

	The National Assessment of Educational Progress	The NAEP Design and Analysis Committee
---	--	---

To: Peggy Carr
National Center for Education Statistics

From: Brian Junker and Betsy Becker
Co-Chairs, NAEP Design and Analysis Committee

Date: November 12, 2020

Subject: Technical, Practical, and Ethical Infeasibility of NAEP 2021 Administration
and Reporting

Executive Summary

Due to the impact of the COVID-19 pandemic on sample size, sample representativeness, practicality of testing, and the health of students, teachers, and test administrators, the NAEP Design and Analysis Committee concludes that *it is infeasible to “conduct a national assessment and collect and report assessment data, including achievement data trends...in a valid and reliable manner,” as required by U.S. Code 9622*. Therefore, the DAC recommends in the strongest possible terms that Congress grant a waiver allowing the assessment to be called off for 2021, and to resume in 2022 or 2023, assuming that by then the pandemic will have abated.

Background

The most optimistic, best-case survey outcome for NAEP in 2021 would produce results that do not meet the validity and reliability standards to which technicians and policy-makers justifiably hold NAEP. K–12 students will continue to experience a variety of instructional modes in 2021—face-to-face, hybrid face-to-face/online-at-home, and entirely online-at-home—depending on the ever-changing severity of the COVID-19 crisis in their area. Because of this, NAEP would need multiple visits to schools to test students in the first two modes (face-to-face and hybrid), requiring additional test administrators, personal protective equipment, and procedures to limit the spread of the virus. Students in the third mode (online-at-home) cannot be tested at all without prohibitive costs and challenges.

Because no additional funds are available for these added expenses, NCES has proposed a best-case reduced survey design that makes it possible to pay for them out of existing funds. This design targets less than half of the normal NAEP sample size ($\approx 200,000$ students in

≈5,000 schools, down from ≈473,000 students in ≈12,000 schools), and entirely omits students in online-at-home instruction, to control costs. If successful, this design would produce reports with *reduced validity*, because students in either under-performing or high-achieving districts (or both) may be restricted to online-at-home instruction in proportions that are unpredictable and changing over time as the COVID-19 crisis evolves, and *reduced reliability* because of the reduced sample size.

This best-case survey design is *extremely unlikely to be successful in the field*, however, producing a sample dramatically smaller than intended. The data from this much smaller sample would be unusable to report national or state achievement levels or trends. Information on instructional modes currently being gathered by NCES shows that:

- (1) some states may be entirely omitted from the NAEP sample, and others are likely to contribute highly biased samples because such a large proportion—65%, 85% or more—of their students are in online-at-home instruction;
- (2) for ethical and liability reasons NCES does not expect the data collection contractor, Westat, to send assessment personnel to assess in schools located in counties with an elevated level of COVID-19 infection; and
- (3) many schools which would otherwise allow student testing limit COVID-19 contagion by preventing non-essential workers such as NAEP assessment personnel from visiting.

What is left is a small handful of states for which the instructional modes and visitor restrictions *are not yet known*, of which only two would potentially contribute enough students—1,200 per grade—to be included in state-level reporting. Moreover, based on current information, Westat estimates that less than half of the national sample needed for the required statistical analysis of test items for each subject at each grade level (14,000 students) would be available. Sufficient sample sizes to accurately estimate achievement levels, especially for the smaller reporting groups, are also unlikely to be available. As a result, *NAEP would be unable to report either achievement levels or trends*.

Serious ethical, liability, and reputational issues are also involved in a potential 2021 assessment. Even in schools where NAEP is allowed to operate, mixing NAEP personnel and testing materials with students and staff, potentially in close quarters, puts participants in a heightened risk category, according to Centers for Disease Control guidelines¹. Data collection under these circumstances may also violate the professional ethics codes to which most NAEP

¹ Centers for Disease Control K-12 Readiness and Planning Tool
(<https://www.cdc.gov/coronavirus/2019-ncov/downloads/community/School-Admin-K12-readiness-and-planning-tool.pdf>)

personnel are bound²³⁴. Similarly, the Common Rule⁵ for ethical research to which the Department of Education is a signatory makes clear that approved research must ensure that “risks to subjects are minimized,” that “selection of subjects is equitable,” and that “additional safeguards have been included in the study to protect the rights and welfare [of vulnerable populations]” such as children. In addition, should a COVID-19 outbreak be associated with NAEP assessment efforts, the Department of Education, including the National Assessment Governing Board, National Center for Education Statistics, and NAEP contractors, would be exposed to lawsuits and liability. Finally, NAEP’s reputation as a beneficent and authoritative provider of educational data could be damaged if its data collection threatened the vulnerable population of minors on which it depends, for results that are virtually guaranteed to be of unusably low quality.

Conclusion

The DAC fully endorses the role of NAEP as “the Nation’s report card,” and understands the desire to use NAEP to learn what has been lost in educational progress due to the COVID-19 crisis. Indeed, the DAC has grappled with how NAEP 2021 data could be analyzed to answer this question, if it were collected. However, after reviewing plans for NAEP 2021 and considering the sample size, sample bias, ethical, and other issues raised above, we conclude that it would be a waste of valuable taxpayer dollars to pursue data collection for NAEP in 2021. As the COVID-19 crisis abates, we anticipate that a complete, reliable, and valid assessment that meets the highest ethical and data-quality standards will be possible at a later date. This future assessment could be used to gauge what has been lost in the intervening years. We recommend in the strongest possible terms that Congress grant a waiver to allow the assessment to be called off for 2021, and to resume in 2022 or 2023 as conditions permit.

² American Educational Research Association Code of Ethics
([https://www.aera.net/Portals/38/docs/About_AERA/CodeOfEthics\(1\).pdf](https://www.aera.net/Portals/38/docs/About_AERA/CodeOfEthics(1).pdf))

³ American Association for Public Opinion Research Ethics Standards
(<https://www.aapor.org/Standards-Ethics/AAPOR-Code-of-Ethics.aspx>)

⁴ American Statistical Association Ethical Guidelines
(<https://www.amstat.org/ASA/Your-Career/Ethical-Guidelines-for-Statistical-Practice.aspx>)

⁵ Common Rule, section 97.111
(<https://www2.ed.gov/policy/fund/reg/humansub/part97-2.html#97.111>)

About the NAEP DAC

The Design and Analysis Committee (DAC) is a committee of 14 independent experts in statistics, educational testing, psychometrics and cognitive psychology, administered by Educational Testing Service (ETS) under Contract 91990019C0045, to provide advice to ETS and other NAEP contractors on all technical and related aspects of design, analysis and reporting of NAEP data. The DAC, which has operated in its current form since 1987 (and in other forms going back to the inception of NAEP in 1964) has played an integral role in setting and helping to achieve the high standards of quality and innovation that make NAEP the leader in educational achievement survey methodology and reporting.

The following are the members of the NAEP DAC:

Betsy Becker, Florida State University (Co-Chair)

Brian Junker, Carnegie Mellon University (Co-Chair)

Peter Behuniak, Criterion Consulting, LLC

Daniel Bolt, University of Wisconsin, Madison

Lloyd Bond; Professor (Emeritus), University of North Carolina, Greensboro;
Senior Scholar (retired), The Carnegie Foundation

Derek Briggs, University of Colorado, Boulder

Richard Duran, University of California, Santa Barbara

Ben Hansen, University of Michigan

David Kaplan, University of Wisconsin, Madison

Ken Koedinger, Carnegie Mellon University

Sophia Rabe-Hesketh, University of California, Berkeley

Michael Rodriguez, University of Minnesota, Minneapolis

S. Lynne Stokes, Southern Methodist University

Chun Wang, University of Washington