

November 18, 2014

# High School Longitudinal Study of 2009 (HSLS:09) Second Follow-up Field Test and Main Study Panel Maintenance 2015

## Supporting Statement Part B

Request for OMB Review  
OMB# 1850-0852 v.15

National Center for Education Statistics  
U.S. Department of Education

## TABLE OF CONTENTS

Section	Page
B. Collection of Information Employing Statistical Methods.....	1
B.1 Target Universe and Sampling Frames.....	1
B.1.a Student Universe.....	1
B.1.b Institution Universe.....	1
B.2 Statistical Procedures for Collecting Information.....	1
B.2.a Field Test and Main Study Student Sample.....	1
B.2.b Institution Sample.....	2
B.2.c Institution Sample Phase I and Phase II.....	2
B.2.d Student Sample for Phases I and II.....	2
B.3 Methods for Maximizing Response Rates.....	2
B.3.a Locating.....	2
B.3.b Interviewing Procedures.....	3
B.3.c Quality Control.....	5
B.3.d Panel Maintenance.....	5
B.3.e Postsecondary education transcript study and financial aid record collection.....	6
B.3.f HSLs:09 Postsecondary Education Transcript Study (HSLs:09 PETS) collection.....	7
B.3.g HSLs:09 Financial aid record collection (HSLs:09 FAR) collection.....	13
B.4 Field Test Experiments.....	14
B.5 Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study.....	18
Recent Responsive Design References Using HSLs:09 Data.....	18

## EXHIBITS

Number	Page
Exhibit B-1. HSLs:09 results: Percentage taking Algebra 1 in ninth grade, by phase of data collection.....	14
Exhibit B-2. Schedule of interventions in the HSLs:09 Second Follow-Up Field Test.....	16

## **B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS**

This section describes the target universe for this study and the sampling and statistical methodologies proposed for the field test and the main study. This section also addresses suggested methods for maximizing response rates and for tests of procedures and methods, and introduces the statisticians and other technical staff responsible for design and administration of the study.

### **B.1 Target Universe and Sampling Frames**

#### ***B.1.a Student Universe***

The base-year target population for HSLS:09 consisted of 9<sup>th</sup> grade students in the fall of 2009 (main study) or fall of 2008 (field test) in public and private schools that include 9<sup>th</sup> and 11<sup>th</sup> grades. The target population for the second follow-up is this same 9<sup>th</sup> grade cohort in 2016 (main study) or 2015 (field test). The target population for the postsecondary transcripts and student financial aid records collections is the subset of this same 9<sup>th</sup> grade cohort who had postsecondary education enrollment as of 2016.

#### ***B.1.b Institution Universe***

The HSLS:09 second follow-up study will include a sample of institutions; these institutions will be identified as part of the second follow-up survey when respondents indicate having enrolled at the postsecondary level. These institutions will constitute the sample from which postsecondary transcripts and student financial aid records will be collected.

### **B.2 Statistical Procedures for Collecting Information**

#### ***B.2.a Field Test and Main Study Student Sample***

A subset of the students who were sampled in the 2008 base-year field test will be recruited for participation in the second follow-up field test. The subset comprises 1,100 base-year field test students who responded in the base year or first follow-up. The second follow-up field test sample includes 748<sup>1</sup> of the 754 sample members included in the 2012 Update field test and 352 base-year respondents who were included in neither the 2012 Update field test nor the first follow-up field test<sup>2</sup>.

In the main study, the same students who were sampled for the 2009 base-year data collection, excluding deceased sample members, those who elected to withdraw from HSLS:09, and those who did not respond in either the base year or first follow-up, will be recruited to participate in the 2016 data collection. The main study sample will include 25,184 sample members. Of these, 23,316 will be fielded; the remaining cases have been identified as final refusals or as sample members who did not respond in either the base year or first follow-up and will be classified as nonrespondents.

---

<sup>1</sup> Study withdrawals and deceased sample members are excluded from the second follow-up field test.

<sup>2</sup> A subset of 26 of the 41 base-year field test participating schools and associated students were included in the first follow-up field test and 2012 Update activities. The other schools and associated students were not pursued; 352 of those students participated in the base year and have been re-introduced into the second follow-up field test sample.

**B.2.b Institution Sample**

The HSLS:09 PETS postsecondary transcript collection and financial aid record (FAR) collection will be divided into two phases.

- Phase I. The set of student self-reported institutions will be asked to participate in the PETS transcript collection and FAR student record collection.
- Phase II. Institutions with new student-institution pairs identified in the phase I (and phase II) collected transcripts will be contacted to provide postsecondary transcript data.

The total number of postsecondary institutions is estimated to be 3,800. Using eligibility and participation rates observed in the ELS:2002 postsecondary transcript data collection, an expected 98 percent of the 3,800 schools will be eligible and approximately 3,230 postsecondary institutions (approximately 87 percent) will provide transcripts and financial aid data.

**B.2.c Institution Sample Phase I and Phase II**

The institution sample for the first phase of transcript and financial aid record collection will comprise all eligible institutions reported by second follow-up sample members. The second phase involves reviewing transcripts collected during phases I and II, identifying new student-institution linkages not otherwise previously reported by second follow-up sample members, and following up with the institutions so identified. Note that some institutions identified following transcript review will have previously been contacted during phase I or phase II and will be re-contacted in order to retrieve information for the newly linked students. Similarly, some institutions may be identified that were not included previously. These institutions will be contacted in order to retrieve information for the associated students.

**B.2.d Student Sample for Phases I and II**

All second follow-up sample members who reported attending, or having attended, one or more institutions in the institution sample – as identified using student self-reports – will be included in the first phase student sample. All second follow-up sample members who attend, or attended, one or more of the institutions in the second phase institution sample – as identified during processing of phase one and phase two transcripts – will be included in the second phase student sample.

**B.3 Methods for Maximizing Response Rates****B.3.a Locating**

The response rate for the HSLS:09 data collection is a function of success in two basic activities: locating the sample members and gaining their cooperation. Several locating methods will be used to find and collect up-to-date contact information for the HSLS:09 sample. Panel maintenance activities between rounds of data collection have been and will be conducted. Activities include contact information updates by matching the information to national change of address and phone number databases; requesting information updates from the sample members and their parents, providing \$10 to sample members for such updates as approved in April 2013 (OMB# 1850-0852 v.11); and pre-data collection intensive tracing for the main study sample for cases lacking any current address information. During the HSLS:09 second follow-up field test and main study, batch searches of national databases will be

conducted prior to the start of data collection. Follow-up locating methods will be employed for those sample members not found after the start of data collection. The methods chosen to locate sample members are based on the successful approaches employed in earlier rounds of this study as well as experience gained from other recent NCES longitudinal cohort studies.

Many factors will affect the ability to successfully locate and survey sample members for HSLs:09. Among them are the availability, completeness, and accuracy of the locating data collected in the prior interviews. The locator database includes critical tracing information for nearly all sample members and their parents, including address information for their previous residences, telephone numbers, and e-mail addresses. This database allows interviewers and tracers to have ready access to all of the contact information available and to new leads developed through locating efforts. To achieve the desired locating and response rates, a multistage locating approach that capitalizes on available data for the HSLs:09 sample will be employed. The proposed locating approach includes the following activities:

1. **Advance Tracing** includes batch database searches and contact information updates.
2. **Telephone Locating and Interviewing** includes calling all available telephone numbers and following up on leads provided by parents and other contacts.
3. **Pre-Intensive Batch Tracing** consists of the Premium Phone searches that will be conducted between the telephone locating and interviewing stage and the intensive tracing stage.
4. **Intensive Tracing** consists of tracers checking all telephone numbers and conducting credit bureau database searches after all current telephone numbers have been exhausted.
5. **Field Tracing and Interviewing** will be employed during the main study to pursue targeted cases (e.g., dropout students) that were unable to be reached and interviewed by other means.
6. **Other Locating Activities** will take place as needed and may include additional tracing resources (e.g., matches to Department of Education financial aid data sources) that are not part of the previous stages.

The steps described in the tracing plan are designed to locate the maximum number of sample members with the least expense. The most cost-effective steps will be taken first so as to minimize the number of cases requiring more costly intensive tracing efforts. Approved in April 2013, the address update (panel maintenance) procedure will be conducted to encourage sample members to update their contact information several months prior to the start of data collection. All cases will be offered an incentive of \$10 if they or a parent confirm or update sample member contact information via the study website prior to the start of data collection.

### **B.3.b Interviewing Procedures**

**Interviewer training.** Training for individuals working in survey data collection will include critical quality control elements. Contractor staff with experience in training interviewers will prepare the HSLs:09 second follow-up field test Telephone Interviewer Manual, which will provide detailed coverage of the background and purpose of the study, sample design, the questionnaire, and procedures for the telephone interview. The manual will also serve as a reference throughout data collection. Training staff will prepare training exercises, mock interviews (specially constructed to highlight the potential of definitional and response problems), and other training aids.

**Interviewing.** Interviews will be conducted using a single web-based survey instrument for self-administered and telephone data collection (as well as field data collection in the main study). The data collection activities will be accomplished through the Case Management System (CMS), which is equipped with the following capabilities:

- online access to locating information and histories of locating efforts for each case;
- questionnaire administration module with input validation 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;
  - sorting of non-appointment cases according to parameters and priorities set by project staff;
  - restriction on allowable interviewers;
  - complete records of calls and tracking of all previous outcomes;
  - flagging of problem cases for supervisor action or supervisor review; and
  - complete reporting capabilities.

The integrated system reduces the number of discrete stages required in data collection and data preparation activities. Overall, the scheduler provides a highly efficient case assignment and delivery function, reduces supervisory and clerical time, improves execution on the part of interviewers and supervisors by automatically monitoring appointments and callbacks, and reduces variation in implementing survey priorities and objectives.

In the main study, field interviews will also be conducted via Computer Assisted Personal Interviewing (CAPI) to increase participation among sub-groups of interest. Field interviewing can be an effective approach to locating and gaining cooperation from sample members who have been classified as “ever dropouts<sup>3</sup>” or who have not enrolled in postsecondary education. CAPI can also be effective in increasing response among other potentially underrepresented groups as identified by responsive design modeling. Approximately 6 weeks after the start of outbound calling, CAPI cases will be identified and field interviewers will be assigned in geographic clusters to ensure the most efficient allocation of resources. Field interviewers will be equipped with a full record of the case history and will draw upon their extensive training and experience to decide how best to proceed with each case.

**Refusal Conversion.** Recognizing and avoiding refusals is important to maximize the response rate. Supervisors will monitor interviewers intensively during the early weeks of data collection and provide retraining as necessary. In addition, supervisors will review daily interviewer production reports to identify and retrain any interviewers with unacceptable numbers of refusals or other problems. After encountering a refusal, comments are entered into the CMS record that include all pertinent data regarding the refusal situation, including any unusual circumstances and any reasons given by the sample member for refusing. Supervisors 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.

If follow-up to an initial refusal is not appropriate (e.g., there are extenuating circumstances, such as illness or the sample member firmly requested no further contact), the case will be coded as final and no additional contact

<sup>3</sup> Dropout students included students who stopped attending school for a period of 4 weeks or longer, not including early graduates or home-schooled students.

will be made. If the case appears to be a “soft” refusal (i.e., an initial refusal that warrants conversion efforts – such as expressions of too little time or lack of interest or a telephone hang-up without comment), 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 skilled at converting refusals. Refusal conversion efforts will be delayed until at least 1 week 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. Project staff sometimes receives refusals via email or calls to the project toll-free line. These refusals are included in the CATI record of events and coded as final when appropriate.

### **B.3.c Quality Control**

Interviewer monitoring will be conducted using Quality Evaluation System (QUEST) as a quality control measure throughout the field test and main study data collections. QUEST is a system developed by a team of RTI researchers, methodologists, and operations staff focused on developing standardized monitoring protocols, performance measures, evaluation criteria, reports, and appropriate system security controls. It is a comprehensive performance quality monitoring system that includes standard systems and procedures for all phases of quality monitoring, including obtaining respondent consent for recording, procedures for interviewing respondents who refuse consent, and for monitoring refusals at the interviewer level; sampling of completed interviews by interviewer and evaluating interviewer performance; maintaining an online database of interviewer performance data; and addressing potential problems through supplemental training. These features and procedures are based on “best practices” identified in the course of conducting many survey research projects.

As in previous studies, QUEST will be used to monitor approximately 7 percent of all completed interviews plus an additional 2.5 percent of recorded refusals. In addition, quality supervisors will conduct silent monitoring for 2.5 to 3 percent of budgeted interviewer hours on the project. This will allow real-time evaluation of a variety of call outcomes and interviewer-respondent interactions. Recorded interviews will be reviewed by supervisors for key elements such as professionalism and presentation; case management and refusal conversion; and reading, probing, and keying skills. Any problems observed during the interview will be documented on problem reports generated by QUEST. Feedback will be provided to interviewers and patterns of poor performance (e.g., failure to use active-listening interviewing techniques, failure to probe, etc.) will be carefully monitored and noted in the feedback form that will be provided to the interviewers. As needed, interviewers will receive supplemental training in areas where deficiencies are noted. In all cases, sample members will be notified that the interview may be monitored by supervisory staff.

### **B.3.d Panel Maintenance**

A panel maintenance mailing is proposed to collect updated contact information in summer of 2015 prior to the second follow-up main study scheduled to take place in early 2016. Sample members and their parents will receive the panel maintenance mailings. The proposal for the panel maintenance activities is modeled after the activities approved for HSLS:09 in April 2013 (OMB# 1850-0852 v.11) and used successfully on ELS:2002/12.

ELS:2002/12 conducted an experiment with its field test sample that demonstrated the effectiveness of a \$10 incentive offer to increase participation in the panel maintenance. In this experiment, half of the sample members in the field test sample were offered a \$10 check if they or their parents confirmed or updated their contact information. No incentive was offered to the other half of the sample. A cost-benefit analysis was also conducted to evaluate the difference between the cost of the incentive offer and the difficulty of cases that responded. The impetus behind this analysis was determining if information was received from more difficult cases, as the benefit would be reduced if the “easy-to-track” cases were the ones to respond. Overall, the \$10 treatment group had a higher participation rate (25 percent) than the control group (20 percent,  $t=1.90$ ,  $p < .05$ ). Higher panel maintenance participation for the treatment group as compared with the control group was also observed by various characteristics of cases, such as those with postsecondary education experience, males, and those with a high school diploma. Further evaluation of the data indicates that the contact information provided largely new information not already in the study database; for 82 percent of the responding cases, at least one new address, phone number, or email address was provided for the student, parent, or both. Being able to make direct contact with the sample student during data collection saves time and costs, and is likely to increase interview participation.

The \$10 incentive offer was then implemented with the ELS:2002/12 main study sample based on the field test results. Forty percent of the ELS:2002/12 third follow-up main study sample participated in panel maintenance at some point. Among that group, 97 percent responded to the third follow-up survey compared to a 74 percent response rate among those that did not participate in panel maintenance. Among those that responded to one or more panel maintenance requests, the ELS:2002/12 survey response rate exceeded 90 percent across numerous categories including but not limited to: those without known postsecondary experience (91%), those without a regular high school diploma (93%), males (96%), and ever dropout cases (96%). HSLS:09 proposes to implement the ELS:2002 model for the 2015 panel maintenance activity and offer a \$10 incentive to sample members who provide updated contact information.

There is also a panel maintenance activity planned for early 2018. It would involve the same procedures but would include both the field test and main study samples in order to collect updated information for use in locating the sample members in future collections using the HSLS:09 cohort.

### ***B.3.e Postsecondary education transcript study and financial aid record collection***

The success of the HSLS:09 postsecondary education transcript study (HSLS:09 PETS) and financial aid record collection (HSLS:09 FAR) is fully dependent on the active participation of sampled institutions. The cooperation of an institution’s coordinator is essential as well, and helps to encourage the timely completion of the data collection. Telephone contact between the project team and institution coordinators provides an opportunity to emphasize the importance of the study and to address any concerns about participation.

**Proven Procedures.** HSLS:09 procedures for working with institutions will be developed from those used successfully in other studies with transcript collections such as the ELS:2002 Postsecondary Education Transcript Study and the 2009 Postsecondary Education Transcript Study (PETS:09) that combined transcript collections for



the B&B:08/09 and BPS:04/09 samples, as well as studies involving student records data collection from institutions, specifically, ELS:2002 Financial Aid Feasibility Study, NPSAS:08, and NPSAS:12. HSLS:09 will use an institution control system (ICS) similar to the system used for ELS:2002 PETS to maintain relevant information about the institutions attended by each HSLS:09 cohort member. Institution contact information obtained from the Integrated Postsecondary Education Data System (IPEDS) will be loaded into the ICS, then used for all mailings and confirmed during a call to each institution verifying the name of the registrar and financial aid director (or other appropriate contact), address information, telephone and facsimile numbers, and email addresses. This verification call will help to ensure that transcript and student records request materials are properly routed, reviewed, and processed.

**Endorsements.** In past studies, the specific endorsement of relevant associations has been extremely useful in persuading institutions to cooperate. Endorsements from 17 professional associations were secured for PETS:09. Appropriate associations to request endorsement for HSLS:09 PETS and the HSLS:09 FAR collections will be contacted as well; a list of potential associations is provided in appendix F.

**Minimizing burden.** Different options for collecting data for sampled students are offered. Institution staff 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.

Another strategy that has been successful in increasing the efficiency of institution data collections, encouraging participation, and minimizing burden is soliciting support at a system-wide level rather than contacting each institution within the system separately. A timely contact, together with enhanced verification procedures, is likely to reduce the number of remail requests, and minimize delay caused by misrouted requests.

### ***B.3.f HSLS:09 Postsecondary Education Transcript Study (HSLS:09 PETS) collection***

NCES recently completed ELS:2002 PETS, a collection of approximately 20,600 postsecondary transcripts, and previously completed PETS:09, a collection of approximately 45,000 postsecondary transcripts for the BPS:04/09 and B&B:08/09 samples. The same processes and systems found to be effective in those studies will be adapted for use for the HSLS:09 PETS. They are described below.

Data request materials and prompting. Transcript data will be requested for sampled students from all institutions attended since high school. The descriptive materials sent to institutions to request data will be clear, concise, and informative about the purpose of the study and the nature of subsequent requests. The package of materials sent to the transcript coordinators, provided in appendix F, will contain:

- A letter from RTI providing an introduction to the HSLS:09 PETS;
- An introductory letter from NCES on U.S. Department of Education letterhead;
- Letter(s) of endorsement from supporting organizations/agencies (e.g., the American Association of Collegiate Registrars and Admission Officers);
- A list of other endorsing agencies;

- Information regarding how to log on to the secure NCES postsecondary data portal website and access the list of students for which transcripts are requested as well as a form in which they can request reimbursement of expenses incurred with the request (e.g., transcript processing fees); and
- Descriptions of and instructions for the various methods of providing transcripts.

The institutions will also be asked to participate in the HSLs:09 FAR by providing financial aid student records data (described in greater detail in a separate section below).

Follow-up calls to ensure receipt of the packet and answer any questions about the study will occur 2 days after the initial phase I mailing. Telephone prompting also will be required to obtain the desired number of transcripts. Despite the relatively routine nature of the transcript request, many institutions give relatively low priority to voluntary research requests. Telephone follow-up is necessary to ensure that the request is handled within the schedule constraints. In addition to telephone prompting, institutions will be contacted by email prompts, letters, and postcard prompts. Experienced staff from RTI's Research Operations Center will carry out these contacts and will be assigned a set of institutions that is their responsibility throughout the process. This allows RTI staff members to build a relationship and maintain rapport with the institution staff and provides a reliable point of contact at RTI. Project staff members will be thoroughly trained in transcript collection and in the purposes and requirements of the study, which helps them establish credibility with the institution staff. Because institution coordinators are a critical element in this process, communicating instructions about their transcript collection tasks clearly is essential.

**Data submission options.** Several methods will be used for obtaining the data including: (1) institution staff uploading electronic transcripts for sampled students to the secure NCES postsecondary data portal website; (2) institution staff sending electronic transcripts for sampled students by secure File Transfer Protocol; (3) institution staff sending electronic transcripts as encrypted attachments via email; (4) RTI requesting/collecting electronic transcripts via a dedicated server at the University of Texas at Austin for institutions that already use this method; (5) institution staff sending electronic transcripts via eSCRIP-SAFE™, in which institutions send data to the eSCRIP-SAFE™ server by secure internet connection where they can be downloaded only by a designated user; (6) institution staff transmitting transcripts via a secure electronic fax after a test submission of nonsensitive data confirms that the institution has the correct fax number; and as a last resort, (7) sending transcripts via Federal Express. Each method is described below.

A complete transcript from the institution will be requested as well as the complete transcripts from transfer institutions that the students attended, as applicable. To track receipt of institution materials and student transcripts, a Transcript Control System (TCS) will be added to the IMS and refined for HSLs:09 PETS. The TCS will track the status of each catalog and transcript request, from initial mailout of the requests through follow up and final receipt.

**Uploading electronic transcripts to the secure NCES postsecondary data portal website.** Goals for HSLs:09 PETS include reducing the data collection burden on institutions (thereby reducing project costs), expediting data delivery, improving data quality, and ensuring data security. Because the open internet is not

conducive to transmitting confidential data, any internet-based data collection effort necessarily raises the question of security. However, the latest technology systems will be incorporated into the web application to ensure strict adherence to NCES confidentiality guidelines. The web server will include a Secure-Sockets Layer (SSL) Certificate, and will be configured to force 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 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.

Files uploaded to the secure website will be immediately moved to a secure project folder that is only accessible to specific 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. 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 loss of data. ITS will use their standard procedures for backing up data, so the backup files will exist for 3 months.

**Institution staff sending electronic transcripts by secure File Transfer Protocol.** FTPS (also called FTP-SSL) uses the FTP protocol on top of SSL or TLS. When using FTPS, the control session is always encrypted. The data session can optionally be encrypted if the file has not been pre-encrypted. Files transmitted via FTPS will be copied to a secure project folder that is only accessible to specific staff members. As with uploaded files, 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 immediately deleted from the FTP server. 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 loss of data. ITS will use their standard procedures for backing up data, so the backup files will exist for 3 months.

**Institution staff sending electronic transcripts as encrypted attachments via email.** RTI will provide guidelines on encryption and creating strong passwords. Encrypted electronic files sent via email to a secure email folder will only be accessible to a few staff members on the project 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 email folder. The files will be stored on the network that is backed up regularly to avoid the need to recontact the institution to provide the data 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 3 months.

**Institution staff sending electronic transcripts via eSCRIP-SAFE™.** This method involves the institution sending data via a customized print driver which connects the student information system to the eSCRIP-

SAFE™ server by secure internet connection. RTI, as the designated recipient, can then download the data after entering a password. The files are deleted from the server 24 hours after being accessed. The transmission between sending institutions and the eSCRIP-SAFE™ server is protected by Secure Socket Layer (SSL) connections using 128-bit key ciphers. Remote access to the eSCRIP-SAFE™ server via the Web interface is likewise protected via 128-bit SSL. Downloaded files will be moved to a secure project folder that is only accessible to specific 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. 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 loss of data. ITS will use their standard procedures for backing up data, so the backup files will exist for 3 months.

**Institution staff transmitting transcripts via a secure electronic fax (e-fax).** It is expected that few institutions will ask to provide hardcopy transcripts. In such cases, institutions will be encouraged to use one of the secure electronic methods of transmission. If that is not possible, faxed transcripts will be accepted. 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, the protocol will only allow for transcripts to be faxed to an electronic fax machine and only if institutions cannot use one of the other options. To ensure the fax transmission is sent to the appropriate destination, a test run with nonsensitive data will be required prior to submitting the transcripts to eliminate errors in transmission from misdialing. Institutions will be given a fax cover page that includes a confidentiality statement to use when transmitting individually identifiable information.

Transcript data received via e-fax are stored as electronic files on the e-fax server, which is housed in a secured data center at RTI. These files will be copied to a project folder that is only accessible to project 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-fax server. The files will be stored on the network that is backed up regularly to avoid the need to recontact the institution to provide the data 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 3 months.

**Institution staff sending transcripts via Federal Express.** When institutions will ask to provide hardcopy transcripts, they will be encouraged to use one of the secure electronic methods of transmission or fax. If that is not possible, transcripts sent via Federal Express will be accepted. Before sending, institution staff will be instructed to redact any personally identifiable information from the transcript including student name, address, data of birth, and Social Security Number (if present). Paper transcripts will be scanned and stored as electronic files. These files will be stored in a project folder that is only accessible to project 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. The files will be stored on the network that is backed up regularly to avoid the need to re-contact the institution to provide the data 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 3 months. The original paper transcripts will be shredded.

**Collecting electronic transcripts via a dedicated server at the University of Texas at Austin.**

Transcripts will also be requested electronically via a dedicated server at the University of Texas at Austin for institutions that currently use this method. Approximately two hundred institutions are currently registered to send and receive academic transcripts in standardized electronic formats via a dedicated server at the University of Texas at Austin. The server supports Electronic Data Interchange (EDI) and XML formats. Additional institutions are in the test phase with the server, which means that they are preparing and testing using the server but not currently using it to send data. The dedicated server at the University of Texas at Austin supports the following methods of securely transmitting transcripts:

- email as MIME attachment using PGP encryption
- regular FTP using PGP encryption
- Secure FTP (SFTP over ssh) and straight SFTP
- FTPS (FTP over SSL/TLS)

Files collected via this dedicated server will be copied to a secure project folder that is only accessible to specific staff members. The same access restrictions and storage protocol will be followed for these files as described above for files uploaded to the NCES postsecondary data portal website.

Active student consent for the release of transcripts will not be required for HSLS:09 PETS. For certain agency purposes, the Family Educational Rights and Privacy Act of 1974 (FERPA) (34 CFR Part 99) permits institutions to release student data to the Secretary of Education and his authorized agents without consent. In compliance with FERPA, a notation will be made in the student record that the transcript has been collected for use only for statistical purposes in the HSLS:09 longitudinal study.

Despite the relatively routine nature of the transcript request, it is anticipated that telephone prompting will be required to obtain the desired number of transcripts. Email prompts, letters, and postcard prompts, which have proven to be effective tools in gaining cooperation, will also be used. Email, in particular, has proven to be a low-cost and effective means of reaching institution officials who cannot be reached by phone. Because institutions can request reimbursement for expenses incurred in handling the request, it is unlikely that refusals will become a significant problem. However, in the event that an institution expresses resistance to the transcript request, seasoned institutional contactors and other project staff are trained to sensitively listen to institutional concerns, address any roadblocks to participation, and negotiate with institution staff to resolve them.

**Quality control and initial processing.** As part of quality control procedures, the importance of collecting complete transcript information for all sampled students will be emphasized to registrars. Transcripts will be reviewed for completeness. Institutional Contactors will contact the institutions to prompt for missing data and to resolve any problems or inconsistencies.

Transcripts received in hardcopy form will be subject to a quick review prior to recording their receipt. Receipt control clerks will check transcripts for completeness and review transmittal documents to ensure that transcripts have been returned for each of the specified sample members. The disposition code for transcripts received will be entered into the TCS. Course catalogs will also be reviewed and their disposition status updated in the system in cases where this information is necessary and not available through CollegeSource Online. Hardcopy course catalogs will be sorted and stored in a secure facility at RTI, organized by institution. The procedures for electronic transcripts will be similar to those for hardcopy documents—receipt control personnel, assisted by programming staff, will verify that the transcript was received for the given requested sample member, record the information in the receipt control system, and check to make sure that a readable, complete electronic transcript has been received.

The initial transcript check-in procedure is designed to efficiently receipt returned materials into the TCS as they are received each day. The presence of an electronic catalog (obtained from CollegeSource Online) will be confirmed during the verification process for each institution and noted in the TCS. The remaining catalogs will be requested from the institutions directly and will be receipted in the TCS as they are received. Transcripts and supplementary materials received from institutions (including course catalogs) will be inventoried, assigned unique identifiers based on the IPEDS ID, reviewed for problems, and receipted into the TCS.

Data processing staff will be responsible for (1) sorting transcripts into alphabetical order to facilitate accurate review and receipt; (2) assigning the correct ID number to each document returned and affixing a transcript ID label to each; (3) reviewing the materials to identify missing, incomplete, or indecipherable transcripts; and (4) assigning appropriate TCS problem codes to each of the missing and problem transcripts plus providing detailed notes about each problem to facilitate follow-up by Institutional Contactors and project staff. Project staff will use daily monitoring reports to review the transcript problems and to identify approaches to solving the problems. Web-based collection will allow timely quality control, as RTI central staff will be able to monitor data quality for participating institutions closely and on a regular basis. When institutions call for technical or substantive support, the institution's data will be queried in order to communicate with the institution much more effectively regarding any problems. Transcript data will be destroyed or shredded after the transcripts are keyed, coded, and quality checked.

**Transcript Keying and Coding.** Once student transcripts and course catalogs are received, and missing information is collected, keying and coding of transcripts and courses taken will take place. As part of PETS:09, the course classification structure used on NELS:88 has been updated and enhanced. The result, the 2010 College Course Map<sup>4</sup>, is a hybrid PETS coding taxonomy that makes it easier for Keyer-Coders (KCs) to select an appropriate code for the courses they identify on transcripts. This taxonomy was used for coding ELS:2002 PETS and will be used for HSL:09 PETS. KCs will have full access to all transcript-related documents including course

---

<sup>4</sup> Bryan, M. & Simone, S. (2012). *2010 College Course Map* (NCES 2012-162REV). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. Retrieved from <http://nces.ed.gov/pubsearch>.

catalogs or other course listings provided. All transcript-related documents will be thoroughly reviewed before data are abstracted from them.

**Transcript Keying and Coding Quality Control.** A comprehensive supervision and quality control plan will be implemented during transcript keying and coding. At least one supervisor will be onsite at all times to manage the effort and simultaneously perform QC checks and problem resolution. Verifications of transcript data keying and coding at the student level will be performed. Any errors will be recorded and corrected as needed. Once the transcripts for each institution are keyed and coded, transcript course coding at the institution level will be reviewed by expert coders to ensure that (1) coding taxonomies have been applied consistently and data elements of interest have been coded properly within institutions; (2) program information has been coded consistently according to the program area and sequence level indicators in course titles; (3) records of sample members who attended multiple institutions do not have duplicate entries for credits that transferred from one institution to another; and (4) additional information has been noted and coded properly.

### ***B.3.g HSLs:09 Financial aid record collection (HSLs:09 FAR) collection***

Several options will be offered to institutions for providing the financial aid student records for students, similar to those used for ELS:2002 FAFS and NPSAS:12 and invite the institution coordinator to select the methodology that is least burdensome and most convenient for the institution. The optional methods for providing financial aid student record data are described below.

**Student Records obtained via a web-based data entry interface.** The web-based data entry interface allows the coordinator to enter data by student, by year.

**Student Records obtained by completing an Excel workbook.** An Excel workbook will be created for each institution and will be preloaded with the sampled students' ID, name, date of birth, and SSN (if available). To facilitate simultaneous data entry by different offices within the institution, the workbook contains a separate worksheet for each of the following topic areas: Student Information, Financial Aid, Enrollment, and Budget. The user will download the Excel worksheet from the secure NCES postsecondary data portal website, enter the data, then upload the data to the website. Validation checks will occur both within Excel as data are entered and when the data are uploaded via the website. Data will be imported into the web application such that institution staff can quality control their data.

**Student Records obtained by uploading CSV (comma separated values) files.** Institutions with the means to export data from their internal database systems to a flat file may use this method of supplying financial aid records. Institutions that select this method will be provided with detailed import specifications, and all data uploading will occur through the secure NCES postsecondary data portal website. Like the Excel workbook option, data will be imported into the web application such that institution staff can check their data before finalizing.

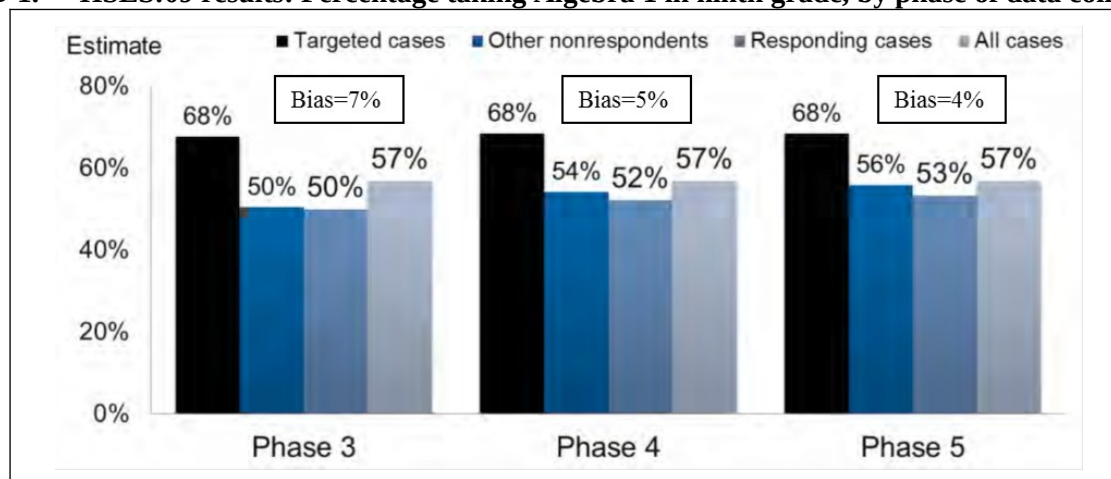
Institution coordinators will receive a guide that provides instructions for accessing and using the website. In conjunction with the transcript collection, institution contacting staff at RTI will make initial telephone calls to

notify institutions that the financial aid student records data collection has begun. Using daily status reports that summarize the progress of the institutions, staff will also call institutions periodically to prompt completion of the financial aid student records collection. Help desk project staff are available by telephone or by email to provide assistance if institution staff have questions or encounter problems.

#### B.4 Field Test Experiments

**Background.** NCES has had recent success with the responsive design approach in which nonresponding sample members believed to be most likely to contribute to nonresponse bias in different estimates of interest are identified at multiple points during data collection and targeted with changes in protocol to increase their participation and to reduce nonresponse bias. Exhibit B-1 presents an example of evidence that this approach helped bring respondent estimates in line with sample frame estimates by targeting underrepresented groups in the HSLS:09 2013 Update; in this example it was the sub-group of students who took Algebra 1 in the ninth grade, but similar results were observed for other variables of interest. At the beginning of Phase 3 (a few weeks after the start of outbound calling), some 50 percent of respondents had taken Algebra 1 in the ninth grade, 7 percentage points lower than the percentage in the overall sample (57 percent)—representing a nonresponse bias of 7 percent. Among the targeted cases, however, approximately 68 percent had taken Algebra 1 in the ninth grade. These cases had been identified through the modeling process, and response among this group was increased through the use of targeted interventions. Through the responsive design targeted intervention approach, the nonresponse bias associated with the estimated percentage of students who took Algebra 1 in the 9<sup>th</sup> grade was reduced to 4 percentage points by the beginning of Phase 5, as illustrated in Exhibit B-1 (Pratt 2014).

**Exhibit B-1. HSLS:09 results: Percentage taking Algebra 1 in ninth grade, by phase of data collection.**



**Proposed Design.** The HSLS:09 2013 Update demonstrated that the models used to identify sample members who are underrepresented with regard to the key survey variables were successful. Likewise, the interventions used (e.g., prepaid incentives, increasing incentive offers) among the cases targeted based on the model results also seemed effective in encouraging cooperation among targeted cases, although they were not evaluated experimentally. In order to adequately assess the effectiveness of specific interventions used among targeted cases, a randomized design is



proposed to allow for experimental evaluation of interventions in the HSLS:09 second follow-up field test. Due to the small sample size, however, the interventions will be distributed based purely on assignment to treatment condition, and not based on some predetermined threshold from the model. In other words, responsive design will not be implemented *during* the field test; however, the models will be run at various points throughout data collection (e.g., weekly) so that post hoc analyses can be conducted of the effectiveness of interventions in relation to the cases that would have been targeted for treatment, as will be done in the main study.

The models that were developed and evaluated in the 2013 Update to identify cases for targeted treatments will be refined, modifying them as appropriate by making use of any relevant data obtained as part of the 2013 Update and high school transcript collection. The HSLS:09 second follow-up field test sample is relatively small (1,100 cases), therefore efforts will be focused on tests of interventions. The assessment of different interventions in a full factorial experimental design will inform the most effective and cost-efficient treatments to be used in the main study. These interventions are described in more detail below.

In the second follow-up field test, four interventions will be experimentally tested:

- Prepaid incentive of \$5 with the start of data collection mailing vs. providing a prepaid incentive of \$5 after several weeks of outbound telephone calls (approximately 6 weeks later);
- Initial contingent incentive of \$15 vs. no initial contingent incentive;
- “Incentive boost” (an additional contingent incentive to supplement the initial offer) in one of two values (\$15 or \$30) with the control treatment of no incentive boost, approximately eleven weeks into data collection; and
- Final response conversion in the case of low response rate, comparing two treatments. Should response rate be below a threshold at which the main study data would be possibly of poor quality or might not be able to be released (<70%), an additional treatment would be needed. To identify an effective treatment at the end of the study, all remaining nonrespondents would receive either 1) an additional \$25 (for some sample members it would be the only) contingent incentive or 2) an abbreviated instrument.

The schedule of interventions is outlined in exhibit B-2. For ease of presentation, the first two interventions are presented as four distinct groups. A full factorial experimental design ( $2 \times 2 \times 3 \times 2$ ) is proposed as some interventions may interact; for example, the baseline contingent incentive may be most effective when coupled with a prepaid incentive. The 1,100 sample cases will be allocated equally across the treatment conditions. To facilitate analysis, cells will be collapsed when warranted based on the results, to analyze the results sequentially (e.g., outcomes by week 10 are not affected by the contingent incentive treatment in week 11, and the outcomes at the point that the final treatments are initiated are not affected by the final treatment), and to leverage the rich frame data for regression-based analysis of the experiment. The latter approach uses available auxiliary data as covariates in the model evaluating the effect of the treatment conditions, reducing the error variance. The effect of this approach is analogous to increasing the sample size as it reduces the sampling variance, to the extent that the auxiliary data are correlated with the response outcome. Since variables such as “ever dropout” are highly correlated with nonresponse in the HSLS:09 sample, this approach can enhance the analyses of the field test.

**Exhibit B-2. Schedule of interventions in the HSLs:09 Second Follow-Up Field Test**

Group Assignment	Group A			Group B			Group C			Group D		
Baseline Contingent Incentive and Timing of \$5 Prepaid Incentive	No baseline contingent incentive			Baseline contingent incentive amount: \$15			No baseline contingent incentive			Baseline contingent incentive amount: \$15		
	Timing of \$5 prepaid incentive: week 6			Timing of \$5 prepaid incentive: week 6			Timing of \$5 prepaid incentive: with start of data collection mailing			Timing of \$5 prepaid incentive: with start of data collection mailing		
Contingent Incentive boost (in addition to baseline promised incentive) (Week 11)	No boost (n=41)	\$15 (n=41)	\$30 (n=41)	No boost (n=50)	\$15 (n=50)	\$30 (n=50)	No boost (n=41)	\$15 (n=41)	\$30 (n=41)	No boost (n=50)	\$15 (n=50)	\$30 (n=50)
<b>Final Treatment -- all remaining nonrespondents</b>												
(Weeks 13 and beyond)	Random assignment of nonrespondents to either \$25 additional contingent incentive or an abbreviated interview											

Note: It is expected that a main effect of five percentage points (i.e., 45% vs. 50%) in response rates due to offering a baseline incentive or the timing of the prepaid incentive could be detected at  $\alpha=.05$  with  $\text{power}=.84$  (one-tailed test). For the contingent incentive boost, it should be possible to detect a difference of five percentage points (i.e., 59% vs. 64%) at  $\alpha=.05$  with  $\text{power}=.47$ , between two of the three conditions (one-tailed test).

This plan tests the timing of a \$5 prepaid incentive that will be provided in addition to the contingent incentive. In an experiment for B&B:08/12, a \$5 prepaid incentive resulted in significantly higher rates of response among targeted cases. However, the B&B:08/12 prepaid incentive was not implemented until mid-way through data collection. It is possible that a prepaid incentive is most effective when offered in the first communication with the sample member at the start of data collection. If this were the case, the cost of sending the prepaid incentive to more sample members could possibly be offset by reducing the need for more costly interventions (e.g., telephone interviewing and field interviewing) later in data collection. To test this, the offering of the \$5 prepaid incentive initially will be compared to the offering of a \$5 prepaid incentive after a few weeks of outbound calling (week 6).

A baseline contingent incentive of \$15 will be assessed in comparison with a baseline no-contingent-incentive control condition in order to gather evidence on whether some sample members in this cohort will complete the survey without any baseline contingent incentive. Incentive boosts (including a no-incentive-boost control condition) will be tested later in data collection for conversion of the more reluctant sample members.

At week 11 of data collection, the effectiveness of a boost in promised incentives will be tested. A test is proposed to evaluate whether 1) the offer of an incentive boost improves response rates (no boost vs  $> \$0$ ) and 2) whether a lower-boost amount (\$15) can be as effective as a higher-boost amount (\$30).

Incentives are an effective tool to motivate sample members to participate in the survey, but for sample members receiving various incentive offers, other interventions may be more effective towards the end of data collection. Another option for an intervention is to decrease respondent burden by offering an abbreviated instrument. Therefore, in the final few weeks of the HSLs:09 field test, remaining nonrespondents will be offered one of two final

treatment conditions: an abbreviated instrument or additional incentive of \$25. This treatment is not intended to be a necessary part of the main study data collection. Its objective is to identify possible effective treatments at the end of data collection, to be implemented only if response rate targets on the main study are not achieved.

**Main study plans.** NCES and RTI are working closely together to make use of findings from evaluations of interventions used to reduce potential nonresponse bias and to increase response. Plans for the HSLs:09 second follow-up main study will be based upon 1) results from the HSLs:09 second follow-up field test, 2) results from related longitudinal studies (such as BPS:12/14, B&B:08/12, and ELS:2002/12), and 3) prior experience with the HSLs:09 cohort.

As an example of an approach based on prior experience with the HSLs:09 cohort, sample members who had ever dropped out of high school will be treated as a separate group apart from the rest of the sample. In the 2013 Update, all “ever dropouts” were offered a \$40 contingent incentive with the initial data collection invitation. Other treatments for the “ever dropout” group included a \$5 prepaid incentive at the start of phase 3 of data collection (a few weeks after the start of outbound calling) and an abbreviated interview for the last few weeks of data collection. This differential treatment of the “ever dropout” group was successful; the resulting response rate for this group was 82 percent, compared to 80 percent for the overall response rate (see Pratt, 2013). Based on the success in the 2013 Update (as well as similar success with the ELS:2002 cohort), the same approach will be used for “ever dropouts” in the second follow-up main study. As another example, the ELS:2002 third follow-up data collection included outbound CATI for the “ever dropout” group approximately three weeks earlier than for other sample members, and also implemented field interviewing (CAPI) with the “ever dropouts” as an additional intervention.

Main study plans for the remaining sample (those not in the “ever dropout” group) will be informed by results of the HSLs:09 second follow-up field test, the HSLs:09 2013 Update, and the most recent evidence collected from BPS:12/14. For instance, the proposed design is similar to that implemented in the BPS:12/14 full-scale study that randomly selected a small subset of the sample for experimentation on interventions. This experimental sample was treated in advance of the remaining cases, and after analyzing the results for the experimental sample and consultation with OMB, the successful treatment was applied to the targeted cases in the remaining sample. In the HSLs:09 second follow-up main study, a similar subset of cases (somewhere between 5 and 10 percent of the main study sample) will be used to experimentally evaluate the impact of various treatments. After analyzing the results and consulting with OMB, treatments that are successful in the experimental evaluation will be implemented for the remaining main study sample cases. As an example of this approach, BPS:12/14 tested an incentive boost of \$0 (no boost), \$25, and \$45 for targeted nonrespondents. Results indicated that only the \$45 incentive boost significantly increased response among targeted cases, and the \$45 incentive boost was used for targeted cases in the main sample.

In the HSLs:09 second follow-up main study, data collection for the experimental sample will begin approximately 7 weeks prior to the rest of the main sample, allowing the results of experiments to inform the main sample data collection. Treatments under consideration for main study implementation include: differing baseline incentive amounts, the timing of a \$5 prepaid incentive, differing incentive boost amounts, CAPI, and an extended

data collection for targeted cases (approximately 4 weeks, where participation is still open to all nonrespondents, but active telephone- and field- efforts for all but targeted cases will be discontinued).

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

The following statisticians at NCES are responsible for the statistical aspects of the study: Dr. Elise Christopher, Dr. Sean Simone, Dr. Jeffrey Owings, Dr. Chris Chapman, Dr. Marilyn Seastrom, Dr. Tracy Hunt-White, Dr. Sarah Crissey, and Mr. Ted Socha. The following RTI staff work on the statistical aspects of the study design: Daniel Pratt, Melissa Cominole, David Wilson, Steven Ingels, Andy Peytchev, and Jeff Rosen.

#### **RECENT RESPONSIVE DESIGN REFERENCES USING HSLS:09 DATA**

- Rosen, J. A., Murphy, J. J., Peytchev, A., Holder, T. E., Dever, J. A., Herget, D. R., & Pratt, D. J. (2014). Prioritizing low-propensity sample members in a survey: Implications for nonresponse bias. *Survey Practice*, 7(1), 1–8.
- Pratt, D. J. (Invited Speaker). (2014, March). *What is adaptive design in practice? Approaches, experiences, and perspectives*. Presented at FedCASIC 2014 Workshop Plenary Panel Session, Washington, DC.
- Pratt, D. J. (2013, December). *Modeling, prioritization, and phased interventions to reduce potential nonresponse bias*. Presented at Workshop on Advances in Adaptive and Responsive Survey Design, Heerlen, Netherlands.