

Comments on Presentations at the NRC Symposium on the Use of School-Level Data for Evaluating Federal Education Programs

Mitchell D. Chester
Associate Superintendent for Policy and Accountability
Ohio Department of Education

December 14, 2005

This paper summarizes my comments during *Session 5: State and District Reactions* of the December 8-9, 2005, Symposium on the Use of School-Level Data for Evaluating Federal Education Programs. My comments relate to the questions that are salient to policy makers when gauging program effectiveness; the importance of taking policy context into account; and the potential for incorporating longitudinal, student-level data.

Key Research Questions

I would like to raise three sets of considerations related to research questions – the value of conditioning findings, employing achievement targets as the counterfactual, and identifying salient policy attributes. I believe that each of these considerations is critical to the use of school-level data to evaluate federal, as well as other, education programs.

Conditioned findings – For many policy makers the critical evaluation questions relate to the contexts and conditions under which a given treatment is likely to be effective. Most interventions will have varying impacts and understanding what features differentiate schools where a given intervention is effective from those where it is not is key. It is unlikely that any particular intervention will be universally effective or ineffective.

Two examples illustrate this consideration. First, during the initiation of a statewide charter school program the salient evaluation question may be whether charter schools are effective. Once charter schools are part of a state's education landscape, the question of interest may shift to refining charter school policy based on an understanding of the differences in the contexts and conditions that differentiate charter schools that are effective from those that are not. Second, as states consider interventions with low performing schools, evaluations that provide insights into the contexts and conditions under which a given treatment is likely to be effective will be much more valuable than evaluations that do not consider this. Again, no intervention is likely to be universally effective or universally ineffective, although some are likely to be useful in certain situations.

* Copyright © 2005 National Academy of Sciences. All rights reserved. No part of these pages, either text or image may be used for any purpose other than personal use. Reproduction, modification, storage in a retrieval system or retransmission, in any form or by any means, electronic, mechanical or otherwise, for reasons other than personal use, is strictly prohibited without prior written permission.

Opinions and statements included in the paper are solely those of the individual author(s), and are not necessarily adopted or endorsed or verified as accurate by the Planning Committee for the Symposium on the Use of School-Level Data in Evaluating Federal Education Programs or the National Academy of Sciences, including the National Academy of Engineering, Institute of Medicine, or National Research Council.

Achievement targets as the counterfactual – Much of the discussion about research design focused on the identification of appropriate control groups, including the potential for random assignment. Several presenters and respondents mentioned the importance of matching control groups on as many variables as possible. While the ability to draw conclusions by comparing outcomes in treatment schools to outcomes in control schools was the focus of most discussion, the effectiveness of treatments in moving schools to achievement targets is of substantial importance to many policy makers and educators. When benchmarking the impact of an intervention against an achievement target the counterfactual is failure to meet the target, not simply whether the treatment schools outperformed control schools. Intervention schools may outperform control schools without moving schools to their achievement targets.

The following are examples of research questions that rely on failure to meet an achievement target as the counterfactual. Is the implementation of a given intervention likely to result in a school meeting Adequate Yearly Progress (AYP) or achieving a higher rating in a state accountability system? Under what circumstances is a treatment likely to result in a school meeting AYP or achieving a higher rating?

Salient policy attributes – Understanding the district and state policy attributes that contribute to intervention effectiveness would be an important contribution of research utilizing school-level data. In his comments, David Francis alluded to this importance when he identified the “mechanism of action” as an important point of interest for researchers and evaluators. Variation in district and state theories *in action* likely contributes to variation in treatment outcomes. If states and districts are to scale up reform, then understanding the intersection of policy and implementation is essential.

Interaction of Treatment and Context

As alluded to earlier, greater insight into program effects will be realized if the impact of context is considered. It is likely that no treatment is universally successful or universally unsuccessful. The context (e.g., school and district resources, capacity, leadership; state policy attributes) within which a school is situated is one determinant of the impact of a given intervention. By understanding the interaction of treatment and context, districts and states will be better able to match schools and interventions in a manner that increases the likelihood that the treatment will be effective.

From the perspective of cross-state or national research, variation in state policy is a key context variable. Evaluations of federal programs that do not consider differences in state policy approaches may miss factors that are critical to understanding the between-state variation in the programs’ impacts. For example, national studies designed to examine the impact of NCLB-mandated school choice options would want to account for school choice opportunities that existed prior to NCLB (e.g., state-mandated choice, charter schools, open-enrollment; district-sponsored magnet programs, open-enrollment). By accounting for state and district choice opportunities, the national study would have the opportunity to distinguish between effects (or lack of effects) attributable to NCLB-mandated school choice and effects attributable to non-NCLB-mandated school choice.

Longitudinal, Student-Level Data

There was much support during the symposium for incorporating student-level data in federal school-level databases. I, too, support the recommendation that the United States Department of Education explore the addition of student level data. In light of the growing number of states that are developing longitudinal, student-level data systems, there is an opportunity for the USDE to explore the potential for such a data system. The growing number of states developing such data systems has several impetuses. The NCLB requirement of annual testing in grades three through eight provides states with yearly test results in reading and mathematics for individual students as they move through the grades. As longitudinal student data becomes available, more states are adopting a unique student identifier system to track and store the individual student data. Additionally, the USDE recently awarded grants to 14 states to support the development of longitudinal data systems.

There are at least two basic ways to think about a cross-state database of longitudinal, student-level data. One is as a stand-alone, federally maintained database. The second is a more fluid concept whereby states that have longitudinal student records would make their data available for cross-state research purposes wherein the student records are merged with the federal school-level database.

For the concept of a cross-state, student-level database to advance, there are a number of roadblocks that will need to be resolved. These roadblocks include procedures for securing the confidentiality of individual students, state statutory restrictions on the use of individual level data, and state education agency capacity to collect and store the data.