu and Cooper, 1983). Furthermore, Singer et al. found that response rates increased as the incentive amount increased⁹ (Singer, Van Hoewyk, Gebler, Raghunathan, McGonagle, 1999).

Several experiments regarding the effectiveness of various incentive plans have been conducted during the field tests of NCES postsecondary studies. With the increasing use of internet surveys, experiments have focused on an “early response” incentive, which attempts to encourage early, self-administered interview completion, thus reducing the cost and time required for these sample members who would otherwise have not responded early in the data collection period.

The Baccalaureate and Beyond (B&B:93/03) was the first NCES study of postsecondary students to include the web-based self-administered option. With this new mode of administration, an experiment was designed to determine whether the offer of an early response incentive would encourage sample members to complete their interview via the Web early in data collection. Results of this experiment suggested that the early response incentive offer may be a useful strategy for encouraging early, self-administered survey completion. Those who received the incentive offer of \$20 responded at a rate of 12.7 percent within the specified time frame (3 weeks) compared with 8.7 percent for those who did not receive an early response incentive offer ($z = 1.9, p < 0.05$).

In the field test for NPSAS:04, another experiment was conducted to compare three different amounts for the early response incentive: \$0, \$10, and \$20. While response rates for both the \$10 and \$20 groups were significantly higher than the \$0 group (23.2 percent vs. 13.1

⁹ This study was a meta-analysis of interviewer-administered surveys that offered incentives.

percent; $t = 4.43$, $p < 0.0001$), no difference was observed between the \$10 and \$20 amounts in terms of response rates¹⁰ (24.3 percent vs. 22.2 percent; $t = 1.65$, $p = 0.24$).

Most recently, the BPS:04/06 field test offered a \$30 incentive for early response to all sample members. The \$30 amount was used because that was the same amount offered to these sample members as a nonresponse conversion incentive at the end of the base year study (NPSAS:04). While an experiment was not conducted to evaluate the use of the \$30 either as a nonresponse conversion incentive for NPSAS or as an early response incentive for BPS, it did appear to be effective. Of all completed interviews in the NPSAS:04 full-scale study, approximately 35 percent received the \$30 nonresponse conversion incentive. When the \$30 incentive was used in the BPS field test, 20 percent of the entire field test sample participated during the early-response period, and nearly 40 percent of all completed interviews were obtained during the early period. For comparison, the NPSAS:04 full-scale study offered a \$10 early response incentive, and obtained a response rate of 17 percent of all eligible sample members, accounting for about 28 percent of all completed interviews.

A \$10 early response incentive was budgeted for the NPSAS:08 field test. As described above, in NPSAS:04, \$10 was more effective than no incentive, but \$20 was not significantly better than \$10. However, when \$30 was offered in NPSAS:04 as a nonresponse conversion incentive, response rates improved substantially. We therefore propose to test whether a \$30 early incentive for NPSAS:08 will increase response rates more than a \$10 incentive amount.

Null Hypothesis to be Tested

There will be no difference in response rates at the end of the early response period for those who receive an offer of a \$10 or \$30 early response incentive.

Estimated Cell Sizes

Control group		Treatment group		Detectable difference with 95 percent confidence
Definition	Sample size	Definition	Sample size	
\$10 early incentive	1,179	\$30 early incentive	1,179	3.2

d. Nonresponse Conversion Incentives

Another strategy commonly used to obtain sufficient response to survey data collections is the nonresponse conversion incentive. The model used recently for the NCES studies has typically required that a case meet one of the following conditions:

- refusal to participate,
- hard to locate (e.g., have a mailing address, but not a good telephone number), or
- high call count (e.g., >10 or 15).

Once a case has been identified as eligible for the nonresponse conversion incentive, an additional incentive offer is made in an attempt to obtain a completed interview. In previously conducted NCES studies, the nonresponse conversion incentive amount has been high (or at least as high), as the early response incentive amount. In NPSAS:04, a field test experiment found that

¹⁰ Lack of observed difference may have been due to small sample sizes.

a \$20 nonresponse conversion incentive was associated with significantly higher response rates than those obtained from the group that did not receive a nonresponse incentive (32.5 percent vs. 15.8 percent; $t = 4.84, p < 0.0001$). As described above, the NPSAS:04 full-scale study increased the nonresponse conversion incentive from \$20 to \$30 toward the end of data collection (because there was not much response to the \$20 offer) and saw a dramatic increase in response rates.

Evidence suggests that nonresponse conversion incentives are an effective tool for increasing response rates among sample members who do not respond to early attempts to obtain a completed interview. However, little research has been done to identify the most effective combinations of incentive amount offers over the course of data collection—from early response period to production interviewing¹¹ through nonresponse conversion.

RTI’s experience, particularly in NPSAS:04, has demonstrated that using an early response incentive (e.g., \$10) and a high nonresponse conversion incentive (e.g., \$30) increased response rates. What we have not yet tested is whether prepayment increases response rates more than promised incentive offers. There is much evidence to suggest that prepaid incentives increase response rates more than promised incentives (Dillman, 2000; US Dept of Ed, 2004; Groves, et. al. 2004) However, prepaid incentives are operationally very difficult to administer, especially with a large sample such as this, that includes many cases that require tracing. To accommodate an evaluation of the impact of prepaid incentives, then, we propose to limit our analysis to the difficult cases at the end of data collection—those determined to be eligible for the nonresponse conversion incentive. This will allow us to assess the impact of prepayment on a reduced scale to determine whether it would be effective to implement for targeted groups in the full-scale sample

Null Hypothesis to be Tested

There will be no difference in response rates for those who are offered a prepaid nonresponse incentive (\$10 followed by \$20 upon interview completion) and those who are offered a \$30 nonresponse conversion incentive upon interview completion.

Estimated Cell Sizes

Control group		Treatment group		Detectable difference with 95 percent confidence
Definition	Sample size	Definition	Sample size	
\$30 nonresponse incentive	693	\$10 prepaid followed by \$20 nonresponse incentive	693	5.0

e. Experimental Design

Based on lessons learned from past RTI studies and relevant findings in the literature, RTI proposes to conduct 4 experiments that will test the hypotheses outlined below. The first set of hypotheses test each condition independently, and were discussed in the sections above. Additionally, the interactions will be examined to explore whether a data collection strategy that combines some or all of the independent conditions yields higher response rates than the

¹¹ Production interviewing is the phase of data collection between the early response period and the nonresponse conversion period, during which outbound computer-assisted telephone interviewing (CATI) occurs.

individual conditions alone. The next section illustrates the design option proposed. Finally, we provide detail about the field test sample and its allocation to each of the cells, and discuss the assumptions made in developing the design.

Hypotheses

One-way comparisons

1. There will be no difference in response rates during the early response period for those who receive the study materials and survey invitation via Priority Mail and those who receive the study materials via regular mail.
2. There will be no difference in response rates during the early response period for those who receive prompting calls during the early response period and those who do not.
3. There will be no difference in response rates during the early response period for those who are offered a \$10 or a \$30 early response incentive.

Two-way comparisons

4. There will be no difference in response rates during the early response period for those who receive the study materials and survey invitation via Priority Mail and who receive prompting calls when compared with all others.
5. There will be no difference in response rates during the early response period for those who receive the study materials and survey invitation via Priority Mail and who receive an offer of a \$30 early response incentive offer when compared with all others.
6. There will be no difference in response rates during the early response period for those who receive prompting calls and an offer of a \$30 early response incentive offer when compared with all others.

Three-way comparison

7. There will be no difference in response rates during the early response period for those who receive the Priority Mail, the prompting treatment, and the \$30 incentive offer when compared with all others.

Nonresponse conversion period

8. There will be no difference in response rates for those who are offered a prepaid nonresponse incentive (\$10 followed by \$20 upon interview completion) and those who are offered a \$30 nonresponse conversion incentive upon interview completion.

f. Design Options

The following section illustrates the design option for the proposed experiments to be conducted in the NPSAS:08 field test. The field test will have a sample of 3,000 students, of which 2,775 are estimated to be eligible and 1,942 are estimated to respond to the student interview. Tables 9 illustrates the sample design and allocation.

Table 9. Sample design and allocation

Mail condition	Early response incentive	Prompting	Production interviewing ¹	Nonresponse conversion period
Regular mail	Low (\$10)	Yes	\$0	Prepaid
				Not prepaid
	High (\$30)	No	\$0	Prepaid
				Not prepaid
	High (\$30)	Yes	\$0	Prepaid
				Not prepaid
Priority Mail®	Low (\$10)	Yes	\$0	Prepaid
				Not prepaid
	High (\$30)	No	\$0	Prepaid
				Not prepaid
	High (\$30)	Yes	\$0	Prepaid
				Not prepaid
High (\$30)	No	\$0	Prepaid	
			Not prepaid	

¹ Production interviewing is the phase of data collection between the early response period and the nonresponse conversion period, during which outbound CATI interviewing occurs.

NOTE: This design option allows for the testing of hypotheses 1-8, listed above.

Detectable Differences

As part of the planning process for developing the field test experiment design, the response rate differences between the control and treatment groups necessary to detect statistically significant differences will be estimated. That is, how large of a difference is necessary to be able to say that the response rates between the two groups are different. Table 11 shows the expected sample sizes and statistically significant detectable difference for each of the eight hypotheses. Several assumptions were made regarding response rates and sample sizes. In general, the closer the response rate is to 50 percent (either less than or greater than), the larger the detectable difference. Likewise, the smaller the sample size, the larger the detectable difference.

Assumptions:

1. The sample will be equally distributed across experimental cells.
2. All ineligible cases will be excluded from the analysis even if they are not determined ineligible until the interview.
3. Five percent of the sample members will respond prior to the prompting phase during the early response period (e.g., within the first 2 weeks of data collection, before any outbound prompting calls are made).
4. Fifteen percent of the sample members will not be located and therefore will be excluded from the experiment.

5. The response rate for the control group for hypotheses 1 through 7 will be 18 percent.¹²
6. Thirty-five percent of the eligible sample members will respond before the nonresponse conversion period (18 percent during the early response period, and 17 percent will respond during the production interviewing phase).
7. The response rate for the control group for hypothesis 8 will be 65 percent.¹³

¹² Eighteen percent is used here as a baseline because it is consistent with response rates obtained during the early response period from past studies.

¹³ Sixty-five percent is used here as a baseline because we expect to get less than a 70 percent response rate with a \$20 incentive, and we expect to get about a 70 percent response rate using a \$30 incentive (consistent with NPSAS:04).

Table 10. Detectable differences for field test experiment hypotheses

Hypothesis	Control group		Treatment group		Detectable difference with 95 percent confidence
	Definition	Sample size	Definition	Sample size	
1	Regular mail	1,179	Priority Mail®	1,179	3.2
2	No prompting	1,110	Early prompting	1,110	3.3
3	\$10 early incentive	1,179	\$30 early incentive	1,179	3.2
4	All others	1,665	Priority Mail and early prompting	555	3.9
5	All others	1,769	Priority Mail and \$30 early incentive	589	3.8
6	All others	1,665	Early prompting and \$30 early incentive	555	3.9
7	All others	1,942	Priority Mail, early prompting, and \$30 early incentive	277	5.3
8	\$30 nonresponse incentive	693	\$10 prepaid plus \$20 nonresponse incentive	693	5.0

5. Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study

Names of individuals consulted on statistical aspects of study design along with their affiliation and telephone numbers are provided below.

Name	Affiliation	Telephone
Dr. Lutz Berkner	MPR	(510) 849-4942
Dr. Susan Choy	MPR	(510) 849-4942
Dr. E. Gareth Hoachlander	MPR	(510) 849-4942
Dr. John Riccobono	RTI	(919) 541-7006
Dr. James Chromy	RTI	(919) 541-7019
Dr. Karol Krotki	RTI	(202) 728-2485
Dr. Roy Whitmore	RTI	(919) 541-5809
Mr. Peter Siegel	RTI	(919) 541-6348

In addition to these statisticians and survey design experts, the following statisticians at NCES have also reviewed and approved the statistical aspects of the study: Dr. Dennis Carroll, Dr. James Griffith, and Dr. Paula Knepper.

6. Other Contractors' Staff Responsible for Conducting the Study

The study is being conducted by the Postsecondary Longitudinal Studies Branch of the National Center for Education Statistics (NCES), U.S. Department of Education. NCES's prime contractor is the RTI International (RTI). RTI is being assisted through subcontracted activities by MPR Associates and the National Association of Student Aid Administrators (NASFAA). Principal professional staff of the contractors, not listed above, who are assigned to the study are provided below:

Name	Affiliation	Telephone
Ms. Mary Ann O'Connor	NASFAA	(202) 785-0453
Ms. Vicky Dinger	MPR	(510) 849-4942
Mr. Tim Gabel	RTI	(919) 541-7415
Dr. Laura Horn	MPR	(510) 849-4942
Mr. Jeff Franklin	RTI	(919) 541-2614
Ms. Christine Rasmussen	RTI	(919) 541-6775
Ms. Melissa Cominole	RTI	(919) 990-8456
Ms. Kristin Dudley	RTI	(919) 541-6855
Mr. Brian Kuhr	RTI	(312) 456-5263