

District- and School-Level Reaction Panel – Panelist Remarks

Joe O'Reilly
Mesa Public Schools, Mesa, Arizona
December 9, 2005

I was asked to provide comments on what districts think about the school level database. That's easy. Most of the 14,000 or so districts are completely oblivious to this work. And those in the larger districts are probably, at best, vaguely aware but too busy to give it any extended focus.

In fact, when someone in a district hears that the federal government is interested in collecting data, they probably have an initial reaction like mine. When I heard the title of the session I immediately flashed on a Gary Larson cartoon of two deer in the woods. One, lets say he represents someone from AIR, Westat or a state NAEP Coordinator, puts his arm around the second deer, who represents a principal or LEA staffer, looks at the white and red and white and red concentric circles on his chest, and says 'Bummer of a Birthmark, Bob!'

The first reaction is that we are the ones with a target on our chests – either demands on student and staff time (that is what you call collecting data) or fear of results that reflect poorly on the school.

But we generally do cooperate. Why? Because we have a vested interest in getting information that will help students learn. If this database could be combined with other information to produce information on how different interventions work in different schools and different populations, it would be very useful to schools.

This is how I see schools primarily using the database information – by using the studies generated by researchers to make better informed decisions. Districts may also use the database to conduct studies, but for the most part we will be consumers of the information that come from the database. So, the improvements I would like to see are those that enable us to answer specific questions – how does a specific program work for a specific population such as ELL students? Or, what have schools comparable to mine found to be effective to teach a given topic? In other words, improvements that directly add to the schools' bottom line – student learning.

It would also be welcome if it saved time. David Goodwin mentioned that data could come from state assessments or NAEP or independent tests. If researchers could increasingly use data that are already collected instead of taking time for more surveys and assessments, it would be a great improvement. And, since state assessments are the ones we pay attention to, they are the measures that are most likely to show the impact of an intervention.

Having a comprehensive database would also be of value to researchers because schools are becoming increasingly resistant to allowing researchers in because of the focus on time on task to meet accountability demands. This is an unintended consequence of NCLB for the research field. As schools become increasingly hesitant, databases such as the one we are discussing today become much more important.

Let me address some of the common issues that came up during this session. First, data quality is an issue that a number of speakers brought up. There is a concern that the data is bad and inconsistent across states. There are data problems, but I have

seen a great deal of improvement over the past few years. States are becoming more consistent, which bodes well for this effort:

- Ø States are testing more grades and the grades tested are consistent across states.
- Ø States are collecting, nominally at least, the same demographic information – ethnicity, ELL status, Special Ed status, etc. It is not perfect but through the efforts such as the SIF initiative and the NGA’s current effort to have a consistent definition of a dropout, it is much more consistent than it has ever been.
- Ø The USDOE’s peer review of assessment and accountability systems has detailed guidance and rubrics that are followed by every state. Perhaps equally importantly, SEA staff have been trained as peer reviewers, making their efforts to develop their submissions more consistent with what the Department of Education wants and more consistent with other states. As an aside, the documentation submitted by the states is a rich trove of information for researchers who want to understand, use and compare data from state systems.
- Ø All states have to have a way to report whether schools have tested 95% of enrollment, how those students performed, and to which subgroup(s) each student belongs. Presumably each state has some sort of database by student to accomplish this.

A second issue is the level of detail that should be collected. As the school level database is developed, the potential that this source of detailed information offers should not be ignored. If this database is to be truly valuable to schools and districts, student

level data must be collected. States have this data, and use it to report percent tested and scores by group. The value of this data to practitioners is that we can ask 'what is the impact of intervention A on students with characteristics X, Y and Z?' For example, which of three reading interventions work best with English Language Learners in urban schools? This is only directly answerable if you know who was an ELL student. Losing access to this detailed data would greatly limit the usability of the database.

This highlights for me a disconnect we heard yesterday between the direction that USDOE is going and what is becoming increasingly available. The Department said that it would collect percent mastery by school level and not to expect much more. States have rich databases by student down to the concept level. It is baffling to me why, when so much data is being collected by student down to the concept level, such a rich resource is being lost by focusing just on the percent at mastery by school. That student data is needed to address important questions about subgroups, so I would hate to see that lost.

A third theme throughout the presentations is the need for context information. I agree that context information is very important. Let me give you a non-testing example. A large district in the southwest had won an award for how healthy yet tasty the food was. So one day as the researcher was loading his plate up on Mexican food day he said to the cafeteria worker 'I can't believe this food is healthy, it tastes so good.' The worker looked back puzzled and said 'Oh, you mean the award. We don't follow those recipes. If we did, no one would eat it.' Similarly, Arizona, which is about a year ahead of many states, reset their tests and cut points. At most grades the passing rates jumped 20 to 30%. However, when I approximated the new results to the old scale it appeared that our scores

actually flat lined or declined. Without knowing the context the sharp increase in scores would be misunderstood.

The database would be would be improved if it included access to context information so that researchers can understand the data better. States have a great deal of context information – linking studies, TAC minutes, information on the scales and cut points, and the peer review submissions. All these could be made available through an electronic repository and hot links. Additional information could be collected through surveys of LEA staff. Admittedly this will be more time taken from schools and districts, but if they are parsimoniously designed and provides a tangible payoff for students and schools, it would be worth it.

Fourth, it is important that this database have the features that we are discussing, because if the department does not build a database with the needed features others will fill the gap. And they will not do it with the same level of quality and independence. For example, Just For The Kids has made a major push to collect and report student level data in our state, and many others. They are collecting the data, but it will not be as universally available for research as a database developed by the government.

Similarly, if one cannot use the database to conduct studies others will conduct them without the large samples and the same level of rigor. I saw an ad for a study conducted at one of our schools where the 18 treatment students got $\frac{3}{4}$ of one more item correct which was statistically significant but educationally irrelevant. In another case, a vendor was conducting a rigorous scientific study. A federal monitor even observed and videotaped as teachers were randomly assigned. But the control group is getting the normal reading instruction while the experimental group is getting the normal instruction

plus 45 more minutes of direct instruction. Do you think 45 more minutes of instruction per day might result in higher scores for the treatment group no matter what materials you used? Vendor studies have the illusion of gold standard scientific rigor but they are more likely, to use Michael Scriven's description, just plated gold or even just dross.

So to sum up, from a district perspective –

- Ø Don't lose the rich trove of individual information just because the law only requires a percentage at the school level. This is a research database, not a database for verification of compliance or simple reporting. Collecting individual data raises a number of methodological and policy concerns, but the value of the data outweigh those concerns. And if it can't be done everywhere, it can be done somewhere so we could at least start the effort and see how much more useful this approach would be.
- Ø Collect the abundant context information. Even if we just captured the material submitted for the accountability and assessment peer reviews, we would have a much better understanding of what the data in the database mean.
- Ø Don't let the effort to improve the database as described over the last two days drop because it is too hard. If the Department does not do it someone else will, and it will not be as useful and available.

And finally, keep in mind that what we want is information that helps us help students learn. Our questions are practical and precise, and they involve specific groups of students. If 'what would a principal or superintendent ask' was made a key

principal that is kept in mind when developing and improving this database, rather than what is easy or what is an incremental improvement over what we already have, the database can be very helpful to those of us in schools.