

Institute of Educational Statistics
National Center for Education Statistics

NATIONAL INSTITUTE OF STATISTICAL SCIENCES
EXPERT PANEL REPORT

CONNECTING THE DOTS:
INTEGRATED SAMPLING APPROACH
FOR MULTIPLE SURVEYS

TABLE OF CONTENTS

| | |
|--|----|
| Executive Summary..... | 3 |
| Preface | 5 |
| Introduction | 6 |
| Background | 6 |
| Rationale & Context - NCES | 7 |
| Context – State, District, School, Student/Parent | 8 |
| Objectives..... | 9 |
| Concept | 10 |
| Part I: Technical Feasibility and Requirements | 11 |
| Essentials – Sampling Frame | 11 |
| Essentials – Expanded Sampling Frame | 12 |
| Essentials – Annual Cycle | 13 |
| Essentials – Stratification | 14 |
| Design Options | 15 |
| Design Considerations..... | 19 |
| Part II: Implementation and Impact..... | 21 |
| Recruitment Paradigm | 21 |
| Advance Messaging..... | 23 |
| Messages..... | 24 |
| Foundation for Relationship | 25 |
| Process Priorities..... | 26 |
| Balance of Burden, Benefit, Capacity..... | 28 |
| Considerations | 30 |
| Parts I & II: Principal Findings and Recommendations | 32 |
| Appendix A: Agenda – Panel I | 36 |
| Appendix A: Agenda – Panel II | 37 |
| Appendix B: Charge to Panels | 38 |
| Appendix C: Expert Panel Biosketches | 39 |

NATIONAL INSTITUTE OF STATISTICAL SCIENCES

CONNECTING THE DOTS: INTEGRATED SAMPLING APPROACH FOR MULTIPLE SURVEYS I: SAMPLING DESIGNS & II: IMPLEMENTATION

DECEMBER 2021

Executive Summary

The title “*Connecting the Dots*” was chosen to describe the potential for improving the available information on US education by coordinating the efforts across surveys. The fourfold objectives are simultaneously to streamline the recruitment process at state/district/local levels, to reduce burden especially at district/local levels, to diminish non-response, and to leverage data from multiple data sources both at NCES and in other federal data collections.

Conceptually, integrating the sampling approach across multiple surveys would enable presentation annually of a single, compiled package of NCES survey proposals to the decision-maker at each level for approval of surveys individually. This approach would separate the recruitment process, removing it from individual survey contracts which currently handle recruitment independently for each survey.

Two expert panels considered the essential components to a strategy for restructuring multiple surveys into a coordinated process. They then weighed the requirements and the merits of going forward. *Part I* engaged statistical experts who considered the technical issues and feasibility of creating integrated or coordinated sampling plans. The technical solution would have simultaneously to meet the statistical requirements for each individual survey, to avoid inadvertent sample duplications and to enable data leveraging across surveys to improve estimation. *Part II* engaged recruiters (contractors), school district administrators (decision-makers) and social scientists to explore implementation and the impacts of such integrated sample designs on the recruitment process from each of their varied perspectives.

Both *Part I* and *Part II* panels found a coordinated approach to be highly desirable and feasible. They also determined that the work to accomplish this transition would be substantial and a successful strategy is likely to be complicated. Agreement was unanimous that the merits would be well worth the required work and investment.

Part I focused on the set of NCES surveys proposed for a single academic year and the strategy for a coordinated sampling plan. While a naïve approach would not scale beyond two (at most three) surveys, two distinct more computationally intensive strategies (or a combination) are scalable. For these strategies to be efficient, two resources are needed: a history of participation data base for schools, and compatibility with NAEP stratification and sampling design.

Part II placed the NCES surveys into the context of all proposed surveys submitted at each level: state/district/school. The picture differs greatly across states and districts with respect to demand, to

resources, even to the definitions of burden and benefits, and to the valuation of surveys (in the 100s for large districts) competing for school time.

A single package of proposals to review would identify the set of surveys as national with a single parent agency. Dealing with a single, continuing contact would improve dialogue while building a relationship of trust. The review process would be streamlined by eliminating repeated review and redundant discussion of critical issues (e.g., legal security and privacy documents) and by enabling recruitment materials to be passed on to the next level as a single process. The development of an annual calendar (even anticipating a future year) would benefit school calendar planning; it would also facilitate widely disseminated advance messaging. Since both burden and benefits affect response/non-response, understanding these is ultimately critical to overall success; both data and metrics are needed.

The complete findings and recommendations from *Part I* and *Part II* are combined below because the overlap was extensive and nowhere were these in conflict.

Part I and Part II Panels' Principal Findings

Coordination of NCES surveys and assessments is unequivocally desirable and feasible.

Accomplishing this will require extensive changes in several aspects and significant investment of effort and technical expertise to achieve integration across surveys moving toward standardization of critical survey components.

Building trust and open dialogues with gatekeepers at all levels is an immediate step and a continuing objective.

Initial steps can be taken before a comprehensive plan is complete.

Careful launch of these changes requires energetic communication and advance preparation of the education community, also coordination with OMB and current and potential survey contractors.

NATIONAL INSTITUTE OF STATISTICAL SCIENCES EXPERT PANEL REPORT

PREFACE

The title “*Connecting the Dots*” was used as a description of the potential to improve the available information on US education by coordinating the efforts across surveys to diminish non-response and simultaneously to leverage data from multiple data sources both at NCES and in other federal data collections.

In December 2020, NISS convened an expert panel to explore the opportunities and the expectations for surveys set to be conducted following emergence from the Covid period. The central concept arising from this panel’s deliberations and discussions with NCES staff was for IES/NCES to:

Implement a recognized functionally coherent and transparent structure to replace/reorganize the loosely connected collection of separate surveys and assessments.

The problem of rising rates of refusals and non-response at multiple levels further supports consideration of an integrated approach to the multiple surveys that NCES conducts each academic year.

In August 2021, NCES charged NISS with assembling experts to work with NCES staff to formulate a clear objective and a strategy for restructuring multiple surveys into a coordinated process and then to weigh the requirements and the merits of going forward. The task was divided into two parts with two groups of experts bringing appropriate expertise and different perspectives. In September, *Part I* considered the technical issues and feasibility of creating integrated or at least, coordinated sampling plans. The technical solution would have to simultaneously meet the statistical requirements for each individual survey, facilitate explicit management of sample duplications and enable data leveraging across surveys to improve estimation. In November, *Part II* explored implementation and the impacts of such integrated sample designs on the recruitment process from the perspectives of the contractors and of the decision-makers at all levels from state to district to school to individual respondent.

Each Part opened with a Working Session for NCES staff jointly with the assembled experts to define the problem, outline preliminary thoughts and set the agenda for the experts to discuss at subsequent closed panel sessions. Deliberations and recommendations from the panel were presented to NCES staff for discussion at a final meeting. This report presents these findings *from Connecting the Dots, I and II*.

NATIONAL INSTITUTE OF STATISTICAL SCIENCES

CONNECTING THE DOTS:
INTEGRATED SAMPLING APPROACH FOR MULTIPLE SURVEYS

INTRODUCTION

Background

As part of the US Department of Education, the National Center for Education Statistics (NCES) in the Institute of Education Sciences (IES) annually conducts a varying number of national surveys and assessments. The largest of these, NAEP (National Assessment of Education Progress), conducted every two years, is budgeted and managed separately and is the only K-12 survey/assessment with any mandated participation. For the others, including those mandated by Congress, participation is voluntary on the part of states, school districts, schools and respondents.

The NCES data collections are a rich source of national-scale information on a broad range of aspects of US education, including financial and administrative information as well as assessment and attainment. These data collections are widely used by educators, researchers, decision-makers at all levels and the general public. Consequently, assuring national representativeness (and in some cases state representativeness as well) depends on adequate levels of response to these voluntary surveys and assessments.

“Connecting the Dots” describes the exploration of the potential to improve the available information on US education by coordinating the efforts across surveys. The three-fold objectives are to reduce the burden on decision-makers and respondents at all levels (from state to personal), to decrease refusal and/or non-response, and simultaneously to leverage data from multiple data sources both at NCES and in other federal data collections. The goal is to formulate a clear objective and a strategy for restructuring multiple surveys into a coordinated process with a single integrated approach to recruitment that consolidates the recruitment for multiple surveys that NCES conducts each academic year.

In August 2021, NCES charged NISS with assembling experts to formulate and then evaluate options for restructuring multiple surveys into a coordinated process with a unified recruitment for all studies conducted during one academic year.

Rationale

As NCES seeks to explore avenues to improve and enhance the value of the information in their data collections, two previous expert panels¹ were commissioned to focus on new information needs and opportunities for leveraging existing information to enrich the NCES data collections.

The first panel examined new demands and new possibilities for sample designs, especially virtual data capture and information dissemination. The second panel then considered on the potential for incorporating or linking NCES data collections internally and with auxiliary information from other federal sources, particularly for drawing on GIS-linked data.

The primary recommendation to NCES from the first panel was:

Implement a recognized functionally coherent and transparent structure to replace/reorganize the loosely connected collection of separate surveys and assessments.

In addition, of continuing concern to NCES are the rising rates of refusals and non-response at multiple levels. This also supports consideration of an integrated approach to the multiple surveys that NCES conducts each academic year.

Following internal NCES discussions and a memorandum on ways to leverage current NCES surveys and to consolidate information gathered across surveys, Connecting the Dots, I and II was commissioned to continue the earlier explorations. Specifically, the charge was to outline a coordinated multi-survey sampling approach and then to evaluate the technical (statistical) feasibility and the practicality of implementing it.

Context - NCES

NCES conducts a variety of studies, including assessments and/or surveys each year, to provide national, and in certain cases state, data on education. Studies are of multiple types: longitudinal and cross-sectional, student participation and administrative record, education performance, performance and outcome, and economic/financial, national and international, pre-K to post-secondary levels. Of the assessments, only for NAEP (National Assessment of Education Progress) is response mandated (for certain entities – not for individual schools or students). Of the administrative information surveys, for public K-12 schools, information (updated annually) for the Common Core of Data is mandatory; for post-secondary training and education entities, participation in IPEDS is linked to eligibility for federal funding for students and/or educational entity.

With those exceptions, participation is voluntary at all levels from state to student. Hence, examining and improving the recruitment process at all levels is of high priority for NCES.

NOTE: NAEP is separately funded (a federal budget line item) and is managed by its own Governing Board. International studies have their own separate international governing bodies. Therefore, while

¹ Post-Covid Surveys (December 2020) and Priorities for Federal Data Access (March 2021)

both NAEP and the international studies are administered through NCES, they are only considered in this report as they contribute to the overall calendar of requests for participation or as they serve as information resources.

Currently, contracts are awarded by NCES for the actual conduct its surveys and assessments. Each study is contracted separately and handled individually with whichever contractor receives the award. For each study the Government specifies the timing within the academic calendar for administering the study; and it also specifies the particular requirements for this one study (e.g., sample size and precision for required estimates, whether state-level or solely national statistics, and/or by demographic subgroups, etc.). Subject to these specifications, which are included in an awarded contract, the contractor then develops the sample design; upon approved by NCES actually draws the primary and reserve samples.

From this point forward, the contractor takes over the recruitment process. That starts with contractor staff preparing a website, creating online and downloadable, as well as hard-copy, materials for decision-makers at every level including participants, and acquiring support from professional organizations, civic leaders and others. Informational packages are prepared for state offices of education requesting support for the study. On the administrative side, this also involves preparing of full proposals to go to school district superintendents and/or designated district officials., and coordinating NCES and state/district/local legal documents regarding privacy of personal information. Direct negotiation by contractor staff can then begin with decision-makers/gate-keepers, officially starting at the state level but with serious review usually delegated to the districts, then and working through sub-districts (if any) and school level to actual participants. Refusal at any level short-circuits the process and recruitment cannot proceed further with any school within a refusing district or any school (or district) when a state office refuses permission.

Each separate contractor for the various NCES studies independently contacts each office in the approval chain. So each contractor goes through the process of filing a proposal, negotiating with officials with respect to access, with respect to legal considerations, and with respect to any content issues in the survey instruments, at the same time making and presenting the case for participation in that particular study. Without coordination, there is some inevitable duplication of effort (e.g., privacy and data security requirements), some redundancy of information requested (e.g., similar information reported on similar or dissimilar scales), and some reduction in leverage for advocacy of NCES studies, particularly the loss of opportunity to develop a relationship of trust between a single recruiter and the offices in the approval chain for that state or district.

Context – State, District, School, Student/Parent

The review and approval process is sequential, starting at the state level. Upon receipt at the state department of education, and subsequently at the school district office, a survey/study proposal with all the required supporting documents joins the queue of other requests to be reviewed.

Typically, a proposal package sent to a school district for review contains survey-specific information (summary of survey purpose and design, all survey instruments, time/effort requirements, rosters of

sampled schools and schools in a reserve sample), legal documents (privacy and confidentiality measures, statements of personal information usage and rights, permission forms), and survey justification (research value, benefits/costs to participants, schools, . . . , nation). Often the package also includes letters of support (from administrators at higher levels, professional organizations, local representatives outside education). An increasing practice of school districts is to add local/individual requirements to the proposal package, particularly individual research agreements, Data Use Agreements or other Memoranda of Understanding.

States and large school districts each review many hundreds of proposals annually. At the state, and usually at the district, level, NCES proposals make up a small part of the queue (less than 10% at the state level). Some federal survey requests come from other parts of the Department of Education. More numerous requests come from Health and Human Services (including CDC) and NIH/NIMH/NCI, from the Census Bureau and from the Social Security Administration; while still others come from the other federal statistical agencies. States (and large school districts) initiate their own surveys; and localities may use surveys to address local issues. The category with the largest contribution to the queue is university research projects, including projects to meet graduate student program requirements.

Actual review procedures and decision criteria vary widely among states and between districts. States with large education infrastructure and large school districts may have their own research departments who oversee reviews or do so with a network of agreements with peer district research departments and/or with universities. Resources to devote to proposal reviews (both surveys and assessments) may be minimal for small state departments and small districts.

Value placed on *national* education statistics and hence on NCES studies, and even recognition (!), let alone preference, for NCES proposals interleaved into the queue also varies greatly among state and district offices. Prioritizing may be based on many factors. For example, priority is evaluated in terms of a burden-to-benefit ratio, focusing first on instructional time lost and then on other requirements for school resources or scheduling. Priority may also depend upon the remaining “available classroom time” after accounting for approval of studies in earlier reviews, upon individual school circumstances or stresses, upon established state/district priorities, upon the perceived importance of the research for a national understanding, and even upon the sheer number of proposals from students satisfying graduate program requirements and individual or small groups of academic researchers. With approval at one level, endorsements of importance and impact can give valued support for approval at the next level of review.

Objectives

The objective is to develop a strategy for an efficient multi-survey sampling design that expedites the recruitment process, streamlines the proposal review/approval process at state/district/local levels, improves participation and leads to improvement in data quality and information availability.

For NCES this means a strategy that accounts for burden at the school level, more specifically eliminating sample overlaps unless deliberately designed to link information between surveys and to

improve estimates. It also means working jointly with states/districts/schools based on a comprehensive academic year calendar to set up in advance a schedule that minimizes curriculum disruption and administrative workload burden. At NCES this implies compatibility of key information across surveys and consolidation of sampling frame (CCD) updating to leverage data across multiple surveys (including NAEP). Ultimately this will sustain a broader high quality national data base of information on education that supplies timely, relevant information and supports inquiry on more local (small-area) scales.

For districts and schools this means reassessing burden as both immediate and cumulative burden over multiple years, and as burden in the context of other challenges confronting the students in a school or district, and encompassing administrative load, curriculum/schedule disruption as well as classroom time. Specifically, multiple survey requests can largely be eliminated and, when unavoidable, can be coordinated with respect to time and/or data to reduce burden. With streamlining, the district (and state offices of education) can reduce consideration of the basic materials in multiple proposals to a single review. For schools, this strategy also means being able to incorporate survey and assessment scheduling into advance calendar planning of academic year events and activities. With reconsideration of burden-to-benefit balance, benefits may be adapted to individual school or district circumstances.

Concept

Four elements are key to consolidating multiple surveys into a single package:

- A sampling design for multiple surveys with minimal overlap can preserve important statistical properties of estimates and of data quality, both overall and for the surveys individually.
- Proposals and samples for all surveys during an academic year are assembled into a single package for individual review and approval of each survey.
- A single recruiter (as opposed to multiple contractors) would recruit and negotiate participation for all surveys at each level (state, district, school).
- The combined request would be presented early enough in the prior academic year to allow completion of the recruitment process in time for schools to plan their academic calendars to accommodate the surveys.

PART I: TECHNICAL FEASIBILITY AND REQUIREMENTS

The first step in determining whether and how to entertain the idea of a coordinated approach to multiple NCES surveys is to examine the technical feasibility.

The first panel was challenged to start with the existing sampling frames to envision how an integrated multi-survey recruitment process might be structured, focusing on the sampling design and the statistical requirements. Specifically, the panel was asked to identify and address:

- Key components and considerations for an effective and efficient consolidated process
- Technical issues, options to sampling and quantitative methodology solutions to support an integrated approach to multiple surveys
- Illustration, in outline, of how a consolidated process might work.

National Assessment of Educational Progress (NAEP), with participation federally mandated, is supported, governed and implemented independently of the remaining (all voluntary) NCES surveys and assessments. Consequently, NAEP is not included in this report as part of the coordinated planning; however NAEP can be a highly useful reference, especially for sampling frame information.

Essentials – Sampling Frame

The sample design starts with the sampling frame, in this case the Common Core Data (CCD) for public schools and districts. (Private School Universe Survey - PSS - is the parallel for private schools. Although not discussed further in this report, the ideas presented here would apply similarly to PSS.) The CCD contains the information necessary to stratify the public schools into more homogeneous subgroups so that sample inclusion probabilities are well-defined and can be used in producing statistical estimates, estimates of precision and evaluation of bias, especially bias due to non-response. For a multi-survey design, samples for all surveys during one academic year must be drawn from a common sampling frame – in this case, an updated version of the CCD that is fixed at one point in time. Preserving this version with its time-stamp creates a reference file that also permits data sharing across surveys.

Updating of the CCD occurs when samples drawn for NCES surveys/assessments, including NAEP, lead to corrections and whenever a school or district volunteers new information. It is desirable that information be updated on a continuous basis as it becomes available. In this way a consolidated update (at least annually) can provide a time-stamped permanent record to be used by all contractors in making post-stratification edits, in calculating statistical estimates and in planning surveys for a future academic year.

Permanent school (and district) IDs are necessary, with no adjustment for archiving/deletion of ID numbers when schools cease to operate. Permanent IDs will also facilitate calculation of burden and minimization of overlaps within and across years.

Recommendations: Ensure the completeness and the timeliness of the CCD, the primary sampling frame.

- Unify the updating process to be continuous, but with date-stamped versions for in-common use at sampling and in post-stratification adjustment.
- Invest in ongoing management and curation of CCD and other sampling frames, creating and preserving fixed (time-stamped) versions.

Essentials – Expanded Sampling Frame

The value of the CCD as both a sampling frame and an information source could be enhanced with two expansions.

The most important and ambitious expansion to the CCD would develop a restricted data appendix to the CCD to be useful in evaluating burden and in managing repeated resampling of individual schools. Data in this restricted appendix would include a comprehensive history for each school of its survey participation. Elements of a school's data record would be the list of surveys where the school was drawn in an original sample (primary sample, reserve sample or as part of oversampling), together with the outcome chronicle from immediate response at each level (state/district/school) through actual participation and survey completion with reasons specified for any refusal or non-participation at a later stage. These data are needed for quantitating burden, designing NCES surveys to manage overlapping requests, and minimizing, then compensating for, non-response.

This restricted data appendix added to the CCD (updated) will provide the information necessary for understanding burden of participation for schools and districts, and for developing meaningful metrics. These data will provide immediate information highlighting schools that are/have been subject to repeated sample inclusions. In addition, these data will enable a better understanding of non-response, the reasons and the patterns of refusal at each stage from first review (state) through the approval process, the survey implementation and actual completion of the survey/assessment. Since NAEP is arguably the most intense (generally cited by schools as they most burdensome of Department of Education K-12 surveys), it is critical that NAEP records be included in this restricted appendix.

Taking these individual school histories into account can be used to improve efficiency in terms of estimates and their precisions, survey management, and leverage across surveys. For individual sample designs for NCES surveys, the restricted data appendix is the key to avoiding/managing multiple survey requests to the same school or (small) district whether within a single academic year or cumulatively. This opens the door to sampling algorithms that take burden metrics into account – for example, potentially allowing future burden reduction as a benefit of current participation. The availability of participation histories can also enable efficient and broader use of adaptive sampling methodologies, of sample allocations based on (non-response) propensity scores and/or other statistical methodologic advances, and of refined stratification.

Compensating for non-response is accomplished by estimating non-response bias and calculating of bias adjustments. The accuracy of these adjustments depends on the relevance of the covariates used in the calculations. Validation of measures of the impact of non-response on coverage and/or bias depends on having an external source of information that is not subject to the same non-response patterns. Since

NAEP is mandatory for schools, it can serve as one external source of information for validation of coverage and representativeness of the NCES voluntary surveys.

The second expansion, to be publicly available would, as feasible, expand the available data by creating a publicly available appendix that incorporates or links to information in other federal data bases, particularly geographically-indexed information. Such information may currently be accessible for each school using NCES website tools like EDGE, but not be available for integration into the national-scale data base. Similarly, small area, relevant to local school features and contexts, may be available by individual calculation, but not on the larger scale that would be needed for other uses. Small area estimates of correlates would expand the base of information for calculation of metrics for representativeness and for improvement of non-response bias assessment and adjustments.

Recommendations: Expand the sampling frame, both open access CCD and a restricted access expansion to CCD.

- Create a restricted access CCD appendix file of comprehensive history of participation for districts and schools; include for all surveys/assessments (by level – state/district/school) the recruitment attempts, results, reasons for refusals or other non-response, and metrics for burden.
- Expand, as feasible, the open access CCD to incorporate or link to information from other federal agencies and to small area estimates of correlates.
- Moving forward, use the restricted CCD appendix for development of metrics for burden and validation of patterns and reasons of refusal and for estimating the impact of non-response, using NAEP data and external sources to validate.

Essentials – Annual Cycle

Survey preparation schedules are dictated by the dates for putting the surveys into the field, which in turn dictates the timing of the recruitment processes. Coordinating multiple surveys and unifying the recruitment process assume an in parallel preparation cycle for surveys to be conducted during one academic year. Schools ordinarily set their curricula and calendars to include planned dates for classroom exercises and semester testing, for field trips and special events, for teacher conferences and work days, etc., during the spring preceding the start of the next academic year. Consolidated survey recruitment synchronizes the multiple surveys to this date.

Working backward from the recruitment date, a coordinated approach for multiple surveys can incorporate two added features to manage overlap and/or leverage data across surveys. The process list, stated broadly, is from target date back to earliest step:

- Recruit Schools
- Negotiate with District
- Negotiate with State
- Draw Samples (all surveys), *coordinated to manage overlapping samples*
- *Fix the Version for Sampling Frame (CCD)*

- Complete Federal Approval Process (OMB)
- *Set burden limits*
- Finalize NCES Survey Preparation (all surveys)

The additions enable coordination of the multiple surveys' sampling plans. First is the opportunity to set burden limits using metrics that encompass multiple surveys as part of the sampling design. At the same time, opportunities to leverage information between studies may reduce overall burden. For example, the occurrence of paired surveys in some schools would allow a single collection of duplicated information, particularly when one survey will collect more refined school/teacher/principal demographic or economic information. Second is the use of a fixed sampling frame; third is redesign of the sampling plans for the multi-survey context.

Recommendation: Set up a cycle for survey planning through sampling and recruitment that will work for schools, districts and states as well as NCES and contractors.

Essentials – Stratification

Stratification of the population of schools defines (relatively homogenous) subgroups for the sampling algorithm to apply the appropriate probabilities of selection. At a basic level, all NCES surveys and assessments use a few common stratification factors (location-region; school type) for primary samples, possibly excepting some international studies. Further stratum refinements, oversampling, and some reserve samples, are individualized to meet specific survey objectives.

Greater concordance is needed for a multi-survey context; total uniformity is unlikely and might be quite inefficient. The degree of comparability of stratification and auxiliary variable definitions has implications for the advantages from a coordinated approach for multiple surveys. The greater the alignment of primary (at least) and other relevant stratification factors the more efficiency may be incorporated into the overall design. Similarly, the greater the alignment of important auxiliary variables, the greater the information for data users and researchers.

The first principle is to strive for a comparable stratification structure as deep as possible across surveys. This is true, whatever the survey designs. The second principle is for common, precise stratum definitions. This means identical variable definitions and identical boundaries. Wherever a survey requires additional fine-scale stratification, composites of those finer scale strata must be identical to coarser, in-common, strata.

To execute the individual survey designs, the finest granularity needs to be attached to the expanded CCD or the restricted appendix as part of the fixed version the sampling frame. Thus, stratum membership of each of the surveys is indexed for all schools in the frame, whether explicitly (possibly simplest) or indirectly (computable from factors as listed). Individual surveys may utilize aggregated fine strata. For example, a survey of fourth graders, might require a final stratification of school eligibility based on having/not having a 4th grade. For the survey of fourth graders, fine strata "not having a 4th grade" are superfluous and would be assigned a sample size of zero, or equivalently probability zero of selecting a school. For other surveys, the fine strata defined by "having/not having a 4th grade" would be re-paired into amalgamated strata for sampling purposes.

A further benefit to adding the stratification information to the CCD or the restricted appendix is the opportunity to leverage data across surveys both administrative data and participation histories. Comparability of definition will be essential for a limited number of auxiliary variables needed to create small area estimates relevant to individual districts (or even schools).

The “devil in the stratification details” lies in the need for identical variable definitions and boundaries. It is not particularly useful to simply note that definitions differ from one survey to the next since this offers the data user no meaningful way to actually view the information across studies. Potential solutions range from working toward standard definitions for common use (“top down”) to collecting raw data and using computational modules to categorize into strata for each of the surveys (“bottom up”). Midway between is creation of refined categories by absorbing all the boundaries into a single comprehensive list for the common variables. Alternatively, a minimal solution is to anchor every survey to a single benchmark survey (e.g., NAEP) by taking the intersection of each survey’s stratification with the benchmark stratification.

Recommendations:

- Standardize stratum definitions (including cut-offs as well as variables) for the primary variables.
- Standardize a core of key content across surveys and, when possible, anchor these to NAEP or other external source.
- Include in this core content the data needed to form small area estimates for the variety of specifications that meet schools’ and NCES’s needs.

Design Options

Three different approaches to design for a consolidating multiple surveys diverge in different ways from the current complete survey-to-survey design independence. The most straightforward of these is separate sampling followed by adjustment for all repeated selections of particular schools. At the opposite end of the complexity spectrum is combined sampling based on the total combined sample requirements, with subsequent allocation to individual surveys. An alternative coordinated sampling approach is based on partitioning the population and assigning segments to individual surveys for drawing survey samples. Composite strategies may be possible in practice to retain the most important advantages of different approaches, yet still practical for implementation.

Separate Sequential Sampling: Samples are drawn independently for the collection of surveys according to stratifications and sampling algorithms constructed for each survey. Common structure among surveys is minimal, likely including primary stratification variables but with independently set stratum boundaries beyond a core set. Overlaps due to repeated selection of the same school would be dealt with either with discretion on a case-by-case basis or by setting a priority list for surveys and proceeding sequentially to remove replicated selections, replacing them from a reserve sample or by individual substitution.

Separate sequential sampling presents the advantages that little is changed in the development of a sampling plan for each survey construction and stratum definitions can be optimal for every survey. Consequently, separate sampling is open to sample design innovations for individual surveys. Also, a common sampling frame can be used, with a fixed version and collective updating to allow comparable information for all surveys for data sharing, for post-stratification adjustment and analysis.

Procedural issues for separate sequential sampling are the necessity of priority determination among surveys and the development of a process for assigning schools selected repeatedly to one survey. Technical issues include preserving statistical properties when selection probabilities are altered by “removal” of a unit due to prior assignment another survey, and ensuring adequate precision of estimates and valid bias assessment and adjustment. Updating the sampling frame would be limited to the extent of commonality of factor and variable definitions. However, most importantly – based on small (two-survey) experiences – this approach will founder on a larger scale.

Total Combined Sample: The concept of a total combined sample is to take the sample size requirements for all the surveys to define a single aggregate sampling plan. Superimposing the stratification requirements for all the surveys defines a complex common structure for drawing the total combined sample. Then sampled schools can be assigned (in some random fashion) to individual surveys, satisfying the sampling plans of each survey individually.

There are multiple possibilities for the allocation of the total sample to the individual surveys. In general, the greater the agreement on stratum definitions, the wider the choices and the less complicated the allocation process. With broad agreement about strata, one possibility is to use random allocation within stratum to allocate schools to surveys. An alternative is to draw the total combined sample (or partition it after it is drawn) into large collections of mini-replicates within strata to be allocated (at random) to meet the sample size requirements of each survey. For highly refined strata, mini-replicates might be constituted across strata.

The chief advantage to a total combined sample is that coverage can be guaranteed for each individual survey while repeated selection of the same school is avoided. Obviously a common sampling frame is used, data sharing is enabled and updating for post-stratification adjustment is also uniformly applicable across surveys. A secondary advantage is that the (total) reserve sample can be smaller than the total size for independently drawn reserves.

Stratum comparability at a granular level may pose difficulties or at least require greater complexity in the sampling algorithm. In any case, the sampling and assignment processes will be computationally intensive – although not necessarily a limiting factor given the capacity of cloud computing. With a single combined sample, ingenuity will be required to introduce sample design innovations (e.g., selection probabilities adjusted by burden metrics or based on propensity scores) on less than a comprehensive scale.

Total Combined Sample – with Blocking: A variant of the total combined sample uses design of experiments principles to increase estimation efficiency and to address the need of multiple surveys for a small set of schools that share “rare” attributes. The idea is to ensure that every survey includes schools with each rare attribute and also schools with most (but not all) rare attribute pairs. Incomplete block designs are constructed with balanced (mini) replicates that together enable the estimation of the impact of each attribute to be estimated. For example, one mini-replicate schools with attribute pairs {A&B, B&C, C&D} while mini-replicates would include schools with {A&B, B&D, C&D}, {A&C, A&D, B&C}, {A&C, A&D, B&D}, {A&B, A&D, B&D}, and/or {A&B, B&C, C&D}. So while no single survey sample would necessarily contain schools with all possible rare attribute configurations, all surveys would contain schools with each of the rare attributes in most of the pairings.

The (incomplete) blocking depends on having a fine stratification structure incorporating the “rare” attributes at least for the sampling these schools. A two-part sampling process would be possible, using incomplete block mini-replicates to address the specialized sample requirements then using a simpler sampling process for the majority of the total combined sample and assignment of schools to surveys. The concept of incomplete blocking could be applied to the total sample and/or it could be applied for oversampling special populations. Its greatest use would appear to be for reducing the repeated sampling of schools with unusual circumstances or rare attribute configurations.

Partitioned Population Samples: The starting point for a quite different approach to coordination of samples divides the population rather than a combined sample. A collection of mutually exclusive subpopulations could be constructed using only limited common structure with random assignment to a particular subpopulation (survey) within each high-level stratum. Then each subpopulation would be assembled from the stratum partitions. Or a finer structure could be used to ensure a greater degree of representativeness for each subpopulation. Samples for each survey are drawn from the assigned subpopulation.

Details of this process could be adapted to meet additional sample requirements. For example, extra partitions could create a common reserve sample. To accommodate different sample sizes or sampling rates for different surveys, multiple equivalent mini-replicates could be created within each stratum to allow the number assigned to each survey to meet specific survey requirements. If (perhaps partial) overlap is acceptable or even desirable between specific surveys (e.g., an administrative survey and a student assessment), the same population partitions could be deliberately assigned to those surveys on either a selective (by stratum) or complete basis.

Flexibility gives this approach potential advantages in implementation. By first partitioning the population, then drawing samples, the use of mutually exclusive, representative subpopulations allows surveys to be designed independently according to their specific objectives. Innovative sampling designs and algorithms can be utilized by individual surveys. Stratum definitions can be optimized for each survey; sampling rates, including oversampling of demographic groups, are determined to meet survey requirements. The fixed, common sampling frame is used and then updated by all surveys. From the point of view of recruiting at all levels (states/districts/schools/students), nothing new has been added to the process except the important elimination of multiple survey requests to individual schools.

Connecting the Dots

Concern about potential for insufficient coverage of schools with “rare” factors or factor configurations comes from two sources. The first is the reduced population that is available to any single survey. The second, that random assignment of schools to partitions might result in unequal assignment of “rare” schools to subpopulations. (This shortcoming could be addressed as part of the partitioning process.) Other issues to consider with partitioned population samples derive from the independence that gives this approach its flexibility. These include limitations on sharing data across surveys and on updating the sampling frame. This concern could be mitigated if stratification information is specific enough and complete enough to permit categorization of each school using stratum definitions for the other surveys or at least using a fairly comprehensive common definition as a benchmark.

Table: Design Features for Three Design Options to Coordinate Sampling across Surveys

| | Separate Sequential | Total Combined (w/& w/o blocking) | Partitioned Population |
|---------------------------------|--|---|--------------------------------------|
| Sample Specifications | | | |
| Common Structure | Minimal | Comprehensive | Limited to Complete (better) |
| Population | Whole Population | Whole Population | Subpopulations (mutually exclusive) |
| Survey Integration/Independence | Independently Drawn Samples | Combined Sample | Independently Drawn Samples |
| Individuation | | | |
| Stratification | Optimal | Comprehensive | Optimal |
| Sample Size & Sampling Rate | Optimal | Optimal | Optimal |
| Sampling Methodology/Innovation | Individual Choice/Option to Innovate | Uniform for Total Sample | Individual Choice/Option to Innovate |
| Technical Issues | Standard | Computationally Intensive | Computational Solution |
| Special Target Population | Standard | Automatically Incorporated | Standard |
| Reserve Sample | Multiple Reserve Samples | Single Reserve Sample (option) | Single Reserve Sample (option) |
| Practical Issues | Prioritization of Surveys | Coordination of Survey Requirements/Methodology | None |
| Sample Composition | | | |
| Coverage/Representativeness | OK for original sample | ✓ | ✓ |
| Overlap/Duplicated Selections | Possible; most likely for schools with “rare” features | None | None unless deliberate |
| Scale Up | NO | Automatic (computational) | Automatic |

| | Separate Sequential | Total Combined (w/& w/o blocking) | Partitioned Population |
|--------------------------------------|--|-----------------------------------|--|
| Sampling Frame | | | |
| Fixed Sampling Frame (Version Saved) | ✓ | ✓ | ✓ |
| Updating | Requires Extra Information | ✓ | May Require Extra Information |
| Data Leveraging/Data Sharing | OK (with Complete information on record) | ✓ | OK (with Complete information on record) |

Design Considerations

Roles of NAEP. Since NAEP is managed and administered independently, it is excluded from the coordinated sampling plan for other NCES studies. However, NAEP becomes important in the alternating years when it is conducted because of its large contribution to burden, its large sample size, and its priority over other studies. This impact suggests considering a two-year coordinated sampling plan, if feasible; or at least a burden metric based on two years – a NAEP year and a non-NAEP year.

For the same reasons, NAEP has a potential role as an anchor or benchmark for other NCES studies.

Coordinated Covariates and Auxiliary Variables. Success in assessing and adjusting for non-response bias in a survey or study is limited by the extent of relevant information that is missing for non-respondents. Consequently, unless covariates or auxiliary information sufficiently distinguishes non-respondents from respondents, adjustments cannot be precise and impact can only be measured in broad terms. Improvement requires an external source of information about the non-respondents or at least about others that are close enough to exchangeable to act as surrogates.

NAEP data on respondent profiles and demographics could serve as such an external information resource for other NCES studies if basic covariates/auxiliary information were consistent. This would require either harmonizing stratum definitions and auxiliary information or else incorporating NAEP definitions into data records for other NCES studies. Of course, anchoring descriptive data in this way also would link together other NCES studies to enable data leveraging.

The most important variables to coordinate are those used in defining strata, calculating weights, estimating and adjusting for non-response bias, post-stratification adjustment. In addition the variables used in modeling and in small area estimation are essential to being able to provide relevant information at a more local level (“schools like ours”).

Common Features and Compromises. The starting point for coordinating surveys is a series of compromises, then gradually expanding the commonalities. In-common stratification information needs to be comprehensive – not just the intersection of variables used by all NCES surveys. To be useful, the data records for all surveys need to use uniform variable definitions and uniform boundaries even if the study design itself is based on variants and that information recorded as well. Of course, the more extensive the actual uniformity, the better. Note that footnoting to explain differences serves to warn the data user than sharing information is not possible, but does not enable that sharing.

Connecting the Dots

The same principle applies generally to enable leveraging survey estimates. A common core of variables across all surveys and/or common variable sets across directly related surveys, once constituted can be expanded over time. Ideally, commonality implies identical definitions, even identical wording of items. While this may be reasonable for adult respondents, practicality of instruments for student responses may need to be adjusted for age or learning level.

The value is great for data users and researchers of a common set of definitions comes in being able to augment information and expand research questions by combining (at least at a small summary stage) data from multiple studies. For individual studies, additional information can be used in the sample design and/or used to improve estimates and increase precision.

Design Strategies. Striving to improve data quality is a perennial task; response rate has long been taken as a primary indicator. Refocusing on representativeness as the objective sets a more direct objective, still minimizing bias and maximizing precision. An objective that opens sampling plans to design innovation. Immediate challenges are to increase efficiency by drawing on more *a priori* information, probability on non-probability-based and to base selection on measures of burden, propensity for non-response, and other properties or predictions. New challenges will surely arise in developing multiple survey coordinated or combined sampling plans. Solutions will likely be technical and technological. Non-response won't go away; representativeness and uncertainty are the natural terms for expressing the impact, the estimation and adjustment for bias.

Further new possibilities will arise for future thought when considering the collection of studies for one year. For example, these could be separated into administrative-only and student-participant studies, then either partitioned or paired deliberately. Schedules could be harmonized to be advantageous for those schools participating in multiple surveys.

Recommendations:

- Unify the process for drawing the samples for all surveys & assessments for each academic year (possibly expanding to two years – one year with NAEP and one interim year).
- Manage overlap to leverage across surveys, with overlap with NAEP constituting a special case.
- Redirect focus on response rates to quality of information: representativeness, precision and bias of estimates.
- Support the technical development of one or more approaches to coordinated sampling, with theoretical, simulated or data-based reconstructions to evaluate their relative efficacies and efficiencies.
 - Redirect focus on response rates to quality of information: representativeness, precision and bias of estimates.
 - Manage overlap to leverage across surveys, with overlap with NAEP constituting a special case.

PART II: IMPLEMENTATION AND IMPACT

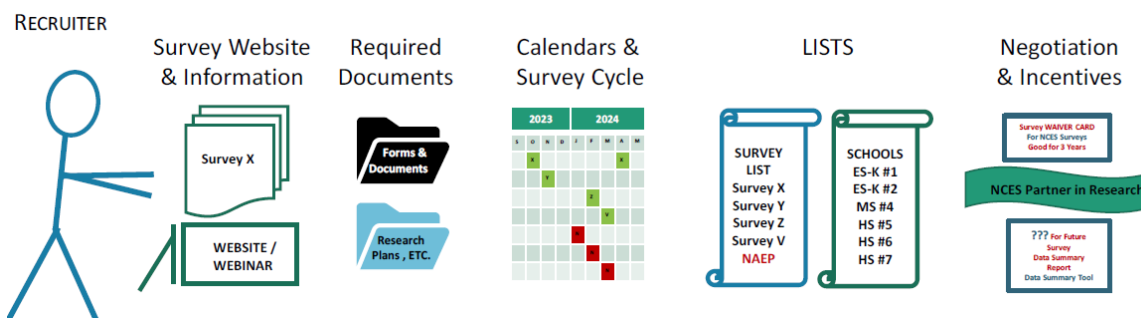
The starting point for Panel II is at the beginning of the recruitment process. It is assumed that technical/statistical problems of a combined recruitment design and structure have been solved, samples have been drawn, instruments have been approved (IRB and OMB) and all the essentials for initiating actual recruitment are in the hands of a single recruiter for any one state/district. Attention now turns to the recruitment process itself: first to identifying the key considerations for making it efficient and effective, i.e., successful, then to identifying and addressing the implementation issues.

For NCES, improving participation will improve the national data base. At the upper administrative levels (state/district), streamlining the process is a primary consideration, while at the school level the consideration is rather controlling the burden and clearly defining benefits that provide a balance. The foundation for success is relationships of trust and dialogues.

Recruitment Paradigm

Role of Recruiter

The recruiter brings three kinds of materials to the review-negotiation-approval transaction: informational, negotiable, and persuasive. Increasingly what is considered to be negotiable is expanding.



Informational materials include study-specific information such as survey descriptions, web-based explanations and examples, classroom time requirements for participants and effort requirements for school staff, FAQs tailored to administrators/educators at various levels and potential participants, etc., plus the actual survey instruments. Individual states and/or districts may also specify requirements for a research plan with the projected benefits to the state, district, locality and/or school.

Participation by the school and scheduling are open to negotiation, as always is the case. Necessary materials include listings of schools by study or reserve sample (whether individual or combined). What is different in the multi-study paradigm is that the calendar is a composite and consequently can enable comprehensive plans for the approaching academic year and can set up the coordination of studies when more than one will take place at the same school, school system or within the same locality.

Agreements must also be reached with respect to requirements to ensure privacy of personal information and security of data. These include documents from high-level standards to highly specific consent forms. It becomes troublesome when federal and local specifications are not consonant. In this case, negotiations can be difficult and may leave the contractor fielding the survey in an untenable, and

Connecting the Dots

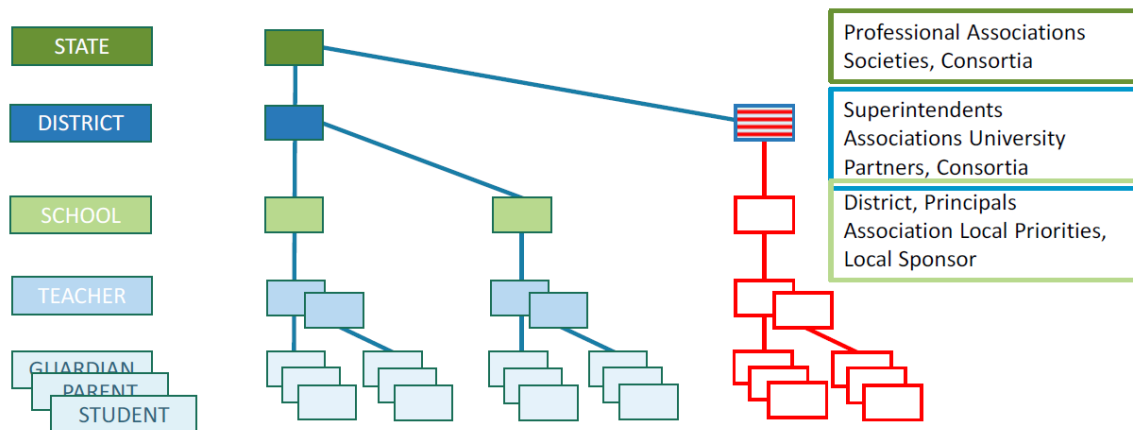
litigable, position. Having a single recruiter offers the opportunity to limit these more difficult negotiations to a single occasion and a single resolution whenever the standards and requirements do not differ materially across multiple studies.

Persuasive materials include incentives, recognition, rewards at the student, school and district levels plus any specific benefits presented by the study itself.

The recruiter's goal is to reach an agreement that brings the greatest participation achievable, to work toward the position of trusted voice, and to obtain the support of administrator as an advocate in proceeding onward to the next level of recruitment.

Gatekeepers and Advocates

State Offices of Education, at the highest level of the chain of gatekeepers, often review basic information but pass the serious, detailed review of study proposals down to the level of school district administration. The district's role in review and approval is critical because declination by the district eliminates all schools in that district from the study.



At the subdistrict (if the district is large enough to be partitioned) and at the school levels, recommendations from district and state officials are important and peer group organizations may be influential by their assumed knowledge of issues farther down the hierarchy.

Advocates at each level in the chain of gatekeepers include relevant organizations and respected individuals whose voices are credible and trusted, and whose endorsement will carry weight in promoting participation. Support for the study from professional education organizations is respected and valued by state and district gatekeepers; but not as relevant locally.

At the subdistrict (if the district is large enough to be partitioned) and at the school levels, recommendations from district and state officials are important and peer group organizations may be influential by their assumed knowledge of issues farther down the hierarchy.

In some cases, localities are now specifically requesting such support from disinterested local voices, i.e., advocates for the local community who have no connection to the school itself, to education administrative hierarchy or to the source (in this case NCES/federal government) running the study. This is a second instance where the NAEP coordinator with knowledge of the local communities may be of assistance by identifying local leaders who can be approached for endorsement.

Recommendation:

- To smooth the process, it is desirable for the NAEP coordinator, as a trusted voice, to initially introduce the single NCES recruiter to the appropriate administrators at the state and district levels.

Advance Messaging

The real beginning of recruitment comes much earlier than shown in the paradigm in Figure 1. Recruitment starts with preparatory messaging to all the involved communities and the outreach from NCES to engage advocates and create partnerships with societies, professional organizations and other potential influencers at all levels. Timing is important – the earlier the better, and definitely well in advance of the start of actual recruitment.

Advance Messaging – Professional Education Community

Advance messaging begins by painting the big picture for the upcoming and one, two, or more future years. This calendar of firm plans and schedule for the upcoming and projections, both firm and tentative, for the future years, can meet the needs of education administrators to plan for NCES proposed studies. For this purpose it is important that the calendar be comprehensive and include NAEP and international studies. Consideration should be given to inclusion of planned IES studies if these appear under the same brand.

Messaging before studies are launched offers the chance to highlight for each proposed study the impact of antecedent studies. It enables administrators to anticipate the proposed study's demands on classroom time and school resources and to project the importance of the data to be collected.

Advocates and Trusted Voices

NCES has well-established relationships with professional societies and associations as well as state education offices, large school districts and consortia. Advance messaging positions them to take on a partnership role and to actively advocate with gatekeepers.

At each level, the gatekeeper can become a trusted voice and advocate for the next level in the chain – the state education office for the district; the district superintendent for the school. Another trusted voice is the NAEP coordinator. With introduction and support from the NAEP coordinator, the NCES recruiter could be added to this list.

On a local scale, in addition to recommendations from the district and state education officials, there are professional associations and peer groups (e.g., principals' associations, state level organizations). A local sponsor, now required by some locales, may be highly influential. Often trust has already been established with the NAEP coordinator, enabling sharing experience and local connections with the NCES recruiter.

Recommendations:

- Use Advance Messaging extensively and sufficiently far in advance to prepare for academic year surveys.

Messages

To be effective, recruitment messages need to be delivered directly to each audience. On a national scale, professional meetings and NCES-hosted events continue to be successful platforms to reach state education offices and district administrators.

These avenues don't serve as direct communication with most teachers, principals, and superintendents of small school districts who do not regularly or even irregularly attend national or major professional meetings. The virtual world for education, imposed by Covid, opens an opportunity for broader virtually direct communication through forums and other opportunities for discussion and feedback.

"It's on the website" is an ineffectual argument, except as a reference. Like any website, it is useful only for someone who either already knows what to find or strongly expects to find it. The IES/NCES website is particularly difficult to navigate – definitely not a direct communication.

Tailored Recruitment Messages

Both content and extent of detail of messages prior to actual review of full recruitment packages vary for their differing audiences. In general, concise, focused messages with links to extensive information are more likely to be read. What each audience needs is a clear picture of what participation looks like from their particular vantage point: what it is they are asked to do; what specific benefits are returned and what the consequences are for not participating.

Comprehensive messages are needed by state and district offices, covering the complete set of studies. A succinct overview could address the plan, the study value, and study requirements, with links to more information. A two-year calendar (even if the second-year plan is tentative) would give administrators a basis for thinking in terms of a pair of NAEP and non-NAEP years.

For each proposed study, information should include highlights of its impact in previous years with a link to the full report online. Identifying on the one hand the usefulness and relevance to specific decisions (past and/or future), and presenting on the other hand an overview of the burden for the proposed study can address the balance of value and burden that comes with approval and participation. Declining approval means the opportunity loss: What happens to national representativeness if the state or district declines?

At the local level (small district or school), the focus is on impact of the study for the school. A page document can explain the value and impact of the study, especially (past and proposed) study usefulness and relevance to specific decisions. The study-specific requirements are critical to local implementation the nuts and bolts that affect the school, the students, the classroom time and schedule. Clear statements of the value to the school and of the individual importance of this school to the proposed study are at the core of this summary. Again, what happens to national representativeness if the school declines?

Branding

The need for consistent branding is implicit in undertaking a coordinated approach to recruitment. Consistent, pervasive, visible, recognized branding is the link that connects the NCES studies from recruitment through completion.

Simply stated, the universally recognized brand is the Department of Education with its logo. At state and large district offices, IES/NCES are recognized. But at more local levels IES/NCES may simply be part of the unending collection of acronyms and initials, with logos – beautiful or not – that are not distinguishable from the rest. Using the Department logo as the primary, with the IES/NCES as secondary (and contractor logos as ancillary) will take advantage of current brand/logo recognition while clearly linking NCES studies.

Recommendations:

- Tailor recruitment materials to recipients; stress the impact – past and anticipated for future. Omnibus information for districts; relevant survey information for school.
- Brand everywhere consistently with US Department of Education (primary); NCES/IES (secondary).
- Improve the Department of Education website, specifically the IES/NCES website, to be more inviting, user-friendly and easily navigable.

Foundation for Relationship

Trust

Establishing a successful continuing relationship depends on building trust. The success of NAEP coordinators provides a model that may be instructive if not (budgetarily) feasible to duplicate. Once the coordinator – and, hopefully, now the recruiter – becomes versed in the state, district and/or local context, communications can be deeper, more detailed and more candid.

Both the state or district and the recruiter benefit from a true dialogue in the recruiting process. Streamlining the process is an obvious benefit. Even more important is the dialogue about perception of benefits, identification of obstacles and characterization of burden. Burden associated with study participation is then put into a larger context that also includes burden coming from other sources. Having a common context for discussion eliminates the need to revisit issues or to resolve inconsistencies when a mutual understanding or a compromise has already been reached.

Districts differ widely in size and in resources; so a valuable opportunity to understand what comprise the potential benefits and the obstacles for a given district and for the schools within that district comes with dialogue. Large districts and consortia with research units can regularly analyze data from multiple sources to report on education practices and trends or can collaborate with universities for in-depth examination of contemporary issues. Small districts may have neither capable staff nor resources to use NCES data to compare themselves to “other schools like ours.” With trust in the relationship, an NCES recruiter may gain a clearer perception of the benefits, obstacles and burdens of the individual districts. And as a result, a partnership in the recruiting process can benefit district, school and NCES.

District policies for recruiting schools also vary. However, a partnership facilitates the process when the superintendent vouches for the recruiter who may in turn introduce the contractor for the study to the schools’ principals.

Dialogue and Partnership

Dialogue, especially at the district level, creates a space for input on the process and on the information priorities. The specific district context, with demands arising out of its local issues and initiatives and its

potential barriers, can give insight into the decision to endorse or not to approve one or more schools' engagement in a study. This understanding can also open a discussion of what constitutes burden and how might it be measured or tracked.

On a practical level, this dialogue may lead to adaptations of the recruiting plan in order to accommodate specific district policies or practices. As another example, superintendents could welcome prepared materials to distribute from their offices to schools in support of participation in a study. Such materials would address the value of the study (past and future) and the specific requirements of time and resources; they could also lighten the load for districts and schools by including template emails and letters.

Time for local dialogues is likely to be even more limited, and either an open (virtual) forum or debriefing session with participating principals and (separately) teachers may be more fruitful. The need for dialogue is just as acute as at the district level. The focus is on specific school needs, recognizing that these will differ widely across schools, although the desire for information on impact and its relevance to individual schools is universal.

Opportunities to meet school needs at some schools would include tools such as templates for tailored summaries and tutorials to make NCES data accessible and relevant. On the other hand, gathering administrative workload information is necessary input for understanding the resource costs of study participation.

Open discussion could lead to recommendations for determining and/or measuring burden. It also could identify individual school priorities for potential direct benefits as value in return for participation.

Recommendations:

- Build trust and open dialogues with gatekeepers at all levels.
- Build trust through a continuing relationship with a single (state/district) NCES recruiter.

Process Priorities

Streamlined Process

The big advance in streamlining the review and recruitment process is reducing from many to one the recruiter relationship with the district. This places all these transactions in the hands of a single recruiter with knowledge of the district context and the decision-makers on its staff. From this point, further tasks are also reduced from many to one, or a few. Standard sections, especially descriptive material about NCES and its policies, need only to be examined once. Specific legal agreements involving security, privacy and rights and ownership of information can be negotiated once for the collection of studies.

Pre-study preparations, especially the gathering of background and administrative information can be done once for delivery to multiple studies, if the school is involved in more than one. Other pre-study preparations such as sampling of individual students, obtaining parental permission can be batched into a single distribution of requests for all students in all studies during the academic year. With both careful planning and serendipity, multiple studies at a single school can be coordinated into the same few days on the school calendar.

Connecting the Dots

Shortening, delegating or sharing sources further streamlines the study data collection process at the school. Using the single “best source” for core data, for example, streamlines by requesting once from administrative records (or district or consortium records) the necessary data for inclusion in all studies, without duplicating this request for school information from students. Even for this task, administrative effort can be minimized by pre-filling information from current/or concurrent studies and surveys or even from past surveys, from CCD data or from the school’s own website.

Elimination of redundant items across surveys, primarily administrative information. But some teacher and principal information is now gathered in cross-sectional student studies or even longitudinal studies that is also collected but in greater detail, for example, in the Teacher and Principal Survey. Where these studies are both conducted in the same school, the less detailed information from the student studies can be omitted and replaced. Similar logic applies to asking for parents’ occupations – clearly the parents’ information suffices (and student information in this case is notoriously unreliable).

Finally, standardizing streamlines because it allows one-time responses. For multiple NCES studies this requires a level of coordination because both the survey question and the allowable responses must both be standard across surveys for the information to be sharable. A problem arises when the particular variable has categorical responses, whether qualitative (e.g., ethnicity) or numeric (e.g., age spans that define “teenage”). This problem is important but it is solvable.

The solution of working to set standard definitions for all such variables appears unattainable. However, for a given time period (e.g., a pair of NAEP and non-NAEP years) all studies could add to their own categorizations the NAEP definitions. That would allow sharing of information at least for the common set of NAEP response categories. Alternatively, a set of “refined” categories that include all partitions used by any one or more of the two years of studies could be implemented. This would allow complete sharing but with simple to create computational modules, each study could still follow its own category definitions, some of which have been in use for a series of past studies. Finally, the whole problem could be eliminated at least for numeric data by collecting the raw data and allowing each study to re-assemble categories according to its own needs.

Calendar

The purpose of a comprehensive calendar is to meet school planning schedules for the academic year, regardless of the dates that individual studies are to be conducted. For schools invited to participate in multiple studies, this may also offer an opportunity to amalgamate at least the administrative information during a single data collection.

Thus, the schedule is created for all studies by working backward through the process to arrive at the point of recruiting individual schools:

9. Recruit Schools
8. Negotiate with District
7. Submit for State Approval & Support
6. Draw Samples (all surveys)
5. Set Version for Sampling Frame (CCD)
- 4. Initiate Advance Messaging**

3. Complete Federal Approval Process (OMB & IRB)
2. Set Burden Limits
1. Finalize NCES Survey Preparation (all surveys)

Advance planning and cooperation with OMB in particular will facilitate their handling of multiple studies arriving for review in a short period of time.

Recommendations:

- Standardize to eliminate redundancy and improve information return (“schools like ours”).
- Streamline the recruitment process to reduce the process time required by states/districts/schools.
- Set up an annual cycle for survey planning: sampling through survey administration that works for schools/districts/states, NCES, contractors.
- Unify the process for all surveys & assessments for each academic year (expanding, or at least projecting, to two years – one year with NAEP and one interim year).

Balance of Burden, Benefit, Capacity

The greatest disparity among schools and also among districts seems to lie in what truly constitutes burden, what constitutes a benefit and how to strike a balance. Size and resources are two contributing factors to this disparity; pressures from local issues, time already promised to other studies, introduction of initiatives, and recent events affecting an individual school are other factors. The vulnerability to loss of learning time for the students and the makeup and stability of the student population stability can be equally critical.

Burden

With respect to the study itself, the burden varies depending on whether data collections recur (longitudinal v cross-sectional), who is involved (administrative data v principals or teachers v students and/or parents) and whether it includes an assessment.

With respect to the school or the district, the time-honored measure of Minutes x Number of Students/respondents is the tip of the iceberg. Differences among schools are greater for other aspects of burden. The administrative preparation from data requests, permission letters, sampling, to scheduling for individual students in the sample, can be part of daily work flow in a well-resources school. For a school with small staff, the same tasks can add to an existing overload. Obviously, studies that multiply the number of kinds of respondents multiply the differential among schools or districts.

Another dimension of differences among schools is in the vulnerability of their students to lost learning time. Poverty, language learning, disabilities, and students facing external challenges inflate the cost of missed classroom time. The effect of removing some students from a classroom goes further because it also affects the teaching and activities that go on while those students are out. Loss due to disruption of class schedule and curriculum plans and progress occurs with each day of one or another study is underway in the school. These are the kinds of components to burden that need input from schools of varied size with differing population compositions that present a range of challenges to learning progress.

Connecting the Dots

However burden is defined, it accumulates over years. Understanding how big a contribution to NCES data has already been made for each selected school is needed information for improving sampling designs and also for information to the district office when the study is under review for approval.

Value Returned

For a participating school, the benefits from participation can be of many kinds, for example, they can come in the form of information, input to future studies, direct relief to burden or recognition. It is generally agreed that cash incentives need to be larger than is possible for NCES to underwrite in order to be effective. Once again, the needs of the individual district or school vary greatly, so that provide one or two new computers would be valued by a small school with limited funds for administration but have minimal impact for a research department of a large district.

Information can take the form of documentation of study or survey impact: What does the country get from this national survey that it could not get otherwise? This could include past results and their impact, with more specific information to allow comparison with (current) local information. This question also prompts explanation for prioritizing availability of current information.

Information can also take the form of school-specific data and summaries for “schools like ours” that are more relevant to local decision-making than the national estimates. For schools or districts without their own research resources, a school-tailored data tool would extend access to data from the current data collection and, to the extent feasible, to data summaries from other NCES studies as well. Another potential benefit to offer to this cohort of schools and districts, would be an onsite tutorial on how to use NCES data tools to adapt for issues of current local concern. The tutorial could be delivered during in conjunction with the data collection process.

Taking the opportunity for input, perhaps at the close of the data collection or during tool demonstration/tutorial, benefits both the teacher, principal or superintendent and NCES and the contractor. Dialogue rather than formal debriefing can reveal topics of concern that might be covered in future studies (especially timely pulse or other brief surveys) or identify particularly valued elements of burden and/or benefit.

Direct relief of future burden could be a waiver from one or more years of study eligibility. For studies that require substantial administrative effort or data collection(s) with multiple days onsite, an administrative supplement could cover reimbursement of costs for additional personnel or resources.

In the past recognition of participating schools has been hampered by privacy guidelines. Across a broader time-frame (e.g., five years) and encompassing other forms of contribution or participation in addition to surveys and assessments, recognition might be possible. A banner, “Partner School in US Department of Education Research,” and publication of the list of partner schools would be inexpensive acknowledgment of the value of the partnership to NCES and would recognize the joint responsibility for collecting and reporting data on US schools.

Recommendations:

- Differentiate among small/medium/large districts/states.
- Work to define Burden in district/school/respondent terms. Account for administrative tasks (pre-to-post), also days in school as well as {minutes x number of students}.

- Work to define Benefits in district/school/respondent terms.
- Differentiate among small/medium/large districts/states with respect to needs and to capacity for participation.

Considerations

History of Participation Data Base

A resource and research data base of schools' participation histories is sorely needed. This comprehensive data base of schools' participation histories across all studies will serve several important goals. On the practical side these records of participation will document the short term (within an academic year) and accumulated (two-year and five-year) burden and be the basis for quantitation. This information on accumulated burden will have immediate direct use in the sampling and recruiting processes. Participation – actually, non-participation – data is the direct input needed to improve non-response bias estimation and adjustment.

As research data, they create a foundation for intelligent design for coordinating multiple surveys and for introduction of advances in sampling methodology such as use of propensity scoring (relative to non-response). For the important problem of defining metrics for burden, especially to measure burden, as perceived by districts, by schools and by teachers and principals, these data are needed for input and for validation of metrics proposed. Since these data can document the approval/decision-making at each level other important research questions can be investigated. For example, refusal patterns can be studied, propensity to decline can be estimated, and effectiveness of burden-indexed compensating benefits can be examined.

Data Quality Indicators

NCES released data is widely acknowledged to be consistently of high quality. In setting thresholds for surveys, response rate has been a prime indicator. But for each study, the value of the data collected depends on more than the response rate and accuracy. For the users, including the participating districts, schools, educators and local public, the value depends on representativeness, timeliness and relevance.

Response rate by itself is inadequate, and increasingly poor as a sole proxy. For an intended goal of representativeness, the sampling design might be reconsidered that could reduce bias, improve precision or increase cost effectiveness. Emphasis is shifted to the impact of each non-responder rather than just the frequency of non-response. For example, knowing the likelihood of non-response could eliminate plans for multiple attempts and could also enable prediction of the contribution to bias. Finding supporting information from either probability or non-probability sources, could streamline sampling by allocating sample sizes to minimize predicted bias or maximize predicted precision (the familiar criteria for sample survey design).

Timeliness is a separate issue; data curation necessarily demands time and NCES is regarded as relatively efficient and definitely meticulous. Other federal agencies (e.g., the Federal Reserve Board, the National Agriculture Statistical Services) also under pressure for timely reporting publish prompt “preliminary reports” with a series of scheduled “updates.” NCES can weigh the pros and cons of such an approach, perhaps an initial, still carefully curated, representative subset to be used for “preliminary” findings with

the associated larger uncertainties. Or, for the future, rethinking the curation process to take greater advantage of technology in order to utilize more sophisticated real-time automated monitoring to allow early presentation of broad summary information, at least.

The desire for relevance is the easiest to understand and probably the hardest to address but maybe the most important of all. The question arises at all levels: *What will data from this study tell us about schools (districts) like ours?*

The statistical methodology to respond to this question exists and is in use at other federal agencies. Tools have been developed and could be tailored to work with NCES data. So the issue is not the feasibility. For technically able users such as the large districts and consortia with their own research offices and academic researchers just need the access to fine-grained data. Smaller districts and certainly individual schools are left out. There is an opportunity to create a general tool, applicable to data from multiple studies, that could be tailored for a participating school to use (perhaps offered as a benefit for participation?).

These small-area estimates are predictions with associated, sometimes large, uncertainties. How useful they are depends on what “like ours” means. Stratum means or margins for two- or three-way tables are not precise enough. Imposing the cross-section of all stratification variables simultaneously would be a good starting point, particularly if the more refined stratifications suggested elsewhere in this report are applied. Conversation with education administrators at all levels could be valuable in identifying other factors to include that govern their decision making.

Launching Change

What is discussed throughout this document is change that is ambitious and will affect – and bring advantages to – many communities: educators, education administrators and decision-makers, researchers, groups with interests in education policy and practice, and others. Therefore the launch, probably going forward in parts, needs to be carefully made. The risks of haste or insufficient buy-in are high.

The first step is the advance preparation of the education community broadly. This begins with all the influencers, advocates, trusted voices whose support can pave the way for making the changes welcome and successful.

In tandem is the preparation by NCES staff and contractors to facilitate the integration of multiple studies, identify the critical points and define the mechanisms needed for coordination, and to find appropriate paths toward standardization sufficient for the level of data sharing required.

Recommendations:

- Create a History of Participation data base for districts and schools, linked to CCD for basic information school information.
- Manage overlap to leverage across surveys, with overlap with NAEP constituting a special case.

PARTS I & II: PRINCIPAL FINDINGS AND RECOMMENDATIONS

At the conclusion of its deliberations, each panel independently presented its findings and made specific recommendations, noted in the text above. The complete findings and recommendations from *Part I* and *Part II* are combined below because the overlap was extensive and nowhere were these in conflict.

The Panels' Principal Findings

Coordination of NCES surveys and assessments is unequivocally desirable and feasible.

Accomplishing this will require extensive changes in several aspects and significant investment of effort and technical expertise to achieve integration across surveys moving toward standardization of critical survey components.

Building trust and open dialogues with gatekeepers at all levels is an immediate step and a continuing objective.

Initial steps can be taken before a comprehensive plan is complete.

Careful launch of these changes requires energetic communication and advance preparation of the education community, also coordination with OMB and current and potential survey contractors.

Specific Recommendations

NCES Supporting Infrastructure

- Expand the sampling frame, both open access CCD and a restricted access expansion to CCD.
 - Assign a permanent school ID for use in sampling and analysis of sampling, participation, and burden histories.
 - Create a History of Participation data base for districts and schools; include recruitment attempts, results, reasons for refusals or other non-response, and metrics for burden.
 - Moving forward, develop a research base for studying patterns of refusal and for estimating the impact of non-response, using external sources to validate.
- Unify the CCD updating process to be continuous, but with date-stamped versions each year, used by all surveys for sampling, in post-stratification adjustment and in data sharing/integration.
- Standardize a core of key content across surveys and, when possible, anchor these to NAEP or other external source.
 - Include in this core content the data needed to form small area estimates for the variety of specifications that meet schools' and NCES's needs.

Connecting the Dots

- Standardize stratum definitions (including cut-offs as well as variables) for the primary variables; at a minimum incorporate NAEP stratum definitions and stratum boundaries into records for all surveys to enable data sharing/integration across surveys and validation of survey estimates.

Calendar

- Set up a cycle for survey planning through sampling and recruitment that will work for schools, districts and states as well as NCES and contractors.
- Set up an annual combined calendar that includes NAEP and international studies for planning purposes and burden evaluation. Preferably expand to include a contemplated calendar for the following year or two years.

Trust, Dialogue and Partnership

- Build trust and open dialogues with gatekeepers at all levels.
- Build trust and understanding of local issues through a continuing relationship with a single (state/district) NCES recruiter.
- Differentiate among small/medium/large districts/states because of differing resources, needs, magnitudes of disruption by surveys, and frequencies of repeated sampling.
- Work to define Benefits in district/school/respondent terms, to identify a collection of potential benefits for different circumstances.
- Work to define Burden in district/school/respondent terms.
 - Account for administrative personnel time and tasks (pre- to post-survey), also days present in school as well as {minutes x number of students}.

Messaging

- Brand uniformly and visibly with US Department of Education (primary); NCES/IES (secondary) survey contractor (tertiary).
- Disseminate targeted advance messaging widely to set expectations for surveys scheduled for the next one or more academic years.
- Use trusted voices as advocates at every level; provide suitable materials for advance messaging and for recruitment discussions (e.g., District support to prepare local recruiting).
- Tailor recruitment materials to recipient; stress impact – past and anticipated for future. Omnibus information for districts; relevant survey information for school.

Streamlined Recruiting

- Use a coordinated sampling approach to reduce burden on individual schools, on school districts, on administrators, coordinators, recruiters and NCEES and contractors. The coordinated request would still allow selective approval/rejection of individual studies for implementation in each state/district/school.
- Streamline the recruitment process by reducing, if not entirely eliminating, duplicative explanations, justifications, requests and negotiations.
- Establish a single, long-term contact (recruiter) in order to improve the quality and efficiency of interactions while reducing the number of meetings between recruiter and education office or administrator and overall decreasing the process time at state/district/school levels.

Sampling Design

- Unify the process for drawing the samples for all surveys & assessments for each academic year (possibly expanding to two years – one year with NAEP and one interim year).
- Support the technical development of one or more approaches to coordinated sampling, with theoretical, simulated or data-based reconstructions to evaluate their relative efficacies and efficiencies, and their scalability.
- Manage overlap proactively to leverage across surveys, with overlap with NAEP constituting a special case.
- Use stratum definitions that are either standardized across surveys or retain comprehensive information on stratum variables, sufficient to conform to a common definition, when possible, anchored to NAEP or other external source.
- Redirect focus on response rates to quality of information: representativeness, precision and bias of estimates.
 - Redefine Data Quality – expand beyond accuracy and response rate, to include timeliness and usefulness for information integration.
- Moving forward, support research to identify patterns and calculate propensity for non-response (as input to designs for future sample) and to estimate the impact of non-response, using external sources (including NAEP) to validate non-response bias assessments and adjustments.

APPENDICES

- A. Agenda
- B. Charge to Panels
- C. Expert Panel Biosketches

Panel I

Jason M. Fields, *US Census Bureau*
Daniel Pratt, *RTI International*
Trivellore E. Raghunathan, *University of Michigan*
Jerome P. Reiter, *Duke University*
Keith Rust, *Westat*
Changbao Wu, *University of Waterloo*

Panel II

James Appleton, *Gwinnett County Public Schools*
Deborah Herget, *RTI International*
Rachel Horwitz, *US Census Bureau*
Regina Lewis, *Maine Department of Education*
Jerome P. Reiter, *Duke University*
Robert G. Sivinski, *Office of Management and Budget*
Kevon Tucker-Seeley, *Los Angeles Unified School District*
Gina K. Walejko, *Google*

NISS

Nell Sedransk
Brian Habing
Ya Mo

APPENDIX A: AGENDA – PANEL I



National Center for Education Statistics



**NISS-NCES EXPERT PANEL
DEVELOPING AN INTEGRATED APPROACH FOR MULTIPLE SURVEYS**

August 27 – September 27, 2021 on ZOOM

AGENDA

PANEL I – DEVELOPING POSSIBILITIES AND FRAMING STATISTICAL ISSUES

| Friday, August 27, 2021 – Sampling Structure for NCES Surveys | |
|--|--|
| 10:00 – 11:00 am | NCES Essentials & Constraints <i>Introductions, Nell Sedransk, Director, NISS-DC</i> <i>Charge to Panel, Peggy Carr, Commissioner, NCES</i> <i>Presentation, NCES</i> <i>Presentation, NCES</i> |
| 11:00 – 12:00 pm | Q & A |
| 12:00 – 1:00 pm | Lunch |
| 1:00 – 2:00 pm | External Requirements & Limitations External Perspectives/Issues, Panel Discussion, NCES & Panel |
| 2:00 – 3:00 pm | Discussion |
| 3:00 – 4:00 pm | Discussion & Planning, Panel Only |
| Monday, September 10, 2021 | |
| 1:30 – 3:30 pm | Panel Deliberations Session 1: Exploring the Possibilities, Panel Only |
| Tuesday, September 17, 2021 | |
| 1:30 – 3:30 pm | Panel Deliberations Session 2: Outlining a Strategy, Panel Only |
| Wednesday, September 27, 2021 | |
| 1:30 – 3:30 pm | Debriefing & Discussion, Panel & NCES Staff |

PANEL

Jason M. Fields, US Census Bureau
Daniel Pratt, RTI International
Trivellore E. Raghunathan, University of Michigan

Jerome P. Reiter, Duke University
Keith Rust, Westat
Changbao Wu, University of Waterloo

NISS

Nell Sedransk
Brian Habing
Ya Mo

APPENDIX A: AGENDA – PANEL II



National Center for Education Statistics



**NISS-NCES EXPERT PANEL
DEVELOPING AN INTEGRATED APPROACH FOR MULTIPLE SURVEYS**

October 25 – November 18, 2021 on ZOOM

AGENDA

PANEL II – MAKING IT HAPPEN – DEVELOPING IMPLEMENTATION TEMPLATES

| Monday, October 25, 2021 – Implementation of Integrated Sampling | |
|---|---|
| 1:00 – 3:00 pm | NCES Essentials & Constraints <i>Introductions</i> , Nell Sedransk, Director, NISS-DC <i>Charge to Panel</i> , Peggy Carr, Commissioner, NCES <i>Recruitment Processes & Challenges</i> , Gail Mulligan, NCES <i>Addressing Recruitment Challenges</i> , Andy Zuckerberg, NCES |
| 3:00 – 3:30 pm | Break |
| 3:30 – 5:30 pm | External Requirements & Limitations External Perspectives/Issues, Panel Discussion, NCES & Panel |
| Tuesday, October 26, 2021 | |
| 10:00 – 12:00 pm | Panel Deliberations Session 1: Exploring the Possibilities, Panel Only |
| Thursday, November 4, 2021 | |
| 12:00 – 2:00 pm | Panel Deliberations Session 2: Outlining the Essentials, Panel Only |
| Thursday, November 18, 2021 | |
| 2:00 – 4:00 pm | Debriefing & Discussion, Panel & NCES Staff |

PANEL

James Appleton, Gwinnett County Public Schools
Deborah Herget, RTI International
Rachel Horwitz, US Census Bureau
Regina Lewis, Maine Department of Education

Jerome P. Reiter, Duke University
Robert G. Sivinski, OMB\EOP
Kevon Tucker-Seeley, LA Unified School District
Gina K. Walejko, Google

NISS

Nell Sedransk
Brian Habing
Ya Mo

APPENDIX B: CHARGE TO PANELS

“Connecting the Dots, I & II”: Developing an Integrated Approach for Multiple Surveys

NCES is looking to move forward vigorously to implement some of the key recommendations from two 2019 and 2020 NISS-NCES Expert Panel Reports. The primary recommendation from the panel on Post-Covid Surveys was for IES/NCES to:

Implement a recognized functionally coherent and transparent structure to replace/reorganize loosely connected collection of separate surveys and assessments.

Panel I: Developing Possibilities and Framing Statistical Issues

As a first step to envision how an integrated or combined (multiple survey) recruitment process might work, a Technical Working Session with external experts and NCES staff focused on sampling design: the options, the obstacles, and ways to overcome them. The goal for this session was an unfettered discussion of possibilities, considering how to employ an integrated sampling process using current list frames, but going beyond modification/updating of current sampling designs and management. To illustrate how this process might work and/or what a consolidated sampling plan might look like, the scheduled 2023-2024 surveys were used as one working example for creation of basic outlines for one or more approaches.

One aspect is sample overlaps among surveys, with attention to statistical properties and advantages in specific situations either for deliberate use or for structured avoidance. A secondary benefit of this process could be sharing contextual data across surveys either to reduce burden or to improve precision, amplify covariates or increase information.

Panel II: Making it Happen – Developing Implementation Templates

With possibilities outlined by Panel I of “Connecting the Dots” a new panel of external experts held a Technical Working Session to address implementation issues for a consolidated multi-survey sampling plan, in particular, the recruitment process. Important perspectives include: 1) experience from both sides of negotiation – government/contractor and district/school, both large and small; 2) insight into constraints and/or regulations both government restrictions and participant protections; 3) statistical expertise to ensure that the statistical properties of the design are not compromised. The focus on implementation encompassed identifying and addressing issues from a survey methodology point of view as well. Follow-on sessions (panelists only) considered how to structure the process. To illustrate how this process might work and/or what a consolidated recruitment plan might look like, the scheduled 2023-2024 surveys were used as one working example for creation of basic outlines for one or more approaches.

APPENDIX C: EXPERT PANEL BIOSKETCHES

Panel I

Jason M. Fields, PhD

Title: Senior Researcher for Demographic Programs, Social, Economic, and Housing Statistics Division, US Census Bureau

Jason Fields is the Senior Researcher for Demographic Programs and the Survey of Income and Program Participation, establishing scientific objectives based on the needs of the Census Bureau and the external SIPP user community; establishing survey objectives; determining SIPP questionnaire content; and engaging the data user community and other stakeholders to identify evolving needs for data. Jason began his career at the Census Bureau in the Population Division in 1997 as a family demographer in the Fertility and Family Statistics Branch. Since 2006, Jason has been integrally involved with the re-engineering of SIPP, and he has been the SIPP Survey Director from 2012 - 2018. Between 2015 and 2018, he added the National Survey of Children's Health (NSCH), the National Sample Survey of Registered Nurses (NSSRN), and the National Survey of Psychiatrists (NSP). He has promoted the integration of administrative data into production survey processing, adaptive design, and implemented a proactive paradata evaluation, management and quality assurance program. His research interests include survey methodology, the integration of administrative and survey data, longitudinal survey design, disclosure avoidance, and substantive interests include: family demography, living arrangements, health and well-being, and the long-term social and health consequences of childhood and family poverty and instability

Daniel Pratt, PhD

Title: Director, Center for Education Surveys/Vice President, Education and Workforce Development, RTI International

Dan Pratt is Division Vice President of Education and Workforce Development at RTI International, having served in that role since 2020. He leads a staff of more than 200 who leverage cutting edge methodologies, survey techniques, technologies and practical educator experience to improve education and workforce opportunities and to empower learners of all ages. His research interests include education surveys, longitudinal studies, adaptive and responsive design, mixed-modes data collection, and the use of technology in study implementation. Mr. Pratt has nearly 30 years of experience directing large-scale, nationally representative longitudinal studies conducted by the National Center for Education Statistics. He currently leads the High School and Beyond Longitudinal Study of 2022 (HS&B:22). HS&B:22 focuses on understanding how high school experiences affect young adults' learning and their educational and career choices, and explores their transitions from high school to college, the labor force, or adult roles.

Mr. Pratt joined RTI in 1987 as a programmer. With expertise in computer science, data processing, and software engineering, he developed many of the technical underpinnings of RTI's key projects in survey research. Mr. Pratt has presented his research at conferences and workshops and has prepared technical reports and journal articles.

Trivellore E. Raghunathan, PhD

Title: Professor of Biostatistics, School of Public Health, University of Michigan; Research Professor, Survey Research Center, Institute for Social Research.

Trivellore Raghunathan (Raghu) is a Professor of Biostatistics at the School of Public Health and a Research Professor in the Survey Research Center at the Institute for Social Research. He is also a faculty member in the Joint Program in Survey Methodology at the University of Maryland. He served as Director of the Survey Research Center September 2015 – August 2019 and as the Chair of the Department of Biostatistics from January 2010 – August 2014. He is an Associate Director of the Center for Research on Ethnicity, Culture and Health (CRECH). He is a Fellow of the American Statistical Association and has received several awards, including the Remington Methodology Award from the American Heart Association. He received his PhD in Statistics from Harvard University in 1987. His research interests are in the analysis of incomplete data, multiple imputation, Bayesian methods, design and analysis of sample surveys, small area estimation, confidentiality and disclosure limitation, longitudinal data analysis and statistical methods for epidemiology. He has developed a multiple imputation and sample survey analysis software (www.iveware.org) for imputing the missing values for a complex data set.

Jerome P. Reiter, PhD

Title: Professor & Chair, Department of Statistical Science, Duke University

Jerry Reiter is Department Chair and Professor of Statistical Science at Duke University. His primary areas of research include methods for ensuring data privacy, for handling missing and erroneous values, for combining information across sources, and for analyzing complex data in the social sciences and public policy. He is a Fellow of the American Statistical Association and a Fellow of the Institute of Mathematical Statistics. He is the recipient of several teaching and mentoring awards from Duke University, including the Alumni Distinguished Undergraduate Teaching Award, the Outstanding Postdoctoral Mentor Award, and the Masters of Interdisciplinary Data Science Distinguished Faculty Award. He has advised multiple government agencies on creating data products to share with the public, as well as served on multiple panels and committees for the National Academy of Sciences. He received a PhD in statistics from Harvard University in 1999.

Keith Rust, PhD

Title: Senior Vice President & Senior Statistical Fellow, Westat, Inc.

Dr. Rust is a Senior Vice President and was appointed to the Westat Board of Directors in 2018. He is a senior statistician with extensive experience in sampling methods, the design and specification of large-scale sample surveys, and analysis of survey data. His areas of special expertise include methods for analyzing large, complex data sets; methods of deriving survey weights; and sampling error estimation procedures. He has applied his research and knowledge to a variety of education research projects over the past several years, both national and international. Dr. Rust has also directed work on Government sample surveys related to education, health, and social issues.

Dr. Rust is a Fellow of the American Statistical Association (ASA) and served as President of the Washington Statistical Society. He is an elected Member of the International Statistical Institute. He served as Chair of the Programme for International Student Assessment Technical Advisory Group, Organisation for Economic Co-operation and Development, from 2001 to 2020. Dr. Rust was a member of the Committee for National Statistics (CNSTAT) at the National Academy of Sciences from 1992-98

and has served on CNSTAT panels. He is a research professor at the Joint Program in Survey Methodology at the University of Maryland.

Changbao Wu, PhD

Title: Professor of Statistics, Department of Statistics and Actuarial Science, University of Waterloo

Changbao Wu is Professor of Statistics in the Department of Statistics and Actuarial Science at University of Waterloo. His main research interests include design and analysis of complex surveys, resampling techniques, missing data analysis and causal inference, and integration of data from multiple sources. He is Fellow of ASA, Fellow of IMS, Elected Member of ISI, and was the winner of the CRM-SSC Prize in Statistics in 2012. He has served on several editorial boards including Survey Methodology, The Canadian Journal of Statistics, JASA T&M and Biometrika. He is the lead author of the book “Sampling Theory and Practice” (with Mary Thompson) published by Springer in 2020. He has also served on Statistics Canada’s Advisory Committee on Statistical Methods since 2015.

Panel II

James Appleton, PhD

Title: Executive Director, Office of Research and Evaluation, Gwinnett County Public Schools, Atlanta, GA

James Appleton, PhD, is the executive director of research and evaluation at Gwinnett County Public Schools, a district of over 180,000 students, in Georgia. He is co-developer of the Student Engagement Instrument (<https://checkandconnect.umn.edu/sei/default.html>) with most of his peer-reviewed contributions focused on student engagement with school. As a practitioner, he leads the implementation of board policy on educational research and provides senior support on the conceptualizations around, and system of, outcome measures in support of policies on academic and operational accountability. His interdisciplinary team consists of staff with a range of training (mostly PhDs, a MPH, and MD). The purpose of the Office of Research and Evaluation is to provide direct research and evaluation support to the Superintendent and to the Chief Strategy and Performance Officer for the purpose of accelerating school and system-wide improvement.

Debbie Herget, MS

Title: Survey Director, RTI International

Debbie Herget is a director in RTI’s Education and Workforce Development Division. She has more than 25 years of experience leading school-based studies and data collection efforts, with more than 20 years leading NCES longitudinal and international comparison education studies and data collections. She is an expert in school-based survey design and data collection, with particular expertise in gaining cooperation from school districts and schools. In addition to recruitment and data collection, her roles include project and task management; preparation and management of complex budgets; managing large teams of staff and subcontractors; developing and implementing protocols, plans, and training packages; preparing Office of Management and Budget (OMB) and institutional review board (IRB) packages; and preparing study reports.

Rachel Horwitz, PhD

Title: Lead Scientist, Survey Evaluation and Improvement, US Census Bureau

Rachel Horwitz began her career working on economic surveys and then shifted her focus to household surveys, specifically the American Community Survey. She currently leads methodological research for the Census Bureau’s demographic surveys, including the National Survey of College Graduates and the

Current Population Survey. Her research focuses on reducing measurement and nonresponse error, understanding respondents' interactions with surveys through paradata, and developing design standards for web surveys. She has also examined methods to optimize the number and type of contacts to sample cases to reduce respondent burden and cost while maintaining or improving response rates.

Rachel has a refereed journal article in *Social Science Computer Review* and chapter in the book *Advances in Questionnaire Design, Development, Evaluation and Testing*, both on using mouse movements to predict difficulty in web surveys. She has also authored and coauthored multiple technical papers at the Census Bureau. Rachel has a PhD in Survey Methodology.

Regina Lewis, PhD

Title: Coordinator of NAEP and International Assessments, Maine Department of Education

As a liaison between the National Center of Education Statistics and the Maine DOE, Dr. Lewis provides important information and feedback between the national, state, and local education agencies, with the primary role of supporting the National Assessment of Educational Progress (NAEP) and any surveys administered in association with NAEP. As a part of that role, she has reviewed supportive materials and surveys for the NAEP State Coordinators as well as those provided to state and local agencies. At Maine DOE, she is currently collaborating on the analysis of reading assessment measures, local benchmarks, achievement data as well as reading instructional programs and resources utilized by schools throughout the state. The investigation is focused on which measures, programs, resources, practices, and educator supports are yielding student growth. Dr. Lewis continues to work as an advocate ensure that all students matter, exemplified by her current contributions as the primary author of the Maine DOE Assessment Technical Assistance Plan, an approach to monitoring state assessments focused on improving the consistency and equity of assessment for all students through collaborative partnership. Dr. Lewis is a current and active member AERA and serves on the NCME Committee for Informing Assessment Policy. She has served in multiple roles of the field of education from interventionist and educator to that of a current member of the board local Catholic school. Dr. Lewis earned her PhD in Education, with a specialization in Curriculum, Instruction, and Assessment, at Walden University.

Jerome P. Reiter, PhD

Title: Professor & Chair, Department of Statistical Science, Duke University (see above)

Robert Sivinski, MS

Title: Senior Statistician, Office of Statistical and Science Policy, Office of Management and Budget, Executive Office of the President

Bob is a methodologist for the Office of the U.S. Chief Statistician in the Office of Management and Budget (OMB). His principal roles are to promote methodological and statistical rigor in data collection, regulatory analysis, and program evaluation across the Federal government, and to promote innovative improvements to Federal statistics. His current portfolio at OMB includes demographic, education, and justice statistics.

Before coming to OMB, Bob worked for the National Highway Traffic Safety Administration at the Department of Transportation, where he designed surveys, contributed to benefit/cost analyses, and conducted retrospective reviews of traffic safety regulations. He is a member of the Federal Committee

on Statistical Methodology, chair of the public policy program of the Washington Statistical Society, and an active member of the American Statistical Association and American Association for Public Opinion Research.

Kevon Tucker-Seeley, PhD

Title: Director, Research & Reporting Branch, Los Angeles Unified School District

Dr. Tucker-Seeley has over 25 years of experience in K12 education at the school-, district-, state-, and national level. His areas of expertise include student data privacy, student assessment administration policies and protocols as well as assessment and survey design, development, and implementation. Currently Dr. Tucker-Seeley serves as the Director of Research and Reporting in Los Angeles Unified School District's Office of Data and Accountability. Previously, Dr. Tucker-Seeley has served as the Student Assessment Program Manager for the state of Rhode Island's Department of Education; a Research Associate in the Institute of Education Sciences' Regional Educational Laboratory-Northeast & Islands; an Assistant Editor of the Journal of Technology, Learning, and Assessment; a K12 Gifted Education Specialist; and an Elementary School Teacher. Prior to entering the field of education, Dr. Tucker-Seeley traveled the world as a Satellite Communications Specialist in the United States Air Force and is a proud veteran of Operation Desert Storm.

Gina K. Walejko, PhD

Title: Senior User Experience Researcher, Google

Before working at Google, Gina Walejko worked as a survey methodologist for the U.S. Census Bureau for seven years. Her work has focused on customer satisfaction surveys, adaptive survey design, interviewer compliance, reducing U.S. Decennial Census undercounts, audience segmentation, as well as message design and testing.

She teaches survey methodology courses for the University of Maryland's Joint Program in Survey Methodology, the University of Mannheim's Master of Applied Data Science and Measurement Program, and the International Program in Survey and Data Science.

She is coauthor of many technical papers, and her refereed journal articles have appeared in publications including Science and Public Opinion Quarterly. Her areas of technical expertise include questionnaire design and message testing. She currently serves on the American Statistical Association's (ASA) Survey Research Methods Section Executive Council and on the ASA's Committee for Women in Statistics. She has also been elected and served as both program chair and president for the Washington-Baltimore Chapter of the American Association for Public Opinion Research and program chair for the Government Statistics Section of the ASA.

National Institute of Statistical Sciences

Nell Sedransk, PhD

Title: Director, National Institute of Statistical Sciences-DC

Nell Sedransk is the Director of the National Institute of Statistical Sciences. She is an Elected Member of the International Statistical Institute, also Elected Fellow of the American Statistical Association. She is coauthor of three technical books; and her research in both statistical theory and application appears in more than 60 scientific papers in refereed journals. The areas of her technical expertise include: design of complex experiments, Bayesian inference, spatial statistics and topological foundations for statistical theory. She has applied her expertise in statistical design and analysis of complex experiments and observational studies to a wide range of applications from physiology and medicine to engineering and sensors to social science applications in multi-observer scoring to ethical designs for clinical trials.

Brian Habing, PhD

Title: Associate Director for Education Research at NISS, and Associate Professor of Statistics at University of South Carolina

Brian Habing is Associate Director for Education Activities and Research working with the DC Office of the National Institute of Statistical Sciences (NISS) and Associate Professor of Statistics at the University of South Carolina. His research has focused on psychometrics and scale construction, with a particular emphasis on multidimensional item response theory. His research focus includes analysis of education statistics and his interests also extend to statistical education, including work with AP Statistics and the development of new courses at the undergraduate and graduate level.

Ya Mo, PhD

Title: Research Fellow, National Institute of Statistical Sciences; Assistant Professor, Boise State University

Ya Mo is a research fellow at the National Institute of Statistical Sciences and an assistant professor of Curriculum, Instruction, and Foundational Studies at Boise State University. She received a dual major PhD in Measurement and Quantitative Methods and Curriculum, Instruction, and Educational Policy Programs, and an M.S. in Statistics from Michigan State University, as well as an Ed.M. in TESOL from Boston University. She researches quantitative methods, psychometric measures, and survey statistics; she also applies quantitative research methods to study substantive topics in education, especially large-scale assessments.