

Institute of Education Sciences  
National Center for Education Statistics

NATIONAL INSTITUTE OF STATISTICAL SCIENCES  
TECHNICAL EXPERT PANEL REPORT

**NISS/NCES PANEL ON MAKING  
NCES PROCESS DATA AVAILABLE**

## TABLE OF CONTENTS

I.	Introduction .....	7
	NCES Data .....	8
	Data Policies at NCES .....	8
II.	Data Requirements .....	9
	Researchers' Information Needs .....	9
	Technological Expertise .....	9
	Analysis Methods .....	10
III.	Data Security and Risks .....	10
	Security, Privacy and Disclosure Risks .....	10
IV.	Recommendations .....	11
	1. File Structure .....	11
	2. Tiered Access .....	14
	3. Resources for Data Users .....	15
	4. Repository for Created Measures and Summary Variables .....	16
	Appendix A: Charge to Panel on Making Process Data Available .....	18
	Appendix B: Agenda .....	19
	Appendix C: Expert Technical Report Panel Biosketches .....	20

# NATIONAL INSTITUTE OF STATISTICAL SCIENCES

## NCES–NISS PANEL ON MAKING NCES PROCESS DATA AVAILABLE

October 2019

### Executive Summary

Most National Center for Education Statistics (NCES) assessments and surveys are now conducted using electronic modes. In consequence, the data captured include not only responses but also time-stamps and click-by-click chronicles of the response process. These data offer unique insight into the cognitive processes involved in test-taking. They also have potential use for automated scoring and may be useful from a psychometric point of view in evaluating item properties. NCES is now committed to making these process data available to researchers.

Therefore NCES charged the National Institute of Statistical Sciences (NISS) with convening a panel of technical experts to advise on how to present these data efficiently, effectively and comprehensively without incurring unacceptable security, privacy or disclosure risks.

Process data comprise the time and action information as a student progresses through the assessment so that the pattern of a student's attention is evident as is the time spent with each action. Questions that are not answerable from response data alone can be studied from process data by the broad community of test developers, psychometricians, behavioral psychologists and many education researchers.

From these time/action data researchers can study student's progress through the assessment including use of tools, time spent on prompts and questions and patterns and sequences for question response and review/correction. Analysis of how time is spent during progression through the assessment may illuminate more elusive concepts such as engagement or lead to definition of new constructs. Differences in test-taking patterns and implications for performance scores can be studied for subgroups of students such as English language learners (ELL) or for students with particular needs such as dyslexia. Process data can also be used during the process of test development as psychometric properties are evaluated; and test design could specifically address new constructs or new metrics.

NAEP process data for released items will provide an initial example. Unlike complete assessment data releases, process data will be made available for individual released items combining student records over several years the item appeared in NAEP assessments.

The immediate challenge is that raw process data are extensive and not structured for analysis or direct interpretation. To date, analyses of process data (from other sources than NCES) has required creation of custom data files for each project or proposed analysis. The charge to this panel is to consider the technical issues in structuring NCES process data so that research can be conducted. *Substantive issues of what specific research or policy questions should be addressable from these data are reserved for a separate panel to consider.*

The panel's deliberations covered two broad areas: the data users' information needs and technology capabilities, and the collective concerns surrounding security, privacy, and disclosure risks. In addition the panel pointed out the need to allow for evolving future demands as research leads to new constructs and new metrics, adopts new technical approaches and creates new computational methods.

Data users (for now) can be assumed to have two levels of technological expertise for dealing with these data: those with resources to extract information from "raw data files" into custom files and those without. For the latter group, much research can be done using well-chosen *summary variables*, if NCES provides these in addition to the raw process data files. Such summary variables would be created from the raw process data to constitute a basic set of summary information (e.g., frequencies, sequences and intervals such as time allocation by item/question, frequency, type and sequence of change(s) to original response, use of available tools such as spell/grammar check).

Specific to IEP and ELL research: Many if not most researchers need matched subsets of non-IEP or of non-ELL students as the basis for comparative studies. NCES is best able to create these subsets efficiently and cost effectively.

Risks to be minimized include: privacy of students, also teachers and schools; security against frivolous or irresponsible data use, disclosure of item information especially for unreleased items or any other breach that would lead to loss of trust in NCES integrity in administration of assessments or loss of trust in NCES data quality.

Two risks deserve formal risk assessment to determine their levels and potential impacts. It is NCES practice not to disclose writing samples. Release of raw text process data would violate this practice because these would allow reconstruction of original writings. To be acceptable an alternative would have to allow access to process data without access to actual content. Process information on unreleased items needs evaluation of a different risk. The panel offered the caveat that this impact might take the form of revealing sufficient information to infer item content or to give information about correct/incorrect responses.

The four-tier hierarchy already in place for access to NCES data should function equally well for process data, incorporating one specific provision for raw text process data.

Panel recommendations center on assembling into a three-part record for each student:

- A. Demographic information (the same information that is available for student-level information for NAEP).
- B. Basic set of summary variables (specific for each subject-matter area to item type, when possible, otherwise specified for the particular released item).
- C. Raw process data (for released items) with raw text and writing samples redacted.

Public information (Tier 1) would include summaries from A and B that parallel the summaries of other assessment data. Protected information (Tier 2) would draw on A and B using Data Explorer (or other NCES software) in the same way that other assessment data is accessed via Data Explorer. Restricted information (Tier 3 - NCES license required) would give the researcher the option of the data files including A and B or (for researchers with the resources to extract individual custom files) the complete data files A and B plus C (released items only). Highly restricted information (Tier 4 - FSDRC only) would restore redacted information to C. At NCES option, based on assessment of risk, more extensive information about unreleased items might be made available at this highest level of restriction.

## MAKING NCES PROCESS DATA AVAILABLE

The panel recommends that process data for released items be released in five sets of files. The first file would be the complete population of all students. In addition two pairs of subfiles would meet the needs of IEP (Individual Education Plan) and ELL (English Language Learner) researchers. Much research in these areas requires comparison with non-IEP/non-ELL students. Therefore data for the special group of students, IEP or ELL, would comprise the first subfile. The second subfile in each pair would contain records for (preferably 3) matched non-IEP/non-ELL students.

For unreleased items, the current practice of releasing item properties could reasonably be extended to summary reporting of summary variables at least for those data that: 1) can be meaningfully interpreted in the absence of the item itself and 2) do not violate disclosure or otherwise carry an unacceptable risk.

The panel further recommended that resources on process data content and available analytic tools be made available to researchers. A primer in the NAEP process data with small examples would serve researchers accessing public and protected data (Tier 1 and Tier 2 access). Also a smaller raw data file (“sandbox”) would allow researchers (Tier 3 access) to write, test and validate extraction and analysis code.

A public repository is needed for new summary variables, constructs and metrics to enable data use to evolve as research on process data advances. At least equally important is a repository for code to allow researchers to share the (costly) efforts of writing extraction and data integration code for these process data.

NATIONAL INSTITUTE OF STATISTICAL SCIENCES TECHNICAL EXPERT PANEL REPORT

## Preface

Now that most National Center for Education Statistics (NCES) assessments and surveys are conducted using electronic modes, electronic data capture means that in addition to basic background information and final responses, the data include documentation of the process of responding. These process data comprise a time-stamped, click-by-click record of each student's progress through the assessment or survey. To date the detailed process data have not been made available; however NCES is now preparing to establish a mechanism for release of process data for research purposes. Therefore NCES charged the National Institute of Statistical Sciences (NISS) with convening a panel of technical experts to advise on how to present these data efficiently, effectively and comprehensively without incurring unacceptable security, privacy or disclosure risks. The specific charge to this panel is to consider the technical issues in structuring NCES process data so that research can be conducted.

The panel held several teleconferences and met in person at NCES on 15-16 October 2019.

# NATIONAL INSTITUTE OF STATISTICAL SCIENCES TECHNICAL EXPERT PANEL REPORT

## NCES PANEL ON MAKING NCES PROCESS DATA AVAILABLE

### I. INTRODUCTION

Most National Center for Education Statistics (NCES) assessments and surveys are now conducted using electronic modes. This allows the data captured to include not only responses but also time-stamps and click-by-click chronicles of the response process. These data offer the potential of unique insight into the cognitive processes involved in test-taking. In addition, these data may be useful for automating scoring, and from a psychometric perspective may provide new information on item properties. NCES is now committed to making these process data as available and as accessible to researchers as is consistent with protection against unacceptable security, privacy or disclosure risks.

Process data pose particular technical challenges to researchers because the data are not structured for analysis, unlike standard assessment data that are structured with a fixed set of fields, each field being of fixed length. To prepare the raw process data for analysis, the first step is to extract and organize the relevant data for the individual research project at hand. This is non-trivial and generally requires the skills of an experienced data base technician – beyond the capabilities of most individual researchers. To date, analyses of process data (from sources other than NCES) have required creation of custom data files for each project or proposed analysis. This places a heavy burden on researchers and can put research projects out of bounds due to the cost of preparing the data.

Therefore NCES charged the National Institute of Statistical Sciences (NISS) with convening a panel of technical experts to advise on how to make process data available at a technically sophisticated level and also in a more accessible form to serve a broad range of researchers including both those able to handle process data in a raw form and those who are not. The challenge is to present these student-level data efficiently, effectively and comprehensively without incurring unacceptable risks to security, privacy or disclosure.

In considering how to structure process data to enable research, the panel discussed how to extract and present the data, whether/how to integrate demographic and other descriptive data, what kinds of quantitative research methods the data would support, how to ensure privacy of personal information, and how to preserve the integrity and security of unreleased assessment items.

*The substantive issues of what specific research or policy questions should be addressable from these data was reserved for a separate future panel to consider. However, the data types needed for answering these questions have implications for defining the data structure.*

The panel's deliberations covered two broad areas: a data structure to meet data users' information needs and technological capabilities, and the collective concerns surrounding security, privacy, and

disclosure risks. Current NCES policies and practices address these broad areas for release of demographic and response data. Therefore the panel took these as a starting point and then considered the implications for translating, extending or modifying these policies and practices for the context of process data. Since NAEP will be the initial focus for release of process data, the panel treated NAEP released items as first exemplars, then followed with extension to unreleased NAEP items where possible.

### **NCES Data**

The raw process data are essentially a time-stamped click-by-click record of a student's response. These data allow a one-click-at-a-time reconstitution and visualization of the response record from opening the assessment until its termination. Records therefore are of variable length, and progress through the item is not in a uniform pattern. Data capture has also changed over time, starting out as "everything that could be captured" and gradually regularizing to recording practices (e.g., copied v paste-in of a number or text) as well as key strokes and time stamps.

While the differences in response records offer a rich source of information about test-taking behaviors and strategies, they do not allow direct or easy application of analysis tools. Hence the data extraction and data structuring requiring specific advanced technical skills must be done prior to any analysis. Some researchers have the resources for creation of customized data files and will want to continue to probe the raw data; however still more researchers are unable to make use of these data in raw form.

For each individual released NAEP item, the process data files will include students who took that item at any time over the course of several years. However there will not be a complete assessment file for any student only the student's file for the particular item.

### **Data Policies at NCES**

NCES releases assessment information, and in particular NAEP results, that includes demographic information as well as response data using a tiered access system to meet public and research needs. An important exception is for writing assessments: student essays are not released (except for one or two selected exemplars); therefore discussion of process data for writing is separated from discussion of process data for other item types.

The currently structured four-tier access classifications (from broadest to most restricted):

- *Public Information* including summaries at the national level,
- *Protected Information* allowing summarizing information for subsets of students (e.g., demographically defined) using calculation tools provided by NCES
- *Restricted Information* requiring a license from NCES to access individual student files containing response complete data plus a set of plausible scores for each student (to normalize across specific question sets that make up individual assessments)
- *Highly Restricted Information* requiring a license from NCES and available only onsite at Federal Statistical Data Research Centers only.

For released items, information at the Public and Protected levels includes the item itself, the item properties, demographics and response information in the form of data summaries calculated from the

Restricted level files using NCES analysis tools (NAEP Data Explorer or NAEP Question Tool). Restricted level files have complete demographic and response information on released items for all students who took the released item. For unreleased items, item properties (e.g., IRT information) are available for all scaled items but the items themselves are not made available. Student-level information includes demographics, response information and plausible scores.

It is useful, therefore, to regard the Restricted level files as the Primary data base. Computational tools, NAEP Questions Tool and NAEP Data Explorer, calculate standard descriptive statistics from this primary data base. Variables for subsetting are limited and disclosure protections are built into both tools, for example requiring a minimal subset size for reporting summary statistics and a minimal cell count for cross-tabulations.

## II. DATA REQUIREMENTS

### Researchers' Information Needs

The prospect of applying process data analysis to a large-scale assessment is relatively new and definitely exciting to a wide range of researchers whose information needs differ. Research can be anticipated that will focus on the students in an assessment situation about how they proceed through the test and about their level of engagement. Such measures would include students' time allocations among questions/parts within the item, students' response-checking for individual questions, orders of proceeding with/without returning to earlier responses, tendency to revise responses, as well as overall test-taking strategies. All these ways of characterizing students' progress through the item can be analyzed for subsets of students, for example defined by demographic variables. For subgroups defined by IEP (individual education program) students or ELL (English language learners), how students use the provided tools and accommodations will also be of research interest. Similarly, specific questions related to the mode of testing such as use of tools can be studied overall or for specified subsets of student.

From a different perspective, students' cognition and behavior can be studied from these data to identify, characterize and measure attentiveness, frustration, confidence or latent constructs. Research in these areas might focus on identification of characteristic patterns, whether of response or time or sequence.

A third area of investigation would be the assessment itself, with a goal of understanding not only the particular item but also of designing future tests that not only measure performance but also illuminate behavior or cognition and provide sensitive measures of underlying constructs defined by behavioral or cognitive research.

### Technological Expertise

Data users can be assumed to have two levels of technological expertise for dealing with these data: those with resources to extract information into custom files and those without. For the latter group, much research can be done using well-chosen *summary variables*, if NCES provides these in addition to the raw process data files. Such summary variables would be created from the raw process data to constitute a basic set of summary information (e.g., frequencies, intervals such as time allocation by

item or question, sequences of questions visited, frequency, type and sequence of change(s) to original response, use of available tools such as mouse or spell/grammar check).

As for any research area in its infancy, the sets of core constructs and relevant measures change rapidly as research results accumulate. Whatever set of summary variables is created initially will certainly deserve review, revision and expansion within five years; in consequence the file structure should be planned to allow for change. Also, as new measures are proposed and studied, researchers need a reference source where calculations are defined for the data base that was used.

Specific to special education research, including both IEP and ELL students is the need of many, if not most, researchers for matched subsets of non-IEP and of non-ELL students as the basis for comparative studies. The process of creating these files represents an additional time and effort investment for each research project, and would result in much duplication of effort and expenditure of research resources. NCES is well able to create these subsets efficiently, a cost-effective solution enabling wide use by the IEP and ELL research communities.

### **Analysis Methods**

At the most public levels of access, Public and Protected, the analysis methods are effectively limited by the summaries and tools NCES provides. Expansion of this library to include multiple regression and higher-order cross tabulations would be useful.

At the Restricted access level analysis methods will be limited only by software available to the researcher. New methodology and data visualization for high-dimensional data are now added to the general classes of models, descriptive statistics and graphics that are used with the currently available data on responses. Computationally intensive model building and deeper exploratory analysis is possible with newer more powerful software suited to big data, for example machine learning and AI. As these lead to definitions of new constructs or relationship within the data, development of new measures, indices, models, and graphical displays will follow. Consequently, the basic set of summary variables can be expected to grow over time.

On occasion patterns identified by these computational tools will be serendipitous rather than accurate and relevant. Reproducibility assumes new importance, especially before adoption of new summary variables into the basic set provided at the Protected level.

## **III. DATA SECURITY AND RISKS**

### **Security, Privacy and Disclosure Risks**

Current NCES security policies are designed to minimize risks of several kinds and to protect privacy of students teachers and schools, to provide security against frivolous or irresponsible data use, to prevent disclosure of item information especially for unreleased (NAEP or other assessment) items, and to forestall any other breach that would lead to loss of trust in NCES integrity in administration of assessments or loss of trust in NCES data quality. NCES security practices, requiring licenses for access to restricted data files, already address the five dimensions of data security for assessment files:

- People,

- Projects,
- Settings,
- Data itself,
- Output.

People and Projects are authorized before data files are made available; and specifications for control of physical access must meet NCES guidelines. Output in the form of publication or dissemination of data analyses are screened for improper disclosures. NCES controls the data to be released. So the risks until now have been primarily through frivolous use of the data or access by irresponsible people - and penalties are steep.

What is different about process data? The data itself and whatever potential for disclosure it carries.

For multiple choice, matching and similar kinds of answers, the fact that the sequence of key strokes alone reconstructs the response provides no more information than is available already through the current files of response data. Adding timing, the visit/revisit pattern and sequence of response changes give information about the item itself - precisely the kind of information that researchers want. However, perhaps this information needs to be withheld for unreleased items, although summary measures might be released at the Restricted level but not at Public or Protected levels (consistent with current practice).

At the opposite extreme of complete disclosure is the student essay which is completely reconstructed by just the key stroke sequence. Since it is NCES policy not to release student essays, releasing the key stroke sequence would be unacceptable. Given the motivation and the technical sophistication, with today's technology the likelihood is that even a coded writing sample could be reconstructed relatively accurately. Risk of disclosure from the coded or partially redacted data should be assessed prior to release and will require a sufficiently high degree of technical knowledge. On the other hand, defining summary variables and measures and using process records may not carry the same risk. To enable researchers to create and validate new summary variables and measures NCES could consider allowing strictly onsite work at a Federal Statistical Research Data Center using a single or a few essays at a time.

## IV. RECOMMENDATIONS

The panel emphasized that the following recommendations are a starting point for data files and resources that will need flexibility to adapt as research using process data evolves.

### 1. FILE STRUCTURE

*The recommendation is to create a set of 5 files with each of the five available in complete and compact versions.*

The complete files would include all permitted raw process data; these will be large files with variable length student records. The compact versions would omit the raw process data and only include the student information and the summary process variables; student records in these files would be of fixed

## MAKING NCES PROCESS DATA AVAILABLE

length. Data extraction and analysis engines like the NAEP Questions Tool (NQT) and NAEP Data Explorer (NDE) could be used with the compact files. Thus the student-level information has three components.

- A. Demographic information (the same information that is available for student-level information for NAEP).
- B. Basic set of summary variables (specific for each subject-matter area to item type, when possible, otherwise specified for the particular released item).
- C. Raw process data (for released items) with raw text and writing samples redacted.

### The BASIC FILE including ALL Students

A	A	B	C
<b>STUDENT FILE ID NUMBER</b>	<b>STUDENT INFORMATION</b> Demographics Education/School Information Testing Date Item Score	<b>SUMMARY PROCESS VARIABLES</b> (Fixed length file) Includes: Summary Variables for all parts of Item	<b>RAW PROCESS DATA</b> (Variable length file) Omits: Raw Data for parts of Item not otherwise released in full
Student #1			
Student #2			
Etc.			

### The IEP FILE including only IEP Students

IEP STUDENT FILE ID NUMBER	IEP STUDENT INFORMATION	SUMMARY PROCESS VARIABLES	RAW PROCESS DATA
	Demographics Education/School Information Testing Date Item Score	(Fixed length file) Includes: Summary Variables for all parts of Item	(Variable length file) Omits: Raw Data for parts of Item not otherwise released in full
IEP Student #1			
IEP Student #2			
Etc.			

MAKING NCES PROCESS DATA AVAILABLE

The IEP COMPANION FILE including only Non-IEP Students Matched (3/IEP Student) to IEP Students

LINKED IEP STUDENT FILE ID NUMBER	MATCHED STUDENT FILE ID NUMBER	MATCHED STUDENT INFORMATION Demographics Education/School Information Testing Date Item Score	SUMMARY PROCESS VARIABLES (Fixed length file) Includes: Summary Variables for all parts of Item	RAW PROCESS DATA (Variable length file) Omits: Raw Data for parts of Item not otherwise released in full
IEP Student #1	Matched Student #1X			
IEP Student #1	Matched Student #1Y			
IEP Student #1	Matched Student #1Z			
IEP Student #2	Matched Student #2X			
IEP Student #2	Matched Student #2Y			
IEP Student #2	Matched Student #2Z			
Etc.	Etc.			

The ELL FILE including only ELL Students

ELL STUDENT FILE ID NUMBER	ELL STUDENT INFORMATION Demographics Education/School Information Testing Date Item Score	SUMMARY PROCESS VARIABLES (Fixed length file) Includes: Summary Variables for all parts of Item	RAW PROCESS DATA (Variable length file) Omits: Raw Data for parts of Item not otherwise released in full
ELL Student #1			
ELL Student #2			
Etc.			

The ELL COMPANION FILE including only Non-ELL Students Matched (3/ELL Student) to ELL Students

LINKED ELL STUDENT FILE ID NUMBER	MATCHED STUDENT FILE ID NUMBER	MATCHED STUDENT INFORMATION Demographics Education/School Information Testing Date Item Score	SUMMARY PROCESS VARIABLES (Fixed length file) Includes: Summary Variables for all parts of Item	RAW PROCESS DATA (Variable length file) Omits: Raw Data for parts of Item not otherwise released in full
ELL Student #1	Matched Student #1X			
ELL Student #1	Matched Student #1Y			
ELL Student #1	Matched Student #1Z			
ELL Student #2	Matched Student #2X			
ELL Student #2	Matched Student #2Y			
ELL Student #2	Matched Student #2Z			
Etc.	Etc.			

2. TIERED ACCESS

*Tiered access to process data should be consistent with current NCES practice and policies for NAEP response data.*

The currently structured four-tier access classifications (from broadest to most restricted):

- *Public Information\**: published summaries (national level).
- *Protected Information\**: researcher-constructed summaries using calculation tools (NQT or NDE) provided by NCES.
- *Restricted Information*: license required from NCES to access individual student level demographic and response data.
- *Highly Restricted Information*: license required from NCES for access available only onsite at Federal Statistical Data Research Centers.

\* Information at the Public and Protected levels is calculated from the Restricted level files.

Access to subset information (Protected level summaries) depends on NCES computational tools. Extension of NQT and NDE to operate on the Summary Variables is one solution; creation of equivalent software for the process data and demographics is an alternative.

In particular, Public and Protected level access should be to process information for released items only. Restricted files should include information for all students for released cognitive items (excluding writing assessment items).

For released items, the items themselves are released and the master files should include information for all students for released cognitive items (excluding writing assessment items). For unreleased items, while the items themselves will not be released, process data could be available at the restricted level whenever it is established that 1) meaningful interpretation of the data is possible and 2) release will not disclose information that could distort item properties or result in altered assessment behaviors (again excepting writing assessment items).

- *Public Information* would include summaries (released items) from A and B that parallel the summaries of other assessment data.
- *Protected Information* would draw on A and B (released items) using NCES analysis tools (NQT, NDE or analogous).
- *Restricted Information* would give researchers the option of access to student level data either A&B (released and unreleased items) or (for researchers with the resources to extract individual custom files) A&B (released and unreleased items) plus C (released items only).

*Highly Restricted Information* would restore redacted information (released items only) to C. At NCES option, access could be granted based on assessment of risk. Alternatively, researchers could submit code to a trusted party (at FSDRC) to be run and output returned to the researcher without exposing *the* data. Whether or not a similar process could be used for unreleased items is a decision left to NCES.

NOTE: Student essays present a special case as current NCES practice is not to release any student essays. Following this practice will require more strict protections from disclosure since raw process data would reveal content. Overall writing measures, such as text complexity, in addition to score could be reported as stand-alone information if already generated and included in the student record. Similarly, if meaningful, valid and reliable summary variables can be created without access to content, then these variables could be treated as other summary variables. Until the security requirements for the content is established, summary variables for writing assessments should be available only after review for content disclosure potential.

### 3. RESOURCES FOR DATA USERS

*Data resources are needed as a primer to describe and illustrate the process data and to provide test data for validating computer code and computation.*

Resources on process data content and the available NCES analytic tools should be made available to researchers. A primer in the NAEP process data with small examples would serve researchers accessing *Public* and *Protected* data.

At a second level a smaller raw data file (“sandbox”) would allow researchers to write, test and validate extraction and analysis code. Such a testbed is valuable if code is to be published/shared. It becomes crucial if code is to be written to analyze highly restricted data without direct data access but only by submission to a third party.

#### **4. REPOSITORY FOR CREATED MEASURES AND SUMMARY VARIABLES**

*Two searchable repositories should be created for new variables and measures defined for NCES data and for extraction and analysis code.*

The definitions of summary variables are outside the scope for this panel. Nonetheless it is safe to say that whatever initial set of summary variables is created for an item or item type, researchers able to utilize the raw data will develop additional summary variables to enable further refined analyses.

An open repository is needed for exact definitions and references for new summary variables, new constructs and new metrics together with algorithms for their computation. An open repository is also needed for algorithms for data extraction and integration to enable researchers to avoid duplicate efforts in creating the particular subsets and data bases required for various research foci. Posting could be done at the time of disclosure review to ensure completeness of the repository. In future, as some of these variables come into common use, NCES could review usage and expand the summary variables available at the Protected level.

## **APPENDICES**

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

**Appendix A: Charge to Panel on Making NCES Process Data Available**

*"Since the raw process data are extensive and not structured for analysis or direct interpretation, the first panel will consider the technical issues in how to structure process data so that research can be conducted. Aspects to be discussed include how to extract and to present data, how to plan for integration with demographic and descriptive data, how to preserve privacy, how to ensure statistical properties are maintained, and generally what research methods would the data support. The substantive issues of what specific research or policy questions should be addressable from these data will be reserved for a separate panel to consider."*

**Appendix B: Agenda**

**NATIONAL INSTITUTE OF STATISTICAL SCIENCES**

**Expert Panel on NCES NAEP Process Data**

**October 15-16, 2019**

**PCP Building, Room 5170/6083**

**AGENDA**

**Day 1 – Tuesday – 15 October 2019**

8:30am	Arrival and Building Security
9:00am	Welcome
9:15 - 10:15	NCES Expectations and Discussion with Panel
10:15 - 11:15	AIR Discussion with Panel
11:15 - 12:00	Discussion
12:00 – 1:00	LUNCH (on your own)
1:00 - 4:30	Panel Executive Session
4:30 - 5:30	Clarification Requests from NCES/AIR (as needed)
5:30	Adjourn

**Day 2 – Wednesday – 16 October 2019**

8:30am	Arrival and Building Security
8:45 – 10:00	Panel Executive Session
10:00 - 10:30	IF useful – NCES/AIR Responses to Panel Requests
10:30 - 12:00	Panel Executive Session
12:00 – 12:30	Working Lunch
12:30 – 2:30	Summary Session with NCES
2:30 - 5:00	Panel Executive Session
5:00	Adjourn

## Appendix C: Expert Technical Report Panel Biosketches

### S. Lynne Stokes, PhD

*Title: Professor & Chair, Department of Statistical Science, Southern Methodist University*

Lynne Stokes is an expert in surveys, polls and sampling, as well as in non-sampling survey errors, such as errors by interviewers and respondents. She is a Fellow of the American Statistical Association. She recently has conducted research on evaluating the accuracy of contest judges and on improving estimates of marine fishery yields by the National Oceanic and Atmospheric Administration.

She also contributes to the National Assessment of Educational Progress, or “Nation’s Report Card,” examining the way schools and students are selected for the large study. Stokes became a faculty at Vanderbilt University, but in 1979 began working for the U.S. Government as a statistician, first at the Patuxent Research Refuge of the U.S. Fish and Wildlife Service and then at the Census Bureau. She returned to academia in 1984, at the University of Texas at Austin, and moved to Southern Methodist in 2001.

Research Interests: Surveys, Polling and Sampling, Voter Exit Polling, Sampling Methods, Non-Sampling Errors, Non-Disclosure Methodology, Measurement Error, Order Statistics, and Mark and Recapture Methods.

### Matthias von Davier, PhD

*Title: Distinguished Research Scientist, National Board of Medical Examiners*

Matthias von Davier holds the Distinguished Research Scientist position at the National Board of Medical Examiners (NBME), in Philadelphia, PA, where he works on psychometric methodologies for analyzing data from technology-based high-stakes assessments. Until 2016, he was a senior research director in the Research & Development Division at Educational Testing Service (ETS), and co-director of the center for Global Assessment at ETS, leading psychometric research and operational analyses of international large scale assessments conducted by the center.

He is the executive editor of *Psychometrika* and also an editor of *Large Scale Assessments in Education*; previously he was editor-in-chief of the *British Journal of Mathematical and Statistical Psychology*. He is co-editor of the Springer book series *Methodology of Educational Measurement and Assessment*. Dr. von Davier received the 2006 ETS Research Scientist award, the 2012 NCME Brad Hanson Award for contributions to educational measurement, and the AERA Division D 2017 Award for Significant Contributions to Quantitative Methodology.

His areas of expertise includes topics such as item response theory, latent class analysis, diagnostic classification models, and more broadly classification and mixture distribution models as well as computational statistics, person-fit, item-fit, and model checking, as well as hierarchical extension of models for categorical data analysis, and the psychometric and statistical methodologies used in large scale educational surveys. His current work is concerned with extending, implementing and applying multidimensional IRT, IRTree, and latent response models as well as speed/accuracy to process data research using data from NBME/USLME, PISA, and PIAAC and automated item generation using deep learning.

**Brock E. Webb, M.S.**

*Title: Senior Information Technology Policy Advisor Office of Management and Budget (OMB), Statistical Science and Policy (SSP), Office of Information and Regulatory Affairs (OIRA)*

Brock Webb, is currently on special Detail to the Office of Management and Budget as a Senior IT Policy Advisor. In his permanent post at the US Census Bureau, he led the Cloud Program which has transformed the way Census acquires and consumes IT. He also drove the Census acquisition strategy and FedRAMP sponsorship of the New York University's Administrative Data Research Facility (ADRF), which enables joint data sharing and statistical work across multiple Federal, State, and Local partners in support of the U.S. Commission on Evidence Based Policy. Prior to joining Census, Brock was the Chief Engineer for Cloud Computing in the Chief Technology Office at the Department of Defense (DoD) Defense Information Systems Agency (DISA).

**Piotr Mitros, PhD**

*Title: Research Director, Educational Testing Service, Center of Next Generation Psychometrics and Data Science*

Piotr Mitros is the Research Director of the Center of Next Generation Psychometrics and Data Science at the Educational Testing Service. Prior, he served as the founding Chief Scientist of edX, where among his duties, he created and managed what has become among the largest and most impactful educational research data sets in the world, with many terabytes of process data. He has participated in workshops, working groups, and committees on big data in education and the social science for the European Commission, the NSF/CRA, and the National Academy of Education.

**Jingchen Liu, PhD**

*Title: Associate Professor, Department of Statistics, Columbia University; Statistics, Data Science Institute*

Jingchen Liu, PhD, is currently an Associate Professor in the Department of Statistics at Columbia University. He is interested in rare event analysis for stochastic processes and in cognitive assessment. Dr. Liu research interests are Applied Probability: Extreme/Rare-Event Analysis of Gaussian Random Fields, Heavy-Tailed Stochastic Systems, and Random Ordinary/Partial Differential Equations; Theory Development for Importance Sampling and MCMC Algorithms. Psychometrics: various Inference Problems of Diagnostic Classification Models and their applications to Psychiatric and Education Testing Data; Computerized Adaptive Testing.

**Daniel Chenok, PhD**

*Title: Executive Director, IMB Center for the Business of Government Executive Director, IBM Global Business Services*

Dan Chenok is Executive Director of the IBM Center for the Business of Government. Currently he oversees the Center's activities in connecting research to practice to benefit government, with a special focus on technology, cybersecurity, regulation, budget, and acquisition issues. Previously he led consulting services for Public Sector Technology Strategy, working with IBM, government, healthcare, and education clients. In addition, he is a CIO SAGE with the Partnership for Public Service, Fellow of the National Academy of Public

Administration, Chair of the Cybersecurity Subcommittee of the DHS Data Privacy and Integrity Advisory Committee, and Member of the Cyber Policy Task Force with the Center for Strategic and International Studies; previously, he served as Chair of the Industry Advisory Council (IAC) for the government-led American Council for Technology (ACT) and Chair of the Federal Information Security and Privacy Advisory Board.

Before joining IBM, Mr. Chenok was a Senior Vice President for Civilian Operations with Pragmatics, and prior to that was a Vice President for Business Solutions and Offerings with SRA International. Mr. Chenok served in the Office for Management and Budget as Branch Chief for Information Policy and Technology, responsible for oversight of federal information and IT policy, including electronic government, computer security, privacy and IT budgeting. Prior to that, he served as Assistant Branch Chief and Desk Officer for Education, Labor, HHS, and related agencies in OMB's Office of Information and Regulatory Affairs. Mr. Chenok has won numerous honors and awards, including a 2010 Federal 100 winner for his work on the presidential transition, and the 2016 Eagle Award for Industry Executive of the Year.

***Technical Expert Panel convened by National Institute of Statistical Sciences***

**Nell Sedransk, PhD**

*Title: Director, National Institute of Statistical Sciences-DC*

Dr. 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.

**Alexandra Brown, MS**

*Title: Research Assistant, National Institute of Statistical Sciences-DC*

Alexandra Brown is a Research Assistant at the National Institute of Statistical Sciences working under the direction of Dr. Nell Sedransk on projects in education research. She holds a MS degree in Economics and is currently a PhD candidate in Survey Methodology at the University of Maryland.