



**THE NATIONAL CENTER FOR THE IMPROVEMENT OF EDUCATIONAL
ASSESSMENT¹**

2019 Summer Internship Program in Educational Assessment and Accountability

Detailed Project Descriptions

1. Effective Practices for Subscore Reporting and Use

A persistent tension between test developers and test users is the reporting and use of subscores on large-scale achievement tests. Developers are typically reluctant to support reporting subscores, due to lack of precision and the risks of inappropriate interpretation. However, many test users support subscores, citing their value for diagnostic purposes.

In recent years, a range of practices have been developed for reporting subscores. These include reporting: scaled scores, raw scores, percent correct, and/or performance categories. Some reports focus on relative performance, while others focus on criterion referenced reporting.

The purpose of this project is to evaluate and classify the range of practices with respect to subscore reporting on large-scale state achievement tests. This information will help inform development of guidelines to describe and promote effective practices.

2. Interim Assessment Research Synthesis

In the past 20 years research on interim assessments has exploded. There have been numerous studies dedicated to examining operational programs of interim assessments and their effects on student learning. To date, however, there has not been a systematic review of this extant body of literature. The purpose of this internship is conduct a research synthesis, and if possible a meta-analysis, on all available literature on interim assessments. Doing so will involve drawing on a theory of action framing to characterize how, exactly, interim assessment results are used within these studies. Thus this work is meant to define the specific ways in which programs of interim assessments have been used and what, if any, uses have empirical support on their efficacy. This project is therefore technical, in that the designs of interim assessments will be characterized, but also policy focused, in that the use of the interim assessments within specific policy contexts will also be investigated.

3. Tools for Depicting and Analyzing Achievement Gaps

Reducing achievement gaps has been a primary purpose of federal Elementary and Secondary Education Act (1965, reauthorized as IASA, NCLB, and ESSA). Yet progress has been

¹ To learn more about the Center, please visit www.nciea.org.

agonizingly limited by most accounts, with a few bright exceptions. This project has three main outcomes that will be achieved by working closely with the mentor: 1) Summarize the most interesting and useful definitions of “achievement gap” and how to measure them—and show how the various conceptualizations are related to larger policy concerns and to each other; 2) Identify and/or devise a few powerful depictions and analyses of “(reducing) achievement gap” that would apply to states’ current efforts; and 3) program some in open-source code (e.g., R, D3.js) or widely used commercial software as a demonstration of a toolkit that might be adopted by states and others interested in understanding, communicating to others, and acting to improve educational outcomes in terms of equity. * Highly qualified applicants will have previous experience working with analyses of achievement gaps with the flexibility to work with multiple definitions; a honed eye to know how to portray this technical topic to policymakers and other non-technical audiences; an inventive ability to improve on current approaches; facility with a suitable open-source programming language to produce the demonstration toolkit. Excellent oral English communication required.

4. Evaluating ESSA’s Impact on Equity

The Every Student Succeeds Act (ESSA) requires states to develop school accountability systems to signal what outcomes are valued and provide information about school performance with respect to those outcomes. States have some discretion when it comes to how these systems are designed and implemented. Given that a central purpose of ESSA is to promote equity – improving outcomes for disadvantaged students – it stands to reason that evaluating the effectiveness of these systems with respect to equity is critical.

This project will involve a review of state accountability systems with a focus on the design features (e.g. achievement gap reduction, growth for students below proficient) explicitly tied to equity. We also hope to review available state and national data to better understand progress toward equity goals. Ultimately, we hope to better understand what accountability features seem most important and effective to promote equity outcomes.

5. **Analysis of State ESSA Accountability Systems:** Every state, as required by federal ESSA law, has implemented a state school accountability system. A key characteristic of that system is the state must identify schools for various levels of “support and improvement” referred to as CSI, TSI, and ATSI. The rates of identification vary somewhat between states: for example, ATSI varies from less than 2% of all schools in one state to more than 30% in another. This project entails analyzing the “construct” and measurement reasons for the variations, and implications, especially reliability/precision and cross-state comparability of the determinations. * Highly qualified applicants will have understanding of current ESSA accountability requirements; detailed acquaintance with at least one state’s school accountability system; program evaluation ability to analyze correspondence between policy, accountability “business rules,” and outcomes; and ability to analyze and model school rating systems using SAS, R, SPSS, or a similar programmable analysis software program. Ability to write technical documentation in English quickly, and strong ability to work independently and

with an intense team required. The Center will provide at least one state data set for simulation and analysis.

6. **Assessment Literacy for Policy Makers:** Much of the focus on assessment literacy has been on improving the knowledge and skills of educators to better design and use assessments to improve instruction for students. We have come to recognize that a lack of assessment literacy among state policy makers can lead to considerable instability, weak designs, and inappropriate uses of state assessments. Further, given the rapid turnover of state leaders, we need to create long-term structural supports for improving the assessment literacy of these state leaders. We are just beginning to understand how to meet the assessment literacy needs of state policy leaders. Linn and Herman (1997) tried to address some of the standards and assessment literacy needs of state leaders with their very clear and concise *A Policymaker's Guide to Standards-Led Assessment*, but it was only one shot. Assessment literacy is one of many other competing demands (e.g., budgeting, politics, and communication) for state policy maker's attention. We need a better understanding of what it means to improve the assessment literacy of state policy leaders—what they need to know and understand—and how best to accomplish this. Further, we should identify approaches for state assessment leaders to better communicate the most critical assessment issues to their chief state school officers. The intern selected for this project will build off of the Center's conceptualization of assessment literacy to design tools and modules for improving policy maker's understanding of key assessment topics and then test the efficacy of such tools with current state policy makers.
7. **Large-Scale Assessment Policy Analysis:** Given the powerful influence of state assessment laws and regulations on the design and implementation of large-scale assessment systems, it is surprising how little we, as a measurement community, systematically know about these laws and regulations. It is difficult to propose policies to support improved large-scale assessments such as through the use of matrix-sampling designs without knowing the current policy and political contexts. Further, if we want to make “room” for balanced systems of assessment, we need to propose policies for reducing the footprint (i.e., the influence of the state assessment on the rest of the system) of end-of-year summative tests without reducing assessment quality, in support of balanced-system implementation. Additionally, stability is central to any policy instrument such as a large-scale assessment or accountability program, and we have observed the negative consequences of instability in large-scale assessment policies. We know many states that have had three or more state testing programs over only five or six years. Not surprisingly, local educators often respond to such instability by simply hunkering down and doing nothing (“This, too, shall pass.”). There are many reasons for these frequent changes, but most are political. In addition to enhancing assessment literacy (which entails an understanding of the need for stability in this regard), we propose working with policy experts to develop guidance for policymakers that ensures the stability of large-scale assessment systems. Therefore, the intern selected for this project will analyze a sample of state assessment policies to understand the range of requirements encoded in statute or regulation. This landscape analysis will provide the foundation for generating policy frameworks that can be used as



models for state leaders who want to improve the constraints and requirements associated with large-scale testing programs.