Institute of Educational Statistics
National Center for Education Statistics

NATIONAL INSTITUTE OF STATISTICAL SCIENCES
EXPERT PANEL REPORT

POST-COVID SURVEYS

TABLE OF CONTENTS

NATIONAL INSTITUTE OF STATISTICAL SCIENCES

POST-COVID SURVEYS

EXECUTIVE SUMMARY

The Covid-19 pandemic and the consequent changes to lifestyles and education practices have focused the American attention intensely on education and learning, both processes and progress. The shifts from in-classroom to virtual learning and from social to asocial learning and isolated environments have profoundly affected students and teachers as well as education administrators at all levels who are faced with making decisions based on whatever information is available. In consequence the context for national surveys and assessments through the remaining pandemic period and into the post-Covid era will in many ways be new, while contextual changes already underway are accelerated.

Important changes are variously attitudinal (i.e., increasing reluctance to participate), transactional (i.e., desires for *quid pro quo* for participation), content-related (i.e., information on measures of learning delays and strategies for recouping), and technological (i.e., greater familiarity and facility with electronic media and virtual classrooms).

The expert panel identified three broad areas for addressing these changes:

- Recognition and Trust of IES/NCES
- Relationships with Educators, Respondents and other Stakeholders
- Opportunities Arising from Technical Advances and Available Technologies

Broad recommendations in each of these areas are given below. More specifically targeted recommendations appear in the final section of the full report and are appended in abbreviated form to this summary.

Major difficulties with data collection ultimately manifest in non-response; and this is of significant concern when it impacts data quality. For NCES the impact is most acute at the school district level where refusal eliminates all district schools from all (non-mandatory) surveys and assessments. Therefore the expert panel framed part of its discussion specifically at the district level and offered some specific recommendations.

Recognition and Trust

First and foremost, IES/NCES needs to function and to be recognized as a single entity rather than a loose assemblage of independent surveys, assessments and data resources. In particular, this includes branding with a universal, single, authoritative brand (logo, letterhead, introductory phrase for study and personnel identification, etc.) that dominates individual survey/assessment identifications and contractor brands. This will allow IES/NCES to unify the recruitment process (especially at the district level) for the upcoming (multi-year) collection of surveys and assessments. When IES/NCES is consistently presented as a single entity, requests for particular studies can draw on the overall value of IES/NCES surveys, assessments, data releases and publications.

Second, IES/NCES needs to actively listen to the full range of stakeholders to ascertain the value to them of IES/NCES surveys, and to hear their unmet wants, needs and requirements. The panel notes that the collective of IES/NCES stakeholders is extremely varied: government, corporate and public, including policy and decision makers, educators at all levels – administrative oversight through classroom – plus parents and students who are the survey respondents. IES/NCES currently engages effectively with groups of stakeholders, but this is not the case for all.

Third, IES/NCES needs to facilitate and expand its reliance on trusted voices in the communities of respondents, educators and education decision-makers. These range from professional organizations to community-oriented individual coordinators (following the successful NAEP model) for a few districts with highest probability of inclusion (e.g., Los Angeles county) and for populations that self-define (e.g., Native American, by region). Advocacy from trusted voices can explain and support the importance of the survey to an uninterested or a skeptical population. Such advocates also provide IES/NCES with information on reluctance.

Relationships and Participation

Survey success depends on approval and agreement at every one of multiple levels from state to student. And at each level, the question is: "What is in it for *me* [my school/district/state]?" A satisfactory response is predicated on a two-way relationship rather than a one-way data flow.

Financial incentives (above the level of pizza lunch for the class) are not feasible. Non-financial incentives include: input to data queries/information needs, information returns with local relevance, dual use of survey information by school and NCES, recognition as "proud participant," and input to NCES website products (like "kid zone").A potentially attractive feature to facilitate participation is the efficiency via simultaneous negotiation for multi-survey (multi-year) package, especially with the attractive provision of limiting/eliminating repeated requests within a few years to the same schools for participation . The key here is for open input from administrators, educators and even parents on what would be valuable returns.

One example is the new study with NICHD that offers direct benefit to students who participate. A different example would be setting up brief, rapid response surveys ("pulse") surveys by polling participating districts for topics requiring information for upcoming decisions.

The challenge here is to achieve a balance between disclosure risk and plaudit for participation; different kinds of solutions may be needed at different levels.

Opportunities

As the year of the pandemic wears on in "virtual mode," schooling, work, shopping, other ordinary activities and much communication are primarily electronic, with consequent overall rise in level of competence with electronic media. This creates the opportunity for mounting surveys virtually, with the particular mode suited to the respondent (for example, computer/tablet for students; mobile app for parents, online for teachers) and the data collection type and stage (such as computer for initial survey contact attempt, videocall for follow-up; online data transfer for administrative data).

Preparing for such data collections implies taking advantage of technological possibilities to streamline the process. User Experience (UX) research is extensive and is not just about access or ease of use.

Engagement is a major focus, producing advances in the forms of animation, conversation, responses to free-form inquiry, prediction and individual feedback. Designing data collection software "instruments" *ab initio* to take advantage of these capabilities can short circuit irrelevant items, prepopulate items from federal data files and public sources, verify responses that deviate from expectation, create dialogues, prompt or amuse along the way, and allow free instead of preprogrammed responses by using automated text readers during data compilation. Taking electronic data collection as the starting point, paper versions can still be created from these as needed.

Hardware evolution means that once required a computer, then a tablet, now only needs a flash drive. Extensive experience with virtual environments means that surveys can be largely managed remotely and/or on a broader scale or a shorter time frame with no loss (even potential gain) in individual attention to respondents.

Advances in statistical theory and practice for surveys comprise the third major collection of opportunities. For example, combining non-probability surveys with traditional probability survey designs can potentially improve precision and coverage without increasing cost. Combining information across multiple probability surveys also makes sense when recruitment is consolidated and the survey designs can be integrated.

Summary of Primary Findings and Recommendations

CENTRAL CONCEPT

IES/NCES: *Implement a recognized functionally coherent and transparent structure* to replace/reorganize loosely connected collection of separate surveys and assessments

RECOGNITION AND TRUST

Create: A single universal brand for IES/NCES, dominating all individual study or contractor identifications. Inclusion of NAEP under this brand will strengthen recognition.

Consolidate: Recruitment with sole recruiter for all surveys and assessments; negotiate for program rather than individual survey/assessment.

Facilitate: Advocates for IES/NCES, from professional education associations to more local "trusted voices" including (a few) coordinators for largest districts and for special populations.

RELATIONSHIPS

Create: Partnership relationships with education administrators and educators to incentivize participation and to invite input on process, content and feedback.

Undertake: Study to characterize NCES data uses and users; and a formal study of non-responding school districts to ascertain motivation and identify barriers.

Design: Useful, prompt feedback with local relevance.

Offer: Option to schools to expand a survey/assessment to their (entire) population, with centralized processing, returned with a demographic data file and a summary report or a school profile for assessment or other privacy-protected data.

Find: Ways to create pride in participation at every level; design to circumvent/minimize potential for disclosure.

OPPORTUNITIES

Transition: To electronic/virtual mode data collection with paper form and format only created as needed. Don't request information that is already available (from federal data files or public sources); only request corrections to prepopulated items.

Consider: Using dual frame sample designs, and/or panel designs to avoid redundancy and to integrate across surveys.

Consider: New designs incorporating data from non-probability sample to provide preliminary (stage 1) information while retaining probability samples (stage 2) for estimates.

Develop: Rapid new "pulse" surveys for quick topical information needed for upcoming decisions or summaries.

Work: With OMB to speed approval process, especially for pulse surveys, by separating the approval process into template approval (burden evaluation) and item bank or items (content evaluation).

Create: Analysis tools for schools to calculate local "expectations" based on national data but adjusted to local school characteristics.

PREFACE

Covid-19 resulted in the disruption of education in the United States on a national scale and also altered the modes and opportunities for surveys. In 2020, the National Center for Education Statistics (NCES) charged the National Institute of Statistical Sciences (NISS) with examining the changing context for surveys: changing needs for information both content and timing, changing willingness to participate, and changing facility of the general populace with virtual communication emerging as the norm. The panel was asked specifically to address the changing needs and changing opportunities for traditional and for new data collection modes in order to continue to conduct reliable, efficient surveys of the many facets of information on education. With a national transition to primarily virtual education (and virtual workplaces), the panel was asked to consider next steps and the expanded use of technology by NCES to maintain or increase engagement of survey participants while meeting the national need for information on education.

The panel of experts held four virtual meetings between 2 and 11 December 2020. At the first meeting NCES staff presented details of the panel's charge. Following deliberations in closed sessions, the panel presented preliminary findings to NCES staff at the fourth meeting. Further panel teleconferences were held during the preparation of this report.

NATIONAL INSTITUTE OF STATISTICAL SCIENCES
EXPERT PANEL REPORT

POST-COVID SURVEYS

I. INTRODUCTION

Surveys in the post-Covid era will need to adapt to a context that is dynamic in all aspects: respondents, methodology, and content. Changes already underway in attitudes toward surveys may be accelerated, while opportunities for introducing changes and new practices may become open because of the nation-wide increase in familiarity and reliance on electronic communication and technology that was enforced by the disruption due to the pandemic. For the National Center for Education Statistics (NCES), in face of the upheaval in education in the United States, these impacts of the pandemic are augmented by new needs for information about the pandemic's impacts on learning and about measures taken to mitigate losses.

Foreseeable challenges are increasing reluctance to participate, less willingness to simply "contribute to the common good", more willingness to look for a *quid pro quo*, coupled with greater impatience in seeing returns. The national involuntary experiment in out-of-classroom learning will bring needs in the short term as decision makers, from state education officials to parents, seek rapidly available data on coping with consequences of the long disruption and on strategies for recouping lost education time and progress. In the longer term, researchers will probe more deeply into the learning processes during Covid and the educational and social consequences - both losses and unexpected benefits - data will have to address new questions. Covid has focused the nation on education from immediate adaptations and consequences to prospects for recovery thus heightening the need for relevant and trustworthy information.

Simultaneously, opportunities continue to arise from the expanding array of electronic media and new technology for data collection and information dissemination coupled with advances in statistical survey theory and practice to take advantage of these.

Even before the pandemic, the problems of rising non-response rates and high cost of customary modes of data collection plagued surveys by all federal statistical agencies. For NCES, obtaining responses requires a hierarchy of permissions beginning at the state or district level and continuing down to school, teacher and finally parent/respondent level. Non-response at the (state or) district level carries the greatest consequences because it eliminates all schools within the state or district from participation.

The expert panel considered these challenges and opportunities in terms of the changing context, focus and process, coupled with increasing pressures for ever greater currency and relevance of information released. Taken altogether, deliberations on these issues then led to re-examination of some traditional practices.

Four themes permeated all aspects of the panel's discussions: 1) a consistent, recognized and trusted identity, 2) a unified operational plan encompassing all NCES data collections, 3) development of two-way relationships or partnerships with gatekeepers, stakeholders and respondents, and 4) innovation in survey design and technology.

The overall findings and global recommendations are summarized first. The sections of the report that follow present for each theme the motivation and rationale for broad recommendations and some more specific suggestions for consideration. Section VII contains a hypothetical scenario to illustrate how the four themes can interrelate. This scenario is not intended as a recommendation but rather as a prompt for wide-ranging thinking about opportunities presented by the post-Covid era challenges.

II. OVERVIEW

The challenge to NCES is to provide the information for decision-making going forward without undervaluing either contemporary snapshots or more comprehensive retrospective research. Meeting this challenge will first involve reviewing, rethinking and then taking a global approach to NCES surveys and assessments.

The central concept envisioned for NCES is of a coherent overall structure that reorganizes or replaces the current looser confederacy of surveys and assessments. Externally this implies presenting an immediately recognized, consistent NCES/IES/ED[1] identity and widening its visibility. Internally this includes strategically unifying or integrating NCES operations.

The first of two dimensions of this challenge is defined in terms of the relationships that connect NCES with data providers and data users. In the past respondents supplied NCES with data and NCES delivered information to decision-makers and researchers. Changing these one-way relationships to partnership will require understanding the needs and constraints of data providers and gatekeepers in order to provide a useful return for their investment in participation. Similarly a successful two-way relationship with current and potential future stakeholders depends on their priorities and needs for information and the time windows for information relevance and usefulness.

The second dimension of this challenge is defined by the push to move aggressively into the digital age. This implies full use of digital tools for gathering information, surveys and assessments redesigned *ab initio* for electronic communication, sampling schemes streamlined by incorporating auxiliary (non-probability as well as probability) data, and the use of technology and electronic media to engage respondents and to expand the stakeholders and data users.

Both dimensions of the challenge are already familiar as NCES currently puts considerable effort into relationships with stakeholders and is already piloting and experimenting with digitally-based data collections. (Almost all data collections have been conducted electronically for quite some time.) The next steps are to continue the evolution on a comprehensive scale with an overall coherent, consistent plan.

III. RECOGNIZED AND TRUSTED IDENTITY

Recognition and trust will predictably be more important in the post-Covid era than they were before: Recognition, especially by education administrators from the state and district levels on down; Trust, especially by respondents, from students and parents on up the administrative chain.

[1] NCES/IES/ED: National Center for Education Statistics/Institute of Education Sciences/US Department of Education

Recognition. The long-time practice of operating IES/NCES surveys independently, including NAEP, has not served NCES well for building a consistent, recognizable identity. The result is a cacophony of logos and representations, no one of which stands out except, perhaps for NAEP which separates itself from the rest of IES/NCES. (Figure 1, Appendix C).

Competition for time and attention to surveys and assessments can be expected to intensify post-Covid while at the same time demands for information increase. The array of new and important research questions revolving around the pre-Covid/post-Covid changes in learning and in education promise an abundance of proposals for surveys and assessments. These will surely include requests for students' time in classroom and out, requests for teachers' plans and practices in shifting between in-person and virtual classes, and both teachers' and administrators' evaluations of teaching practices, strategies and decisions along the way. Simultaneously, students' in-school time (and teachers' time) will become an even more precious commodity with the effort to catch up for lost learning.

Presenting the request for each IES/NCES survey individually makes it difficult to argue for priority or otherwise to single out IES/NCES requests. Different recruitment personnel for each NCES study or assessment, likely coupled with changing personnel in education offices, especially district offices, exacerbates the problem.

"*Hello*, *I am* [*name*] *from* [*contractor*] *to talk about the* [*study title*] *study that IES/NCES plans to initiate in 2022 for the US Department of Education*." This self-introduction reverses the order of importance by putting the person, the contractor and the study in front of the agency and department. In so doing it serves to disconnect each study from other Department of Education studies and data collections, making it easy to treat NCES studies as similar in value/priority to the wide range of requests from university and institute researchers and graduate students, and other organizations.

For IES/NCES to function and to be recognized as a single entity, its hallmark needs to be a universal, single, authoritative brand. Branding works when it is comprehensive, consistent and visible. In the context of NCES data collections, the branding begins with the marketing plans and solicitation of approval from state and district administrators down to respondents (students, parents, teachers, local decision makers). Branding continues through the process to the dissemination of results.

In short, successful branding identifies every contact with every individual with a single dominant NCES/IES/ED brand for all surveys and all assessments, i.e., present on all materials, all touchpoints, all media, all contact modes. In deciding on the best single brand, name recognition matters, especially since many if not most stakeholders are unfamiliar and uninterested in the complex structures within the federal government. A particular study may also carry its own secondary brand when this is useful (e.g., for the longitudinal studies), but always under the dominant brand. It is worth noting that when the Bureau of the Census studied options for use in the decennial census, the most official-looking logo dominated others (more colorful, more designed to attract attention) as a brand.

Visibility is the other requirement for recognition. Increasing the numbers and kinds of visitors to the NCES website will always be a work in progress; and attention needs to be continually given to attracting new website visitors. The Covid pandemic has put education on the minds of almost everyone; finding ways to connect to new, broader audiences not already engaged via the website. For example, when someone Googles NCES and/or the name of a study, first in the list should be a link to the NCES or study website with the logo and branding prominent. Search engine optimization might be useful.

TRUST

More than just good will, gaining participation from respondents and gatekeepers often depends on trust: trust that information will not be misused or available for misuse and trust in the value of that information.

The panel notes that the collective of IES/NCES stakeholders is extremely varied: government, corporate and public, including policy and decision makers, educators and administrators at all levels plus parents and students who are the survey and assessment respondents. IES/NCES currently engages effectively with some groups of stakeholders, but this is not the case for all.

Facilitating and expanding reliance on trusted voices in the communities of respondents, educators and education decision-makers can be very effective. These range from professional organizations to community-oriented individual coordinators (following the successful NAEP model) for a few districts with highest probability of inclusion (e.g., Los Angeles county) and for populations that self-define (e.g., Native American, by region). Advocacy from trusted voices can explain and support the importance of the survey to an uninterested or a skeptical population. Such advocates also could provide IES/NCES with information on reluctance.

The success of NAEP coordinators demonstrates the value of trusted voices in communicating assurance on both counts. At administrative levels, professional organizations and peer groups take on the advocacy role; for respondents, especially for hard-to-reach populations, trusted voices are more local or specific in their advocacy for a particular group.

While NCES Is unlikely to be able to invest in reproducing the NAEP model for its other studies, a much smaller scale might be considered. For the couple of largest districts, i.e., those that are included in every major survey or assessment, an NCES coordinator might take on the NAEP model task of interfacing at all levels from district to student and parent. Similarly, for the hardest to reach populations, an NCES coordinator might establish a relationship with their trusted voices.

On a broader scale, the goodwill and trust that NAEP coordinators already hold might be tapped for a few days each year. At the least, personal introduction of NCES study recruiters (district through school level) by NAEP coordinators could open the way for establishing a relationship of trust, particularly if recruitment is consolidated.

IV. UNIFIED PLAN

Unifying NCES surveys and assessments spanning several years under a single umbrella identity with a coherent structural core could help to achieve several goals. An integrated survey design is one example.

Recruitment could be consolidated with a single source for all NCES surveys and assessments (NAEP may be the exception). By employing a single recruiter/recruiting contractor to negotiate with a district, NCES would distinguish itself from the myriad of survey permission requests from researchers, students and others. This would also allow development of trust and stronger relationships with gatekeepers and more local trusted voices.

Having multiple studies under a single umbrella would also allow negotiation of new kinds of arrangements with districts for multi-year commitments. A collection of surveys and assessments could be designated, even allowing for as yet unspecified one-time data collections to take place over a multi-year period.

At the same time, constraints desired by the school district could be factored in. For example, specifications for numbers of schools or classes at each level could be anticipated; specification of the particular schools could be subsequently determined, allowing for district restrictions or limit on returns to individual schools during the multi-year period. Incentives could be offered on a combined basis, whether monetary or non-monetary.

Either to eliminate redundancy or to increase efficiency, parallel or related studies could benefit at both design and analysis stages. For example, the populations overlap for TIMSS and NAEP. With careful planning, some information from the two assessments could be complementary.

NCES already conducts a variety of kinds of studies, and advances in technology will enable more. Working from a unified, multi-year plan would allow strategic utilization of different kinds of surveys with capability for continuing adaptation to new data collection designs and modes.

For rapid surveys, discussed below, taking this longer view would allow more efficient approval processes through the Office of Management and Budget (OMB). Review of templates (e.g., for rapid surveys) could be separated from question/item review to shorten the time to putting the survey into the field.

V. RELATIONSHIPS

Survey success depends on approval and agreement at every one of multiple levels from state to student. And at each level, the question is raised: "What is in it for *me* [my school/district/state]?" A satisfactory response is predicated on a two-way relationship rather than a one-way data flow. In other words, the roles are changed and data suppliers as well as data consumers become partners in IES/NCES data collections, with the expectation of mutual benefits.

VALUE PROPOSITIONS

The use case for collecting education data is not difficult to make when focused on national descriptions and trends. Starting from the input of respondents (often students). through the process of making the data made available. to the kinds of questions that can be answered, to the impact of those answers, gives an "answers from data" justification.

However, when these windows on education practices and outcomes are broad-based and/or present a picture that is more than a year out of date, the pool of interested stakeholders shrinks. Making the case for participation to potential data providers is facilitated when benefits are relevant and timely.

The value proposition needs to be clear for each participant/stakeholder group involved in data collection from approval through data submission. Making this explicit requires including the role at each stage, the requirements, and the benefits. The key here is open input from administrators, educators, and parents/ students on what would constitute valuable returns, going beyond ranking from a preset list.

A formal study to characterize NCES data uses and users, focused in particular on data providers, could also ascertain motivation and identify possible incentives/rewards. Extending the study to characterization of non-responding districts and schools could identify barriers and potential compensating or mitigating strategies.

BENEFITS

Financial incentives (above the level of pizza lunch for the class) are not feasible. Non-financial incentives will differ among groups of stakeholders with different roles. Exploring these individually is worth doing with a formal study to characterize NCES data uses and users. Such a study can ascertain motivation, determine the values of possible incentives/rewards, and at the same time identify barriers leading to non-response.

Potential non-financial incentives can focus on an individual return to a respondent, or useful information with local relevance, an opportunity to contribute ideas, a perspective, or opinions, recognition.

One interesting example of direct benefit to students is the new NCES study with NICHD that offers hearing and vision testing to students who participate.

Benefits that NCES could provide at school or administrative levels might take the form of information, recognition or opportunity for input. When another agency joins with NCES to conduct a survey, benefits could extend, as in the case of NICHD, to those the other agency can provide. The unified approach to data collection, outlined above, also enables joint incentives and rewards for multiple studies.

Locally relevant information could be provided as summary statistics from data and/or model estimates "for schools like ours," using available covariates. Or, by creating an analysis tool for the school or district, local "expectations" could be calculated based on national data but adjusted to local school characteristics. An NCES-provided tool might be created to produce summary statistics from more restricted data and/or general enough to be applied to other studies as well as the one in which the school or district participated.

Inviting specific input emphasizes partnership and underlies additional kinds of benefits that NCES could offer. For example, during negotiation for participation, to offer a local option to adapt (electronic) survey modes to respondents' capabilities and preferences, particularly in the case of basic data.

A different example would be the opportunity for input to topics and questions for a continuing series of brief, rapid response ("pulse") surveys of educators or administrators. This would invite the data providers who are also data consumers to participate in setting the data agenda. Extending this to gatekeepers and policy makers, especially at school and district levels, would focus subsequent brief surveys on those topics requiring information for upcoming decisions. Since rapid response surveys typically have very large sample sizes, the opportunity for input could be appended as an option onto the end of each survey. Alternatively, participating districts or personnel could be polled to suggest questions or to rank topics via a data hub open to participants.

Recognition is commonly used in other circumstances as an incentive or reward. Schools are often adorned with banners or signs declaring "proud participant . . ." Other agencies publish lists of collaborating organizations, even with awards like "Cooperator of the Year." When the eventual uses of the data include important administrative decisions such as allocation of funds, the value of participation may be high. Recognition is one benefits that can be simultaneously "carrot and stick." For example, absence from a list, even more strongly removal upon declining to participate, can be a strong motivator.

For NCES the concern with preserving privacy and assuring confidentiality of personal information has limited such recognitions. This is worth reconsideration, particularly if recognition can span both time (e.g., three to five years) and specific surveys and assessments.

The real challenge is to achieve a balance between disclosure risk and positive recognition for participation, with different solutions for different respondents, gatekeepers, and stakeholders.

OPPORTUNITIES

Finding ways to partner with districts and schools might serve to relax the constraints on recognition. Without even sketching details, the opportunities described below are intended to open wider consideration, especially making use of advances in technology and of flexibility in designing for electronic data collection as the primary mode.

For some surveys or assessments, it may be possible to offer schools the individual option of expanding the data collection beyond the sampled students or classes to larger groups, for example, entire classes or schools. NCES would retain the full data but only for the original sample; the school would receive the summary information for the entire class or group. This would enable dual use of the survey or assessment information by the school and by NCES.

Practical advantages to this kind of plan would be the full participation of the class or group without disclosure of the NCES participants. Because the school data is universe data for its own population, summary statistics at the school level could be computed quickly and easily without weighting or imputation. Processing (scoring and summarizing) could be done by NCES before data on the non-sampled students is redacted.

Technological advances have led to “computer on a flash drive” capabilities, so that survey or test administration may not now require importation of tablets/laptops and setting up a temporary local LAN. All that is needed is a device to plug into with a screen and keyboard. Scaling up (from sampled students to classes to schools) is feasible.

Other local options might be possible as well; for example, addition of a question of local importance selected from a few approved items. Results could be returned to the school but not retained by NCES.

VI. INNOVATION AND FLEXIBILITY

The pandemic drove the nation to immersion in remote access and limited classroom experience to virtual connections. But even before that, the changing world of technology was presenting new opportunities for gathering and recording information: New modes of data acquisition, both passive and active; New instruments designed to be adaptive, dynamic and interactive; New responsiveness; New statistical methods for processing and combining information different kinds of and from different sources. Because of restrictions during Covid the populace, and above all, students from pre-K to senior citizens, are now capable/facile users of electronic media. So the opportunity is wide open to shift paradigms from two-dimensional paper and linear layout of survey instruments to primary electronic forms. At the same time, for those who were left behind because of access or other limitations, the disparities are now even greater.

STATISTICAL INNOVATION

To continue to deliver high quality data while adjusting for increasing non-response and decreasing funds is driving the need for innovation in survey design and implementation. Panel discussion centered on the broad theme of replacing a “probability survey first approach” with a comprehensive data collection

strategy that also includes administrative data, commercial data sources, public records, other nonprobability survey data and all NCES surveys to support probability surveys. Under this umbrella there were two subthemes: i) Integrating data across different sources including probability and nonprobability samples; and ii) comprehensive design across multiple surveys.

Currently NCES has a variety of types of studies underway and planned. These include large (omnibus) surveys with and without assessments incorporated, both longitudinal and cross-sectional. Rapid studies have been conducted in the past and in newer forms are contemplated for the future. All are probability surveys and to date, all have been design-based using appropriate frames. Many incorporate adaptive design and/or responsive design to successfully sample rare populations on the one hand and on the other to address nonresponse during the course of the survey to reduce potential nonresponse bias.

Rapid studies may be of particular value during the Covid through post-Covid periods (and in future periods of disruption) because of quickly evolving issues requiring prompt decisions and the need for current information on which to base them. Pulse surveys are under consideration at NCES and can be conducted either as a series to follow change or as snapshots for immediate decision making on issues of importance. These are also probability surveys but they differ fundamentally because the design must be simplified while the sample may be much larger to attain sufficient precision and controlled nonresponse bias. As used by other agencies, notably the Census Bureau, the metric for acceptable estimates is representativeness and precision of estimation rather than a limit on non-response. One key to very brief pulse surveys, especially in a series, is only to ask the crucial questions that cannot be answered with public information (whether located beforehand or after the response has been submitted).

To gain efficiency, some of the federal statistical agencies are using volunteer as well as probability samples in selected circumstances, often to combine available non-probability information with a probability-based sample. Stats Canada and other nations' statistical agencies have some experience as well. For NCES these alternative data sources might be useful as ancillary information. High quality volunteer data might come from an open solicitation of known individuals, e.g., all district superintendents, so that extensive covariate information is available for both responders and non-responders. When the population is completely known (e.g., US schools, school districts), these data could be handled statistically as a universe survey with non-response.

Nonprobability sample data can be used in several ways. First, nonprobability data on a response of interest can be used to establish the existence of some attribute or condition in a population, but with no information on its prevalence, because the nonprobability data are not representative. Second, nonprobability data on predictor variables can be used to improve the efficiency of a probability sample, via model-assisted estimation or small area estimation with nonprobability data as covariates, or via Bayesian methods with nonprobability data used to construct prior distributions. Third, if both covariates and the response of interest are available in the nonprobability data, then several methods can be used to combine the representativeness of the probability sample with the extent of the nonprobability data to allow proper inference to the population: for example, sample matching, mass imputation, inverse propensity weighting, and doubly-robust estimation. This is an active area of research.

Low quality volunteer samples might come from direct solicitations without access to covariate information, from websites of unknown quality, from web scraping, or from other open calls. Usual problems are coverage issues and unknown probabilities of selection; but although not sufficiently accurate

or precise to be used as prior data for a designed sample, this supplemental information may have value in reflecting a general level of interest, for example the importance of a contemporary topic to a particular audience.

Strategies to compensate for reduction in funding for follow-up of non-response, and general tightening of budgets for surveys, include considering alternative sampling designs and modes of data collection. One alternative sometimes suggested is to reframe a school-based sample design as a cluster sample because district refusal blocks participation by all schools in the district. However, this technical specification change does not appear likely to solve the problem nor is it likely to change estimates in a material way.

New design strategies are one option; although in building on a good design, the relative merits of further altering the design or of compensating through the estimation process - or both - do not yield a universal answer. The available ancillary information may determine what is possible in a given case. Alternatives to consider here are to use dual frames (e.g., school-based and household based) and/or panel designs to avoid redundancy and to integrate across surveys. A secondary goal is to effectively capture rarer populations and also those most likely to decline to respond. Model-based sampling, based on propensity to respond, can draw on the resources of nonprobability data as auxiliary information to inform both design and analysis of the probability survey.

Designing at a larger scale also presents several alternative strategies that could be evaluated with respect to efficiency and precision of estimation. A consequence of using a comprehensive design for multiple surveys is that contextual information can be shared. This is true even if the surveys are not conducted simultaneously, as is often the case for NCES major surveys and assessments. In addition, as each survey is implemented, its progress, its response rates and the characterization of non-responders informs later surveys and provides information for adjusting their sampling plans.

There are several ways to structure a comprehensive design strategy; in every case the analytic objective(s) need to lead the design. One strategy that has been used by other agencies but not by NCES is a rolling sample with a total refresh at four or five years (n/4 replacements annually either with or without annual updates). This approach becomes feasible if the negotiation with school districts is based on a multi-year plan. In another departure from past practices, if retaining schools in samples for multiple years is an accepted option, then a rolling sample is the natural choice.

Study of the comparative advantages of design compared to analysis tools to deal with (selective) non-response would be welcome. In addition to methodology and approaches described above, the value of techniques like frame supplementation, nonresponse bias analysis and adjustment, weighting procedures and imputation still needs to be evaluated.

TECHNOLOGY-BASED DATA COLLECTION

Perhaps the most obvious change for post-covid data collections is a universal rise in confidence and competence for using electronic media. This will allow tailoring the survey mode to the population being surveyed and the kind of data being collected. For example, one survey might use a portal or hub for administrators, offer parents a cell-phone app, and use a custom LAN for students being surveyed in school. It is generally agreed that assessments work better on larger screens, but demographic and background data for participating students could be gathered earlier or later on using a telephone app.

The most important change, however, will be to move beyond presenting a paper/pencil test on a screen with or without a few added interactive features. NCES has for some time increasingly turned toward this type of electronic data collection. Now the real challenge is to design a data collection instrument and individual items from the start for electronic mode. This means starting from the necessary constructs and then taking advantage of electronic media features to optimize the presentation for engaging the respondent.

It is becoming a common practice to preload information that is available and request that the respondent to confirm/update/correct. Available data are often drawn from previous data collections or administrative records (e.g., school or district files, administrative reports, other federal surveys). More recently, the practice has broadened to include various sources of information of public information to preload as well (e.g., local area information from school/district/town or other websites) since there does not seem any reason to ask the respondent to provide information that is already available.

Because electronic mode data collection can be nimble, the data collection process can be adaptive at the scale of a single respondent. There are possibilities at every scale, starting from item selection based on respondent constraints on time or number of questions, and/or conditional choices of questions based on earlier responses. Respondents can be asked not only about language preference but also be offered translations of individual terms or even translations of their own responses to review both English and non-English versions.

Data collection can become a conversation. Based on available prior information, responses can be predicted (not shown to the respondent), and the respondent can be queried about their answers, “Was that what you expected?” Replies can be scaled and then explained briefly, to be analyzed later by text analysis (AI) software.

To encourage input, free responses can be captured and opinions sought (e.g., “As a teacher, what is most challenging about . . . ?“ or “What would you most like to know about in order to assist students who . . . ?”). To maintain focus on the data gathering process, progress monitors (even playful or interactive) can be interpolated on a fixed or a dynamic basis that depends on the respondent’s time course or other indication of flagging interest. (Fun is possible: Fireworks display on a Fitbit when the milestone - 10,000 steps - is reached appeals to adults as much as to children.)

The key is for electronic/virtual mode to determine the data collection process with paper form or onscreen paper format only created afterward as needed. In the immediate future the question of disparity between respondents with electronic/virtual mode access and familiarity and those without will be an urgent one. Measuring students’ learning loss and the differential between these two groups will be important and challenging with mixed survey and assessment modes. Ignoring the critical segment of the population without full electronic access, and especially without successful remote learning experiences, would be a serious mistake.

VII. AN ILLUSTRATION

Under a unified program of NCES surveys and assessments, technology and the advances in survey theory, methodology and practice enable new options for structuring surveys and assessments. These extend from negotiating participation and returning benefits to the respondents, schools and/or districts to employing non-survey sources preparatory to streamlined probability-based sampling and estimation. The example that follows is intended to prompt consideration of a much wider range of new and proven alternatives as they might work together.

EXAMPLE – RAPID SURVEY

New Demand: For the duration of the pandemic (or future wide-scale disruption) and even more importantly as the nation starts to emerge from this stressful era, decisions are going to be required from education administrators on many new topical issues. These decisions will have to be made promptly based on "best available information at the time;" at the same time the cost for basing decisions on poor quality data can be large. As the nation emerges from the constraints of the pandemic, forecasting the issues that will demand decisive action will not always allow the more than two years' lead time required by complex NCES surveys.

As NCES conjectured in its presentations to the panel, a series of ultra-brief surveys issued monthly or more frequently could be implemented to meet this need. Foreseeably, issues will revolve around the requirements for successful virtual education and then around the practical and curricular issues for returning to in-classroom education.

New Design: "Pulse" surveys (as conducted by survey organizations and by other federal agencies) are short surveys/questionnaires that gather data via electronic communications and use statistical methodology to assess/adjust for bias. Solicitation for participation in these surveys is typically targeted but very broad; and response is acknowledged to be self-selective to some degree.

Publication of results does not depend directly upon the response rate, rather results are published as "best information available," i.e., estimates (bias adjusted) with stated uncertainties, even when the uncertainties are wide.

New Focus: A series of ultra-brief surveys (each as short as 5-20 items) could address information needs projected by stakeholders/decision makers only weeks in advance. To make surveys efficient, information already publicly available would be gathered otherwise (e.g., contractor research) so that items could specifically address the "questions of the day."

As a perk for participating, respondents could be given the opportunity to propose questions or topics for upcoming surveys.

Rapid processing of data to meet a goal of a few weeks from actual survey to posted results is possible only by keeping each survey very simple for analysis and free from complications such as disclosure risks that create time-consuming delays. When useful, reports might also incorporate publicly available information.

Preparation of such a series of ultra-brief surveys could be streamlined if the precedent would apply for separating the OMB assessment of burden (based on employing a single survey template) from the OMB approval of the items themselves (either an item bank or topical sets of individual items).

VIII. SUMMARY OF FINDINGS

CENTRAL CONCEPT

IES/NCES: *Recognized functionally coherent and transparent structure* to replace/reorganize the loosely connected collection of separate surveys and assessments.

RECOGNITION AND TRUST

Create: A single universal brand for IES/NCES, dominating all individual study or contractor identifications. Inclusion of NAEP under this brand will strengthen recognition.

Facilitate: Advocates for IES/NCES, from professional education associations to more local "trusted voices" including (a few) coordinators for largest districts and for special populations.

UNIFIED PLAN

Consolidate: Recruitment with sole recruiter for all surveys and assessments; negotiate for program rather than individual survey/assessment.

Integrate: The varied NCES studies into a unified, multi-year plan that allows strategic utilization of different kinds of surveys and continuing adaptation to new data collection designs and modes.

RELATIONSHIPS

Create: Partnership relationships with education administrators and educators to incentivize participation and to invite input on process, content and feedback.

Undertake: Study to characterize NCES data uses and users; and a formal study of non-responding school districts to ascertain motivation and identify barriers.

Design: Useful, prompt feedback with local relevance.

Offer: Option to schools to expand a survey/assessment to their (entire) population, with centralized processing, returned with a demographic data file and a summary report or a school profile for assessment or other privacy-protected data.

Find: Ways to create pride in participation at every level; design to circumvent/minimize potential for disclosure.

OPPORTUNITIES

Transition: To electronic/virtual mode data collection with paper form and format only created as needed. Don't request information that is already available (from federal data files or public sources); only request corrections to prepopulated items.

Consider: Using dual frame sample designs, and/or panel designs to avoid redundancy and to integrate across surveys.

Explore: New designs incorporating data from non-probability sample to provide preliminary (stage 1) information while retaining probability samples (stage 2) for estimates.

Develop: Rapid new “pulse” surveys for quick topical information needed for upcoming decisions or summaries.

Work: With OMB to speed approval process, especially for pulse surveys, by separating the approval process into template approval (burden evaluation) and item bank or items (content evaluation).

Create: Analysis tools for schools to calculate local “expectations” based on national data but adjusted to local school characteristics.

APPENDICES

Appendix A: Agenda

NATIONAL INSTITUTE OF STATISTICAL SCIENCES

NCES EXPERT PANEL ON CLEAR DATA DESCRIPTIONS IN NCES REPORTS

December 2-3, 9, 11, 2020

AGENDA

Wednesday, December 2

9:00 am – 12:30 pm	Open Session
9:00 am – 9:20 am	Welcome and Introduction of Panel and NCES staff
9:20 am – 11:00 am	Review of Charge and Presentation of Work at NCES
11:00 am – 12:30 pm	Questions from the Panel and Discussion
12:30 pm – 12:50 pm	Break
12:50 pm – 1:30 pm	Panel Executive Working Session (closed)

Thursday, December 3

9:30 am – 1:00 pm	Panel Executive Working Session (closed)
9:30 am – 10:30 am	Panel Deliberations (closed)
10:30 am – 10:45 am	Break
10:45 am – 1:00 pm	Panel Deliberations (closed)

Wednesday, December 9

9:00 am – 12:30 pm	Panel Executive Working Session (closed)
9:00 am – 10:45 am	Panel Deliberations (closed)
10:45 am – 11:15 am	Break
11:15 am – 12:30 pm	Panel Deliberations (closed)

Friday, December 11

10:00 am – 12:00 pm	Open Session
10:00 am – 11:00 am	Review and Presentation of Findings by Panel
11:00 am – 12:15 pm	Questions from NCES Staff and Discussion

Appendix B: Charge to Panel

The ramifications of the Covid 19 pandemic include a more technologically savvy populace, increasing unwillingness to participate in governmental efforts, as well as the obvious and enormous shift to virtual mode for education at all levels. At NCES the readily apparent necessity for change is accompanied by the opportunity to redesign data collection practices, also incorporating changes that are overdue. Hence, the NCES Commissioner, Dr. Lynn Woodworth, is asking the panel to take a *tabula rasa* approach to two inextricably related aspects of future NCES data collections and to provide recommendations about the most fruitful avenues to pursue.

- Engagement of participants, whether establishments (school, governmental education office, etc.) or individuals (education administrator, teacher, student, parent)
- Sampling design and data description/estimation methodology

Appendix C: Figure 1: Branding Cacophony

Appendix D: Expert Panel Biosketches

F. Jay Breidt, *Ph.D.*

Title: Professor of Statistics at Colorado State University

Jay Breidt is an expert in survey sampling, time series, nonparametric regression, and uncertainty quantification for complex scientific models. He received his Ph.D. at Colorado State University in 1991 and spent nine years at Iowa State University as an assistant professor and tenured associate professor before returning to Colorado State in 2000. He served as Chair of the Department of Statistics from 2005 to 2010 and Interim Chair from 2019 to 2020. Breidt has presented over 130 invited short courses, conference talks, and academic seminars. Since 1991, his research has been supported continuously by a variety of agencies including the National Science Foundation, National Institutes of Health, Department of Homeland Security, Environmental Protection Agency, US Forest Service, and NASA. Breidt has been an associate editor for eight journals and reviews editor for Journal of the American Statistical Association and The American Statistician. He has served on six review committees for the National Academy of Sciences. He is past Chair of the American Statistical Association National Committee on Energy Statistics, has served two terms on the Federal Economic Statistics Advisory Committee, and is currently a member of the Census Scientific Advisory Committee. Breidt has been recognized with a national prize in environmental statistics, elected membership in the International Statistical Institute, and elected fellowship in the American Statistical Association and the Institute of Mathematical Statistics.

Jason M. Fields, *Ph.D., MPH*

Title: Senior Researcher at the U.S. 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.

Rachel Horwitz, *Ph.D.*

Title: Survey Methodologist at the U.S. 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 Ph.D. in Survey Methodology.

Natalie Shlomo, *Ph.D.*

Title: Professor of Social Statistics at the School of Social Sciences, University of Manchester

Natalie Shlomo is a survey statistician with interests in adaptive survey designs, data linkage and integration, statistical disclosure control and small area estimation. She is an elected member of the International Statistical Institute (ISI) and a fellow of the Royal Statistical Society. She served as Vice-President of the ISI 2017-2019 and is currently serving as an elected council member till 2021. Natalie also serves on editorial boards of several journals and is a member of methodology advisory boards for National Statistics Institutes in the United Kingdom, Sweden, Italy and Canada. She is the UK principle investigator for several collaborative grants from the 7th Framework Programme and H2020 of the European Union, all involving research on improving survey methods and dissemination. She held a number of grants in the United Kingdom and is currently the principle investigator for the Economic and Social Research Council (ESRC) funded research on Methodological Advancements on the use of Administrative Data in Official Statistics 2021-2023. She has over 60 journal articles and refereed book chapters. Homepage: https://www.research.manchester.ac.uk/portal/natalie.shlomo.html

James Wagner, *Ph.D.*

Title: Research Associate Professor at the University of Michigan's Survey Research Center (UM-SRC)

James Wagner's research is in the area of nonresponse and methods for addressing it during data collection. In particular, he has focused on the use of responsive and adaptive designs for controlling nonresponse. He has also worked on statistical decision rules for supporting these types of designs. He has published articles in a variety of journals including Public Opinion Quarterly, Statistics in Medicine, Journal of the Royal Statistical Society, Journal of Official Statistics, and others. He is co-author of a book (2017) entitled Adaptive Survey Design. In addition to research in the area of survey nonresponse, Dr. Wagner has more than 20 years of experience in the area of sample design and has worked on a large variety of samples. He teaches courses on statistics, sampling, and methods for dealing with nonresponse during data collection through the Michigan Program in Survey Methodology and the Joint Program in Survey Methodology. He has served as a consultant to several federal statistical agencies on methods for improving data quality.

Gina Walejko, *Ph.D.*

Title: Researcher at 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.

Panel convened by National Institute of Statistical Sciences

Nell Sedransk, *Ph.D.*

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, *Ph.D.*

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.

Alexi Albert

Title: Research Assistant, National Institute of Statistical Science

Alexi Albert is a Research Assistant at NISS, working with the National Center for Education Statistics (NCES). She earned her MS in Mathematics and Statistics with a concentration on Data Science from Georgetown University. Prior to graduate school she taught sixth grade math for two years in Eastern Washington State.

Ya Mo, *Ph.D.*

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 Ph.D. 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.