

Institute of Education Sciences
National Center for Education Statistics

NATIONAL INSTITUTE OF STATISTICAL SCIENCES
TECHNICAL EXPERT PANEL REPORT

CLEAR DATA DESCRIPTIONS
IN NCES REPORTS

TABLE OF CONTENTS

Executive Summary	3
Preface.....	5
Background.....	6
I. Reports, Publications, Communications and Audiences	8
II. Clarity, Consistency and Harmonization.....	10
Recommendations: What to Harmonize.....	14
Recommendations: How to Harmonize	14
Recommendations: Process, Construct, Review.....	15
III. Overarching Recommendations	15
IV. Reference.....	16
Appendix A: Glossary	17
Appendix B: Agenda	22
Appendix C: Expert Panel Members’ Biosketches	23

NATIONAL INSTITUTE OF STATISTICAL SCIENCES

CLEAR DATA DESCRIPTIONS IN NCES REPORTS

EXECUTIVE SUMMARY

The National Center for Education Statistics (NCES) is most well known as a source for assessments of students' proficiency and their education trajectories. NCES also gathers information on other aspects of education including economics, adult education levels of attainment, post-secondary education completion, and contemporary issues such as crime in schools. NCES education data is collected variously from surveys and assessments originating at NCES, from internationally-generated assessments and surveys, from collaborations with other federal agencies, and from the assembly of states' administrative data with federal records and survey data. Therefore, achieving consistency and clarity in language usage across NCES reports is extremely challenging and not always possible.

This panel was charged to address both principles and processes and to set attainable goals for NCES in addressing this issue. To narrow the scope to an attainable objective, the panel excluded discussion of technical statistical terms and options for presentation of technical materials in tables, graphs and text.

The findings and recommendations are divided into three parts: characterizing effective writing in relation to the specific audience addressed, outlining the process of "harmonization" in the context of clarity and consistency, and addressing the practicalities of moving forward with a holistic process for writing and reviewing that can improve consistency and clarity among NCES reports and data products.

The panel met at an in-person meeting on 5-6 February 2018 with further closed teleconferences as this report was being prepared.

Recommendations for Reports, Publications, Communications, and Audiences

NCES disseminates data and results using many different formats and products, including *First Look Reports*, *Statistics in Brief* publications, two-page *Data Point Reports* and NCES blog posts. The goal is to clearly articulate the audience of each NCES published work product and to use content and language appropriate for each target audience.

1. NCES needs to produce unique work products aimed at the different audiences they serve.
2. To develop audience-appropriate resources, NCES should invest in understanding their stakeholders, for example, policymakers, educators, researchers, journalists, data users, and parents.
3. NCES needs to articulate its clear vision of what can be accomplished through its electronic media.
4. NCES should use the affordances of their website and other digital media to produce work products aimed at different audiences.
5. NCES needs a user-centric website with a useful structure that allows stakeholders from various communities to rapidly access information and NCES reports relevant to them.

Recommendations for Clarity, Consistency and Harmonization

Transparency and accuracy of meaning is an even greater challenge when the readerships are diverse, the writers are many, and the data collections have been gathered for different purposes and according to different plans. “Harmonization” is used here to mean a deliberated agreement: 1) to apply one meaning to a particular word or phrase wherever it appears in NCES reports, documents, website or elsewhere; or 2) to select one word/phrase from multiple alternatives designating the same thing for universal usage in NCES reports, etc.

What to Harmonize

6. Harmonization efforts should focus on terms that identify classifiers or descriptors of populations and subgroups (“independent variables,” or factors) with priority given to those that are used for establishing policy and those that are used by multiple surveys and/or data bases to provide cross-cutting information or to examine information over time.
7. Harmonization efforts should focus on providing stable definitions for constructed measures, especially those used in policy-making.
8. NCES should look to other federal agencies for standard terms in their purviews; simultaneously NCES needs to take the lead in defining standard (harmonized) terms relating to education and education constructs.
9. When harmonization is inappropriate or simply not possible, both report writers and NCES editors need to focus on using plain language insofar as possible and providing easy access to precise definitions.

Process, Construct, Review

10. The guidelines (statistical and stylistic) for each product, the process for review, and the standards for review for each product should be clear.
11. A checklist for authors would allow them to pre-review products before submitting for review by other staff.
12. A reviewer’s checklist for external (within NCES) reviewers would specify components for review.
13. Adopt a style (e.g., APA style) for all products.

Over Arching Recommendations

Wherever harmonization is possible and retains accuracy of meaning, it is advantageous; and the harmonized usages should be institutionalized across NCES products. Further, an efficient review process aided by these tools is needed to ensure adherence to these standard usages and to encourage rather than inhibit production of NCES reports.

Additional specifically focused recommendations are included in this report.

NATIONAL INSTITUTE OF STATISTICAL SCIENCES

TECHNICAL EXPERT PANEL REPORT ON CLEAR DATA DESCRIPTIONS IN NCES REPORTS

PREFACE

The National Center for Education Statistics (NCES) charged the National Institute of Statistical Sciences (NISS) with convening a panel of technical experts to examine the issues of clarity, consistency and effectiveness in reporting NCES study and survey results. NCES is perhaps most widely regarded as a source for assessments of students' proficiency and their education trajectories. However, NCES surveys and assessments also gather information on other aspects of education including economics, adult education levels of attainment, post-secondary education completion, and contemporary issues such as crime in schools. NCES education data is collected variously from surveys and assessments originating at NCES, internationally-generated assessments and surveys, collaborations with other federal agencies, and from the assembly of states' administrative data with federal records and survey data.

Therefore, achieving consistency and clarity in language usage across NCES reports is extremely challenging and not always possible. The charge to this panel was to address both principles and processes and to set attainable goals for NCES in addressing this issue.

The panel met first via teleconference to discuss extensive materials prepared by NCES staff and on 5-6 February 2018 met in person with presentations and discussions with NCES staff followed by closed sessions for deliberation. The panel held further closed teleconferences to prepare this report.

CLEAR DATA DESCRIPTIONS IN NCES REPORTS

BACKGROUND

NCES disseminates data and results from a wide variety of surveys and assessments using many different formats and products. The users of these reports include education professionals, administrators, policy-makers and researchers in addition to parents, students and the general public. Products include *First Looks*, *Data Points* and the longer, more comprehensive *Statistics in Brief Reports* as well as an NCES Blog. These products need to meet a variety of standards: relevance for users, effectiveness of communication, defensible and appropriate statistical analysis and graphic presentation, adequacy of documentation and clarity and accuracy of writing.

The users of NCES reports often need to assemble information from several surveys and multiple publications of results. Providing clear data description that information can be interpreted correctly across studies is an important challenge. As the primary federal entity for collecting and publishing data related to education in the United States, over the next five years NCES will be responsible for more than 20 data collections (in the Sample Survey division alone).

Inconsistencies in usage arise because different data collections have different sampling designs, often different sampling frames, or because collaborating agencies have different “standard” definitions of population subgroups, or because more exhaustive or technical reports utilize terms that are standard within the technical discipline but different possible “lay language” substitutions are used in different reports, or simply because a different writer makes different choices in describing the data.

The need to maintain clarity and consistency of data definitions is equally acute for administrative data collections. Particularly difficult challenges arise from the varied definitions used by the multiple sources of these data.

Illustrating the Challenge

Different kinds of discrepancies that result were presented to panel as illustrations. The first is reconcilable by returning to an accurate “dictionary definition.” Similarly, some terms have very specific definitions established by other federal agencies, such as “unemployment rate” which has been defined by the Bureau of Labor Statistics. However referring to dictionary definitions can also become quite complex. For example, consider reporting data from school children who are asked to check one of only two boxes on a form: boy or girl. Should the data be headed “Sex” or “Gender?” Here, the dictionary answer is “Sex.” Enlarging this question leads to a different choice since the Protection of Pupil Rights Act prohibits NCES from asking minors about sexual orientation, but this restriction does not currently apply to gender/gender identity. Thus Gender Identity is a different question while Gender alone may be ambiguous.

The second example illustrated the difficulty of setting limits on definitions of subgroups, especially when the definitions used in the sampling plan were determined in order to optimizing estimates’ precisions. Essentially the same difficulty can arise when the sampling frame is based on different elements than the population being estimated; for example, schools are sampled but teachers’ education levels are reported. Sometimes, the definitions of the sampling unit can be separated from the sampled individuals. For example, “middle schools are multi-grade schools with no grade higher than 8 and no grade below 4”

whereas “middle grade students are students in grades from 4 to 8 regardless of their school configurations.” The choices of grades 8 and 4 as the delimiters would require deliberation.

The related inverse problem is to find useful (non-technical and brief) language to refer to groups where there is no logical basis for setting the boundaries, but boundaries have been imposed for reasons that are ancillary (e.g., sampling purposes or coordination with another agency’s reporting system). For example, consider individuals in the age range of 13-24, and data collections that subdivide these differently. Some of these qualify as “young adults,” a partially overlapping category with “youth” with no universally accepted age to separate the two.

The third class of inconsistencies presented a problem of choice where the available alternatives carry different meanings. The illustration presented was “drop-out rate,” for which different choices of numerator and of denominator are meaningful, depending on the purpose for quoting the rate.

Not all inconsistencies in definitions are avoidable. For example, data descriptions and definitions for international data collections are agreed upon by the participating countries’ representatives; and these may not agree with US usage. Within the United States, different practices in different states can present the same problem for some administrative data.

Having established the nature of the problem, NCES formed a working group to consider the “harmonization” of terms. Samples of its deliberations informed the panel and provided the illustrations cited above among others. Effective solution to resolving inconsistencies necessarily involves the writing and also the reviewing of NCES reports and other products.

Currently the review of each report within NCES is a cumbersome multi-level process without specific review objectives assigned at each level. Consequently, failures to meet the existing NCES standards and guidelines may not be caught until the final within-NCES level of review. Correction or revision requires re-review. Following NCES review, an external review within IES follows is often lengthy (measured in months). The arduous process has become a sufficient deterrent that the preparation of the more extensive *Statistics in Brief Reports* is now infrequent.

Focusing the Panel’s Charge

The important construct for the purpose of this report is:

Clarity: What the reader understands from the text is what the writer intended the text to mean.

Clear data description involves vocabulary; it also involves the location and accessibility of information relevant to the data. To narrow the scope of the panel’s charge to an attainable objective, the panel excluded discussion of technical statistical terms and options for presentation of technical materials in tables, graphs and text. A subsequent panel with statistical expertise is suggested to NCES to be charged with taking the recommendations of current panel for a framework and a process and applying these to the consideration of technical statistical terms and presentation of statistical concepts.

The findings of the current expert panel are divided into three parts: characterizing effective writing in relation to the specific audience addressed, outlining the process of “harmonization” in the context of clarity and consistency, and addressing the practicalities of moving forward with a holistic process for writing and reviewing that can improve consistency and clarity among NCES reports and data products.

I. REPORTS, PUBLICATIONS, COMMUNICATIONS AND AUDIENCES

NCES creates wide variety of data products and produces written text in diverse forms for different audiences and with different purposes. The NCES Blog is written to interest a wide audience; as a blog it can be more informal in style and text must be accessible to the general public. In contrast, more formal language and terminology common to governmental processes may be permissible in a report commissioned by the Department administration or by Congress. Technical terms may be both most efficient and clearest for researchers reading data documentation while non-technical terms conveying the general concept are unclear because of their vagueness.

Speaking directly and effectively to any of these audiences requires first determining the target audience, then setting the level and the vocabulary of the text. Requiring that all communications be at the same level fails to meet the needs of many of the target audiences. (That said, for decision-makers and others, the need for consistency of terms may require correspondences between non-technical and technical terms for comparisons to be made across documents.)

Background

NCES disseminates data and results using many different formats and products. These include *First Look Reports*, *Statistics in Brief* publications, two-page *Data Point Reports* and NCES blog posts. These products need to meet standards in several aspects – relevance for users, effective communication, defensible and appropriate statistical analysis and graphic presentation, adequate documentation, and clear and accurate writing. The goal is to clearly articulate the audience of each NCES published work product and to use content and language appropriate for each target audience.

Audience and Language Context

NCES needs to produce unique work products aimed at the different audiences the Center serves, from parents to policymakers and journalists to researchers. Currently the target audiences for NCES’s various written work products are not clearly defined either to writers or to NCES website users.

NCES should explicitly identify the primary audience of each written work product from the start of the writing process through its editing, review, and approval. NCES currently keeps an internal “Tip Sheet” that defines differences in content and review between a *First Look Report*, *Data Point Report*, and *Statistics in Brief* publication. The audience of each type of written NCES work product also should be clearly outlined in this Tip Sheet. This Tip Sheet should also outline the purpose of each type of written work product to help writers understand their differences. Redefining each type of NCES publication for a specific audience allows multiple publications on the same topic or data collection at different stages in time for different audiences.

Defining a specific audience for each work product would allow NCES to use content and language appropriate for each target audience. Standards can be built around different publications, simplifying the task of clearly articulating complex definitions and statistical concepts to multiple audiences at once. For example, written products targeted at researchers can contain more nuanced definitions of “drop-out rates” and “statistical significance” and can be written at a college level. At the same time, publications for a general audience might use plain language, be written at a lower grade level and use non-technical expressions for technical concepts (with links to full definitions or technical terms). NCES might also consider creating online write-ups clearly defining complicated concepts such as “drop-out rate.” Then work products written for a general audience can contain links to these definition pages where readers can learn more about a term.

To develop audience-appropriate resources, NCES should invest in understanding their stakeholders, for example, policymakers, educators, researchers, journalists, data users, and parents. NCES should understand both the current publications that each audience reads as well as the mediums each audience uses to access these documents (e.g. website blog, downloaded reports, Twitter). If not prohibited by law or regulations, this understanding might be accomplished via data from customer service surveys, focus groups with stakeholders, and analysis of web traffic.

Website

NCES should be able to benefit from the affordances of their website and other digital media in producing work products aimed at different audiences. For example, when NCES releases a report, a blog post highlighting the report's main findings is accessible from the first page of the NCES website. This post addresses a broad audience that includes both readers of a full study report and a more general public readership. This would be the ideal entrée to additional content available online for specific audiences as well as the full report. For example, electronic versions of reports could have supplemental methodological materials appropriate for researchers but not targeted toward teachers, or vice versa.

The need to understand and then to address different groups of stakeholders has been identified previously (see *Integrity, Independence and Innovation: The Future of NCES*, report of a previous expert panel). To meet this need, the Institute of Education Sciences (IES) should develop a user-centric NCES website that allows stakeholders easily to access information relevant to them. The NCES website also should communicate where different audiences should navigate to find written information appropriate to their specific needs. The structure of the separate National Assessment of Educational Progress (NAEP) website accomplishes this and clearly shows how to navigate to resources for educators versus resources for school administrators and parents.

First, NCES needs to articulate a clear vision of what it can accomplish through its electronic media. The quality of all the several types of NCES reports is high. But how these reports are communicated to the general public so that the target audience for each is reached is not completely clear. Examples of readerships interested in the information provided in one or more of these types of reports include:

- Parents of students who are concerned about the quality of the schools and how their school compares to similar-grade schools in other areas of the country and the world.
- Teachers concerned about their schools' effectiveness in preparing their students for college and employment. Also, teachers who wish to use education data and tables in their classrooms.
- General public with interest in how the educational system in the United States is changing, for example comparison of the performances of traditional public, private and charter schools.
- Decision-makers and administrators and their staff focused on issues including teacher qualifications, curricula, attendance and success rates, or on resources and fiscal issues.
- Decision-makers, officials and lawmakers focused on education policy.
- Researchers who use NCES data for many purposes such as to study the learning process, assessment, attributes of learning and learners of all ages or to study education systems and influences of factors that contribute to successful performance.

States are also aware of the various stake-holders in education and many specifically design their department of education web pages to communicate to different groups. For example, some state home pages directly link to webpages for their key constituencies (e.g., Administrators, Teachers, Parents) making it easy for any constituent to find relevant material written with that constituency in mind.

The NCES website at <https://nces.ed.gov/> does not have a similarly useful structure. The current home page focuses on new reports and data tables prepared by NCES, but it is unclear which stakeholders each report would interest. Clicking on the link to the top report “*2016 National Household Education Surveys Program Data File User’s Manual and Data File*” leads only to a technical page that seems of interest primarily if not exclusively to an education professional – and certainly necessary to a researcher. This is a missed opportunity for NCES where a link to the survey itself could provide information about study results that would be of interest a very broad audience including of study participants, their education systems and decision-makers.

Website structure is also lacking for accessing information not on the opening page. While some *First Looks* are available from the home page, there is no easy way to access all of the reports; the situation is the same for *Data Point Reports*. Locating useful education *data* from the NCES website is an even greater challenge. Navigation to find reports must go through such a long chain of links, that in practice few stakeholders may eventually find the information that they need. If only the experienced researcher can locate relevant data, then what is the purpose of the NCES website? NCES needs a useful structure that allows stakeholders from various communities to rapidly access information and NCES reports.

By contrast, the NCES blog is relatively user-friendly. In particular, content usually contains nice summaries of NCES reports including graphs. The NAEP website goes a step further, allowing interactive exploration of NAEP results. This website provides summary information at the top layer, and then the viewer can divide into specific results for states or districts – a good model to think about emulating.

RECOMMENDATIONS – I

1. NCES needs to produce unique work products aimed at the different audiences they serve, from parents to policymakers and journalists to researchers. Broadly accessible lay reports need to communicate the same essential content as supporting technical reports with greater detail that permit validation of results and interpretations.
2. To develop audience-appropriate resources, NCES should invest in understanding their stakeholders, for example, policymakers, educators, researchers, journalists, data users, and parents.
3. NCES needs to articulate its clear vision of what can be accomplished through its electronic media.
4. NCES should use the affordances of their website and other digital media to produce work products aimed at different audiences.
5. NCES needs a user-centric website with a useful structure that allows stakeholders from various communities to rapidly access information and NCES reports relevant to them.

II. CLARITY, CONSISTENCY AND HARMONIZATION

Background

Transparency and accuracy of meaning is an even greater challenge when the readerships are diverse, the writers are many, and the data collections have been gathered for different purposes and according to different plans. Data users compile information and data extracted from multiple publications, often under the presumption that each term for reporting a measurement or an outcome has a single meaning, used identically and consistently throughout all NCES publications. Policy-makers and researchers, alike, use NCES data and reports in a variety of ways, and misunderstandings have tangible costs.

Concern at NCES about consistency in communication led to the formation of an NCES-wide committee to review the way key, commonly reported concepts are reported in NCES publications and to examine the need for greater standardization of definitions across data collections. This committee has produced the

first part of a draft report and made it available to this panel to help frame the more general issues discussed here.

One source of in clarity is polysemy - multiple meanings for a word or phrase, so that legitimate usages of the same word differ among publications or even just between the writer and the reader.

Another source is the obverse - multiple words with overlapping meanings, so that nuances are lost or so that a reader may not recognize one word as the virtual equivalent of another. This particular problem is even more prevalent when different alternative common terms or lay language phrases are substituted for a precise technical term.

The NCES committee focused attention on “harmonization” for several examples of in clarities arising from each of these sources. “Harmonization” is used here to mean a deliberated agreement: 1) to apply one meaning to a particular word or phrase wherever it appears in NCES reports, documents, website or elsewhere; or 2) to select one word/phrase from multiple alternatives designating the same thing for universal usage in NCES reports, etc.

From this foundation the expert panel considered in turn, why to harmonize, what to harmonize as well as what not to harmonize, how to harmonize when indicated and lastly, what to do when harmonization is inappropriate or impossible.

Motivation

The concept of harmonization is important because often a word or phrase will have many meanings, each one valid and relevant in a particular context. However, as a result, readers may have difficulty determining which of the meanings is intended in a given instance. Institutionalizing a harmonized definition across NCES publications and data collections and flagging alternative usages where these appear would resolve many of these in clarities. Further, the value of harmonization, once a harmonized meaning is determined, depends on the specific definition (including its technical meaning, where appropriate) being readily available for the reader to consult. With electronic publication of documents and data, linking key terms to their definitions in a glossary is a simple solution.

The draft report of the NCES committee report identifies examples of a number of different kinds of words/phrases that may require harmonization, ranging from “school” to the usage of “sex” versus “gender.” With each example, the report also presents various options for harmonization. The expert panel spent time reviewing these examples in considering the criteria to apply and the process for appropriate decision-making and implementation. This led as well to discussion of the need for publication of harmonized definitions. Readers, and authors as well, need a glossary of harmonized terms with precise meanings and a convention for flagging non-standard usages or alternative meanings in the text or data presentations. Since harmonization of important terms must be ongoing, an expanded or second glossary is needed for this process to document the decision criteria applied in each case.

What to Harmonize

Each data element presented in a summary table is effectively labeled by the response reported (“dependent variable”) and the population for which it is summarized (“independent variable,” or classifying factor) – effectively the table’s column and row headings. Each of these terms requires a clear definition; those used for reference groups or reference statistics either across surveys and/or data collections or across time are clear candidates for harmonization.

A first class of terms for harmonization effort includes factors used in the sample design and other independent variables or descriptors, should focus on data elements viewed as either independent variables or factors used in the sample design. A second class or terms of equal or even greater import includes constructed measures (e.g., graduation rate, dropout rate). Challenges for the latter often arise because there are several legitimate ways to measure a constructed variable. Each particular measure tells a slightly different story focusing on different policy-relevant issues.

Thus, it is important that while NCES institutionalizes use of the harmonized terms, at the time there are occasions when data publication should summaries based on an alternative term as well. For example, publication of data from international studies should include summaries using the harmonized (US) definitions and also the study-standard (international) definitions. Of course, harmonization is not intended to restrict researchers or decision-makers from more detailed analysis of NCES data to address specific issues or queries.

In considering the candidate list, harmonization for some terms may be clear. When primary expertise surrounding a particular issue or attribute resides in another federal agency, then NCES might logically adopt definitions used by that agency. Doing so could also increase the compatibility with other agencies' data bases and facilitate the integration or sharing of data among agencies. For instance, when it comes to words like *region* or *school locale*, guidance might come from the Bureau of Labor Statistics or the Census Department since they already have geographical or employment related units for reporting. Similarly, for *race/ethnicity*, a joint committee from the federal statistical agencies is expected to release new standard classifications for universal adoption.

In comparable fashion, when dealing with education-specific terms like *instructional or institutional level* or *highest level of education*, NCES should take the lead among federal agencies. In this manner, inter-agency cooperation will not only allow harmonization within a NCES, but also will maximize consistency among the federal statistical agencies.

From an administrative decision-making perspective, the decision to harmonize or not is actually a relatively simple cost/benefit calculation that balances the gains from reduced ambiguity against the losses associated with allowing multiple versions of the measure. From an operational perspective, this means that issues of impact, magnitude of effect, need for consistency across series and over time, stability of meaning and popular understanding of the term may all be part of the decision calculus.

Of course, harmonization is not always appropriate or even possible. For example, harmonization may be limited for international studies when terms reporting US results can be consistent in use of harmonized terms, but terminology reporting international results (including the US) need to conform to internationally accepted standards. Harmonization may not be completely possible when primary sources summarize data differently as can happen in the provision of state-level administrative data. States submitting administrative-level data to NCES young teens may use different age ranges, 10-14 years for some states, 10-16 years for some, still other ranges for other states. In situations like this NCES' options are limited.

Another powerful example concerns socio-economic status and the different ways that the construct can be measured and partitioned. Because NCES is prohibited from asking students about their family income or social behaviors, proxies for socioeconomic status must be constructed. These proxies differ based on the population that they are drawn from; they also are subject to change with time or with policy revision (e.g., expanded eligibility for free/reduce price lunch). In this case, harmonization to produce a universal proxy would clearly be inappropriate.

When harmonization is inappropriate or simply not possible, both report writers and NCES editors need to focus on using plain language insofar as possible, with footnotes giving technical language where appropriate. In addition, if NCES continues to work toward better aligning their reports with their intended audiences, then simpler language can be used in reports designed for broad audiences, while more technical language is reserved for reports for researchers.

How to Harmonize

Criteria for harmonizing revolve around the relevance to policy, the impact on understanding issues in education, the breadth of applicability and also on practical aspects including frequency of use of the term, continuity of series, and popular understanding of the term.

Resolution of harmonization may be single “harmonized” term or phrase or two or more alternatives or it may be clear definitions for multiple categories.

In the case of multiple candidate terms for a single construct, the resolution may be adoption of a standard term or established reporting group from another source (e.g., dictionary, other federal agency or professional standard terminology). The invention of new words is to be discouraged; and caution is needed in introducing a new word into the education context to be sure the word does not already have a precise technical meaning in another field.

Multiple alternative definitions may be unavoidable in some instances, for example in publication of results of international studies for US data users. Resources for established terminology include other federal agencies, such as the Bureau of Labor Statistics (BLS) and the Census Bureau, both of which publish glossaries, including alternative definitions for some terms (e.g., income measures). The American Association for Public Opinion Reporting publishes multiple definitions for response rate. A useful feature of the websites for BLS and for the Census Bureau is information on which surveys use which measures and where to find specific types of information.

Defining data ranges or compiling refined categories into broader groups can follow established standards where these exist. Or these may be defined by inclusion/exclusion. For example, a “middle schools” might be defined to include “schools with at least two grades from 5 to 8 but no grade above 9 and no grade below 4.”

Resolutions should be documented in an expanded glossary to include the criteria applied and the rationale for the decision.

Process, Construct and Review

NCES disseminates data and results using many different formats and products. These include *First Look*, *Statistics in Brief*, and two-page *Data Point Reports*. These products need to meet a variety of standards – relevance for users, effective communication, defensible and appropriate statistical analysis and graphic presentation, adequate documentation, and clear and accurate writing. Some of these standards are evaluated in one of several reviews before a report is released. Each of these standards requires that both writers and reviewers of these reports, staff that can change over time, locate and learn the standards and conventions for each product. Some of these standards are communicated in a “Tips” sheet, in NCES Standards for “Statistical Analysis, Inference, and Comparison,” and the like. These “tips” and standards provide a basis for developing and centralizing more systematic templates, checklists, and other resources to integrate the process of writing and reviewing the various NCES products.

Existing resources such as the NCES Tip Sheets and the NCES Standards for Statistical Analysis provide useful guidance, but it is not always clear in the tip sheets, for example, which features are optional. A comparison of existing *Data Point* products to the relevant tip sheet, for example, show differences in approach. Although some adaptations may always be appropriate, the framework within which variations are made should be clear. The development of templates provides an opportunity to engage both product developers and users an opportunity to increase the effectiveness of these outreach and communication efforts. For example, a *Data Point* needs to present an adequate brief description of the study and sample on which it is based, but beginning with this description (rather than, for example, placing it in a standard “data” box), momentarily distracts the reader from the topic of the product.

Producing consistent, and consistently high-quality reporting products requires a process that supports staff by helping staff more easily locate consistent standards, speeding reviews, and minimizing the need for revisions. Templates, checklists, style sheets, and such are resources to clarify both the goals and the process of writing and reviewing research products.

Reviewing within NCES is multi-stage, but currently without specific objectives at each stage. Consequently, every reviewer assumes responsibility for reviewing all aspects of content, organization, writing style, conformity to standards for inclusion/exclusion of specific content as well as grammar, spelling and copy editing. A more efficient process would divide these responsibilities so that a first level (peer?) review could ensure grammatical correctness and correct usage of terminology and conformance to the template. A supervisor's review could concentrate on technical correctness; and so forth. This ordered review structure would reduce the need for entire re-reviews following revisions that could have been made at the first or second level of review.

Training Staff and Contractors

Success in institutionalizing harmonized terms and in adherence to existing standards and templates for publication of NCES data will depend on training for staff and contractors. With new staff at NCES and with the very large number of contractors, new and old, who write reports, set up and label tables, and prepare materials for dissemination, NCES is encouraged to provide training in three forms. First, a review of current standards, templates and tip sheets (with updates); second, training in general writing for NCES audiences; third, training in the use of standard terms and the glossary with explanations. It is important that the training opportunities be open to contractors, especially for those who are engaged in writing the project and study summary reports.

RECOMMENDATIONS – II

Recommendations: What to Harmonize

1. Harmonization efforts should focus on terms that identify classifiers or descriptors of populations and subgroups ("independent variables," or factors) with priority given to those that are used for establishing policy and those that are used by multiple surveys and/or data bases to provide cross-cutting information or to examine information over time.
2. Harmonization efforts should focus on providing stable definitions for constructed measures, especially those used in policy-making. However, NCES is encouraged in addition to continue reporting on refined, specific measures that address in greater detail the study objectives and/or special interests of stakeholders.
3. NCES should look to other federal agencies for standard terms in their purviews; simultaneously NCES needs to take the lead in defining standard (harmonized) terms relating to education and education constructs.
4. When harmonization is inappropriate or simply not possible, both report writers and NCES editors need to focus on using plain language insofar as possible and providing easy access to precise definitions.

Recommendations: How to Harmonize

1. Priority for harmonization should be based on:
 - potential for ambiguity,
 - importance to policy-making including mandated reporting,
 - magnitude of effect,
 - availability of standard usage,
 - need for consistent usage across time or surveys,
 - compatibility with other (federal) data bases.
2. When multiple definitions are appropriate in different circumstances, reporting should be at least dual with both the harmonized standard usage and the particular relevant alternative.

3. If standard usages have been developed, especially for non-education related terms, these definitions should be adopted unless there is a strong reason otherwise.

Recommendations: Process, Construct, Review

1. The guidelines (statistical and stylistic) for each product, the process for review, and the standards for review for each product should be clear. Clarity and consistency over products and over time can be improved by using templates for each product (rather than “tips”), checklists for authors and reviewers, style sheets and glossaries, and similar resources, collected in a central location. Templates, checklists, and other resources should also reduce the need for revisions after review, and so speed release of products.
2. A checklist for authors would allow them to pre-review products before submitting for review by other staff. The author’s checklist would clarify what reviews will be conducted (e.g., for appropriate analysis, clear language) and the goals of each.
3. A reviewer’s checklist for external (within NCES) reviewers would specify components for review (e.g., graphs stand alone, standard use of statistical terminology, clear language). These reviewer’s checklists would be available to authors as well as to external reviewers, and, by consulting them, authors could better prepare documents for review.
4. Adopt a style (e.g., APA style) for all products. A style sheet with model tables and graphics, a glossary of preferred terms, and examples of revising for clear language would be additional resources. These resources, templates, and checklists would be provided for those producing products within NCES and contractors.

III. OVERARCHING RECOMMENDATIONS

1. Recommendation
 - a. Wherever harmonization is possible and retains accuracy of meaning, it is advantageous; and the harmonized usages should be institutionalized across NCES products.
 - b. An efficient review process aided by these tools is needed to ensure adherence to these standard usages and to encourage rather than inhibit production of NCES reports.
2. Recommendation
 - a. The NCES committee on harmonization or other mechanism for vigorously continuing to address clarity issues is needed on a permanently active basis.
 - b. Tools, including a published, continually expanding glossary, are needed for writers of NCES products, including both staff and contractors.
3. NCES should provide leadership in defining terms, phrases and subgroupings related to education while NCES should take advantage of terms, phrases and subgroupings defined by other agencies within their purviews (e.g., BLS for economics, Census for demographics, NCHS for health, etc.).
4. “One Size Fits All” approach to language of all NCES products does not serve the multiple target audiences well. Difficulty level of 12th grade may be appropriate for administrators and policy-makers while technical language is appropriate for researchers. For products designed to be broadly accessible to the general public, language should be non-technical at a suitable level. Differing criteria in the review process can establish and support the target audiences for different products.
5. Website and electronic media should address different audiences directly and provide easy (!) access to reports and other products.

IV. REFERENCE

National Institute of Statistical Sciences Technical Expert Panel. (2016). *Integrity, Independence, and Innovation: The Future of NCES*.

Appendix A: Glossary

EXAMPLE GLOSSARY

This glossary is an example of a document intended to provide guidelines for authors and reviewers with terms and constructs that may have different meanings for different consumers of written reports about education. It includes candidate terms and constructs mostly drawn from the panel’s discussion and the report of the “Harmonizing Committee” that preceded it. The comments include thoughts about the likely resolution of each harmonization effort. The entries in this glossary are intended to be *illustrative of the thought process*. Recommendations are not intended to be conclusive. The entries in this example glossary are terms and concepts deemed noteworthy based on one or more of the criteria listed below. Note that terms and concepts - even if polysemous - would not be included if judged unlikely to cause significant misinterpretations of major points in a report.

Criteria for inclusion in Glossary (in rank order):

1. High likelihood of being misunderstood;
2. High policy relevance;
3. Impact of misconstruing the term;
4. Common usage but with multiple meanings that have important distinctions and with no single “standard definition;”
5. Concepts or variables for which the measurement rules vary from survey to survey or among reports from a single survey;
6. Consistency across federal data bases.

Such a glossary is intended to perform three functions: to identify terms and concepts that may be problematic and meet one or more of the criteria above; to note reasons why multiple meanings may result in ambiguity when clarity is needed; to recommend an approach for disentangling the multiple meanings. The approaches generally fall into five categories:

- Harmonize meanings;
- Refer to a “dictionary definition” established by other federal agencies;
- Specify clearly how a term or concept is being used when several meanings are in common usage;
- Specify limits that define classes of observational or analytical units; Adapt a term or phrase from “common usage,” explain it clearly, and use it consistently throughout the report.
- Adapt a term or phrase from “common usage,” explain it clearly, and use it consistently throughout the report.

The examples below are all judged (1) to lead to possible misunderstandings, and (2) to have relevance for education policy and practice. The accompanying comments illustrate some of the interpretive complexities. The approach to clarifying a concept depends on the level of specificity required by the particular audience. (We reiterate the need for an in-depth understanding of the categories of consumers of NCES products, their needs for specified levels of information, and their understanding of technical detail.)

Note that the electronic dissemination of reports can make several approaches accessible for a single term or concept.

GLOSSARY ILLUSTRATION WITH COMMENTARY

School Location	Comments
<p>Urbanicity <i>(Sociological term not in most thesauruses)</i></p>	<ul style="list-style-type: none"> → Common language suffices for many consumers (urban, suburban, and rural). → First Look (NCES 2018-052) lists 11 categories, likely to be useful for reports that focus on geographic distinctions.
<p>U.S. or National Results</p>	<ul style="list-style-type: none"> → Common understanding is 50 states plus the District of Columbia. → Needs clear explanation for other designations. → Needs names of other noncontiguous regions included, (e.g., Puerto Rico, Guam, US Virgin Islands), if separate results given in report for any reason.

School Characteristics	Comments
<p>Sector (P-12 schools)</p>	<ul style="list-style-type: none"> → Public/Private common language misleading due to distinct categories of private schools. → If report compares categories of schools, further delineation needed (e.g., magnet schools, charter schools, technical schools) with harmonized definitions. → Other categories of schools may be listed if important to the report but common language is usually sufficient, e.g., Department of Defense schools, Indian Education schools.
<p>Grade Levels: P - 12</p>	<ul style="list-style-type: none"> → Common language (elementary, middle school, junior high, high school) unlikely to suffice for any report. → Harmonized definitions cannot work across reports with different grade ranges. → Reports must be clear exactly what grades are included/excluded in any report in which grades (or ages) characterize a school, or portion of a school. Often used for policy decisions.
<p>School Size</p>	<ul style="list-style-type: none"> → Common language suffices (i.e., total enrollment) for most reports. → Reports that focus on school size should specify range of grades in school, and distinguish between school enrollment and enrollment per grade.
<p>Racial/Ethnic Composition <i>(see also Student Characteristics)</i></p>	<ul style="list-style-type: none"> → <i>See new federal classifications.</i> → <i>Note that students of Hispanic origin are the most predominant minority in the U.S., but have recognized differences depending on the family's country of origin, geographic region in the US, and generations living in the US. Not all reports need to consider all of these but the variation should be addressed.</i>
<p>Racial/Ethnic Disparities or Disproportion</p>	<ul style="list-style-type: none"> → <i>Needs consideration and alternative definitions; Harmonization to a single definition is unlikely.</i> → Must define and illustrate the measure(s) being used.
<p>Socioeconomic Composition</p>	<ul style="list-style-type: none"> → <i>Needs harmonization to several common definitions.</i> → <i>See also definitions used by other federal agencies.</i>

CLEAR DATA DESCRIPTIONS IN NCES REPORTS

<i>(see also Student Characteristics)</i>	<ul style="list-style-type: none"> → Family income and/or US government poverty standards being used more commonly over time (including percent of students eligible for free/subsidized lunch a/c Dept. of Agriculture). → NCES “traditional” definition has been a composite of family income, education, occupational status a/c Department of Labor, averaged to the school level. → Other approaches include proxy variables, for example, family size, family structure, number of items in the home.
<p>Dropout Rate</p> <p><i>(see also Student Characteristics)</i></p>	<ul style="list-style-type: none"> → <i>Needs consideration and alternative definitions.</i> → Cannot be harmonized to a single definition. Multiple definitions give different results (see <i>How are Dropout Rates Measured</i>. http://www.ncset.org/publications/essentialtools/dropout/part1.2.asp) → The ‘target’ dropout rate is the percentage of students who enter school in 9th grade but leave school without graduating in grade 12 (not including transfers, early graduation, late graduation due to illness, etc.). Referred to as the <i>cohort dropout rate</i> but very difficult to obtain. → The <i>event dropout rate (or incidence rate)</i> is the proportion who drop out in a single year, usually grade 12. → The <i>status dropout rate (or prevalence rate)</i> is the proportion of persons in a particular age bracket (ages 16-24 used commonly, from the US Bureau of Census) who have not acquired a high school diploma. Used most often to report national results and to dictate state and federal policy, this measure suffers from serious methodological problems. → Reports use different numerators and different denominators even in calculating a single type of rate. Recommend stating clearly how a rate is determined and what its limitations are.
Graduation Rate	<ul style="list-style-type: none"> → <i>Needs consideration and possibly alternative definitions.</i> <p>See NCSET report above. Note that graduation rates do not translate to dropout rates or vice versa.</p>

Classroom Characteristics	Comments
Class Size	<ul style="list-style-type: none"> → Common technical definition is fine: The number of students in the room for whom the teacher is responsible; (also the number of students on a teacher’s roster). → Often used erroneously for student-teacher ratio, leading to errors in policy making.
Student-Teacher or Pupil-Teacher Ratio (PTR)	<ul style="list-style-type: none"> → The number of students in a unit (e.g., school, district, state, or country) divided by the number of adults in the same unit. → Adults (“teachers”) needs harmonization because they are counted differently for different reports (e.g., teachers, administrators who may or may not have a teaching role; remedial teachers; special education teachers; Chapter I teachers; part-time teachers; teacher aides; librarian; others). → Not to be confused with class size.

CLEAR DATA DESCRIPTIONS IN NCES REPORTS

Classroom Organization	<ul style="list-style-type: none"> → Common language sufficient. → <i>Give detailed descriptions with specific terms, e.g., Open Classroom, Montessori Classroom, Flipped Classroom, Inclusion Classroom, Blended Classroom.</i>
Cooperative Learning Groups	<ul style="list-style-type: none"> → <i>Needs harmonization.</i>

Student Characteristics	Comments
Sex or Gender	<p>NOTE: Rapid change in usage make it likely that federal standard language directives or best practices will supersede what is written below.</p> <ul style="list-style-type: none"> → If options to a survey question are limited to Male or Female, recommend using Sex. → If a survey delves into orientation or self-described identity, use Gender. Caveat: Some respondents may not distinguish between the terms and may elect to report either biological sex or gender identity.
Race/Ethnicity	<ul style="list-style-type: none"> → <i>See new federal classifications.</i> → Depends on the focus of the survey and the needs of the consumer. → Avoid groupings comprised of diverse populations if possible (e.g., American Indian/Alaskan Native; Asian/Pacific Islander; Hispanic origin).
Educational Attainment <i>(also: Expected Education; Educational Aspirations)</i>	<ul style="list-style-type: none"> → Common language often suffices (less than high school; high school completion or GED; some college but no certificate or degree; complete a 2-year postsecondary program; complete a 4-year postsecondary program; master’s degree; professional degree or Ph.D.). → <i>Harmonization recommended for reports presenting fewer categories.</i> (Note some debate about high-school equivalency programs.) → If focus is on the process of advancing through stages of education, more specific classifications recommended. See, for example, <i>International Standard Classification of Education</i>, UNESCO, 2011. → “Expectations” and “aspirations” are distinct; common language suffices.
Type of High School Program	<ul style="list-style-type: none"> → Common language is sufficient (e.g., general program; college preparatory program; vocational program; specialized program). → Additional detail needed if report is an analysis or comparison of specific programs and/or their requirements.
Type of Postsecondary Institution	<ul style="list-style-type: none"> → Common terminology suffices for reports about general patterns (Public/Private and 2-year/4-year institutions). See <i>Statistics in Brief: What is the Price of College?</i> (NCES 2011-175). → <i>Harmonizing is needed</i> for more detailed classifications of type of institution, including those preparing students for specific jobs. → <i>Harmonizing is needed</i> for detailed classifications of types of vocational programs. See <i>Statistics in Brief: Participation in High School Career and Technical Education and Postsecondary Enrollment</i> (NCES 2018-043).

CLEAR DATA DESCRIPTIONS IN NCES REPORTS

Outcomes of Schooling	Comments
<i>Collaborative Problem Solving (PISA)</i>	→ Words not common terminology within or across countries; need to be described in common language.
<i>Creative Problem Solving (PISA)</i>	
<i>Proficiency or Academic Proficiency</i>	→ <i>Harmonization is needed</i> to clarify distinctions among proficiency, achievement, aptitude, skill, ability.

Appendix B: Agenda

NATIONAL INSTITUTE OF STATISTICAL SCIENCES

NCES EXPERT PANEL ON CLEAR DATA DESCRIPTIONS IN NCES REPORTS

February 5-6, 2018
PCP Building, Room 5170/6082

AGENDA

Monday, February 5th

9:00am-9:30am	Arrive at PCP and proceed through security
9:30am-11:30am	Welcome, Introductions, Clarification & Discussion of issues & panel goals with NCES
11:30am-12:00pm	Executive Session
12:00pm-1:00pm	Lunch
1:00pm-2:30pm	Executive Session
2:30pm-3:00pm	Break
3:00pm-3:30pm	Executive Session
3:30pm-4:00pm	Clarifications with NCES
4:00pm	Adjourn

Tuesday, February 6th

9:00am-9:30am	Arrive at PCP and proceed through security
9:30am-11:30am	Executive Session
11:30am-12:00pm	Further Clarifications with NCES
12:00pm-1:00pm	Lunch
1:00pm-2:00pm	Executive Session
2:00pm-3:00pm	Feedback to NCES
3:00pm	Adjourn

Appendix C: Expert Panel Members' Biosketches

Jim Albert, Ph.D.

Title: Professor, Bowling Green State University

Dr. Jim Albert is Professor in the Department of Mathematics and Statistics at Bowling Green State University. Dr. Albert's expertise is in Bayesian modeling, statistics education, and the application of statistical thinking in sports. Currently, he is the PI of a NSF program to develop a new data science undergraduate major at BGSU. He has developed statistics courses for prospective middle school and high school teachers at Bowling Green, and has given workshops for secondary education teachers in Northwest Ohio. He has participated in the Section on Bayesian Statistics, the Section in Statistics Education, and the Section in Statistics in Sports of the American Statistical Association. He is past editor of the *Journal of Quantitative Analysis of Sports* and *The American Statistician* and has served as Associate Editor of the *Journal of the American Statistics Association*, *Communications in Statistics*, *Journal of Statistics Education* and *Chance* magazine. He has published over a hundred papers in Bayesian modeling and computation, statistics education and the application of statistical thinking in baseball and tennis. In addition, he has written or edited a number of books on Bayesian modeling, introductory statistics, and the interface of baseball and statistics.

Jeremy Finn, Ph.D.

Title: SUNY Distinguished Professor and Chair of the Department of Counseling, School, and Educational Psychology, Graduate School of Education, University at Buffalo, SUNY

Dr. Jeremy Finn is an internationally known scholar, especially due to his extensive work on quantitative research methods and issues of class size, student engagement, school security measures, and dropping out. He has taught at the Ontario Institute for Studies in Education and Stanford University, and has held research fellowships at the National Research Council, Educational Testing Service, and the International Association for the Evaluation of Educational Achievement. He also previously served as an ASA-NSF Fellow at NCES.

Fred Galloway, Ed.D.

Title: Professor, University of San Diego

Dr. Fred Galloway is currently a Professor and 2012-13 University Professor in the School of Leadership and Education Sciences at the University of San Diego; he also serves as a senior research associate at both the Center for Education Policy and Law and the Caster Family Center for Nonprofit and Philanthropic Research at the university. He received his bachelor's and master's degrees in economics from the University of California, San Diego, and master's and doctoral degrees from the Harvard Graduate School of Education. Dr. Galloway's research interests include the economics of education, higher education policy, and research design and methodology, and he has published more than 60 journal articles, policy reports, and book chapters – including recent articles in *Education Administration Quarterly*, the *Journal of Business Administration Research*, the *Journal of Research in Leadership Education*, *Asia Pacific Education Review*, and the *Handbook of Research on Online Instruments, Data Collection and Electronic Measurements: Organizational Advancements*.

He has also provided methodological guidance and direction on 102 completed dissertations (54 as chairperson and 48 as associate member), and has served on numerous technical review and advisory panels, including five technical review panels for the National Center for Education Statistics, the Congressional Budget Office Student Loan Advisory Panel, the General Accounting Office Student Loan Advisory Panel, the Lumina Foundation Technical Advisory Panel, and the Technical Review Group for the U.S. Department of Education Follow-Up Evaluation of the GEAR UP Program. In 1999 Dr. Galloway also

testified before the US House of Representatives as a friendly witness regarding the results of the national Direct Loan evaluation. In addition to his love of research, he is a passionate teacher who has been recognized with several faculty of the year awards by undergraduates as well as graduate students in both economics and education.

Heather Ridolfo, Ph.D.

Title: Agricultural Statistician, Research and Development Division, USDA National Agricultural Statistics Service

Dr. Heather Ridolfo is a survey methodologist in the Research and Development Division of the USDA's National Agricultural Statistics Service. Previously she was a survey methodologist at the National Center for Health Statistics. Her areas of research include questionnaire design, measurement error, respondent burden, and respondent-interviewer interactions. She has a PhD in Sociology from the University of Maryland, College Park, where she focused on self-concept development. She co-authored a book based on her dissertation research titled "Mobility Impairment and the Construction of Identity." She continues to apply her expertise in identity and self-presentation to understanding how respondents interact with survey questionnaires and the impact of those interactions on data quality.

Nora Cate Schaeffer, Ph.D.

Title: Sewell Bascom Professor of Sociology, University of Wisconsin, Madison

Nora Cate Schaeffer is Sewell Bascom Professor of Sociology at the University of Wisconsin, Madison, where she also serves as Faculty Director of the University of Wisconsin Survey Center, teaches courses in survey research methods, and conducts research on questionnaire design and interaction during survey interviews. Her recent service for the National Research Council (NRC) includes Panel to Evaluate the NCSES Approach to Measuring the Science and Engineering Workforce, the Panel on Measuring Rape and Sexual Assault (doi:10.17226/18605), Panel on the Future of Social Science Data Collection (doi:[10.17226/18293](https://doi.org/10.17226/18293)), and the Standing Committee on Integrating New Behavioral Health Measures into SAMHSA's Data Collection Programs (doi:10.17226/21920). She recently served on the Public Opinion Quarterly Advisory Board of the American Association for Public Opinion Research, the General Social Survey Board of Overseers, and the Panel Study of Income Dynamics Board of Overseers. She has also served on the editorial boards of *Public Opinion Quarterly*, *Sociological Methods and Research*, *Journal of Survey Statistics and Methodology*, and *Sociological Methodology*. In 2010 she was selected as a Fellow of the American Statistical Association, and in 2015 she was selected as Fellow of the Midwest Association for Public Opinion Research.

Gina Walejko, Ph.D.

Title: Survey Statistician; U.S Census Bureau

Dr. Gina Walejko has been a researcher and survey methodologist at the U.S. Census Bureau since 2013. She has worked in both the Research & Methodology and Decennial Directorates. Her research has focused on adaptive survey design, interviewer compliance, reducing decennial undercounts as well as audience segmentation and message design. She has been elected and served as both program chair and president for the Washington-Baltimore Chapter of the American Association for Public Opinion Research as well as program chair for the Government Statistics Section of the American Statistical Association. She is coauthor of several technical papers, and her refereed journal articles have appeared in publications including *Science*. Her areas of technical expertise include questionnaire design, field experiments, and message testing.

Panel convened by National Institute of Statistical Sciences

Nell Sedransk, Ph.D.

Title: Director-DC, National Institute of Statistical Sciences

Dr. Nell Sedransk is the Director of the National Institute of Statistical Sciences and Professor of Statistics at North Carolina State University. 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.