



UNITED STATES DEPARTMENT OF EDUCATION

OFFICE OF ELEMENTARY AND SECONDARY EDUCATION

THE ASSISTANT SECRETARY

APR 15 2007

Dr. Gerald Zahorchak  
Secretary of Education  
Pennsylvania Department of Education  
333 Market Street  
Harrisburg, Pennsylvania 17126

Dear Secretary Zahorchak:

Thank you for submitting a proposal for the U.S. Department of Education's (Department) growth-based accountability model pilot project. I appreciate the work you and your staff have done to participate in this effort so far. The Department continues to believe that this pilot project can help determine whether growth models will provide a fair, reliable, and innovative mechanism for holding schools accountable for ensuring that all students reach grade-level proficiency in reading and mathematics by 2013–14.

As you know, the panel of peer experts reviewed Pennsylvania's growth model on March 15-16, 2007. During this review, the peers raised a number of substantive concerns with the structure of Pennsylvania's model, indicating the model was not acceptable for full implementation in the 2006–07 school year. I am enclosing a copy of the peer report for your consideration. Based on the significance of the peers' concerns, the Department has decided not to approve Pennsylvania's proposal for implementation in the 2006–07 school year. I anticipate, however, that there will be other opportunities for Pennsylvania to implement a growth model in the future. As noted in the *Building on Results: A Blueprint for Strengthening the No Child Left Behind Act*, the Department's reauthorization proposal would permit States to include a growth model to measure adequate yearly progress, provided the model is approved by the Department. I urge you to consider carefully the peer reviewers' feedback as you work to refine your growth model for the future. My staff and I are willing to discuss the peer's concerns with you to help refine your model.

The peers noted that the Pennsylvania model shares many characteristics of the Tennessee and Ohio growth model proposals. The peers have noted strengths to this approach, such as the fact that the model uses all available test scores to make projections separately in reading and mathematics.

However, while a similar model has been approved in Tennessee, the peers were concerned that Pennsylvania does not have enough waves of data for one cohort of students through the assessment system in order to implement this proposal in all grades right now. In order to implement the proposal, at least one cohort must have gone through each of the grade spans to which the prediction model will be applied. 2006–07 will be the second year of administration in grades 3, 4, 6, and 7. In addition, the peers

were concerned about the State's capacity to track students across schools and over time.  
(Please refer to the enclosed peer report for details.)

Again, I appreciate your interest in the growth model pilot project and your continued efforts to ensure quality education for all children.

Sincerely,

A handwritten signature in black ink, appearing to read "Kerri L. Briggs". The signature is fluid and cursive, with a large initial "K" and "B".

Kerri L. Briggs, Ph.D.  
Acting Assistant Secretary

Enclosure

cc: Governor Ed Rendell  
Shula Nedley

PEER REVIEW  
NCLB GROWTH MODEL PILOT  
U. S. Department of Education  
March 15-16, 2007

PEER REPORT

State: Pennsylvania

Clarifying Call to State

No call necessary

Questions and responses (please note within each question if the State will provide additional information)

A call to the State was made during the meeting. The State was asked these questions and provided these answers. Additional written answers were provided the next day.

1. Does the state have all of the data necessary to implement this program? How can you implement this program with just two years of data?

A: Yes, we have the data. We don't have 3<sup>rd</sup> grade. For the other grades, we have data from previous analyses. For the 8<sup>th</sup> through 11<sup>th</sup> grade, we don't have the data for 11<sup>th</sup>. For 11<sup>th</sup>, we're going to use status.

Q: To project 4<sup>th</sup> grade to 6<sup>th</sup> grade, do you have data for 4<sup>th</sup>, 5<sup>th</sup>, and 6<sup>th</sup>?

A: Yes

Q: But do you have a subset of kids for whom you have 4<sup>th</sup>, 5<sup>th</sup>, and 6<sup>th</sup> grade scores?

A: We have two longitudinal cohorts, but not one that has grades 4, 5, and 6.

Q: So, you can't do projections for 6<sup>th</sup>, and it looks like you can't do projections for 8<sup>th</sup>, either.

2. It looks like the projections are computed without looking at the schools. It appears that you project if the kid goes to the average school in PA, rather than if they stayed in the school that they were in. It looks like there is bias. Low-achieving schools will be credited with growth projections that they're not likely to achieve. Would you be willing to include school effects?

A: This has come up. I don't know why we decided to do it by state average. Can I get back to you?

Q: You can get back to us tomorrow.

Overall Recommendation

<u>Recommend to Accept</u>	<u>Recommend to Accept with Conditions</u>	<u>Recommend to Resubmit</u>	<u>Recommend not to Accept</u>
0	0	0	16

Comments to support recommendation:

It appears that the State does not have enough waves of data to implement this proposal in all grades right now (because they need three years of data). It is our understanding that at least one cohort must have gone through each of the grade spans to which the prediction model will be applied. According to the Education department they cannot approve a proposal that cannot be fully implemented this year.

The panel is concerned about the prediction method proposed. In essence the method is predicting subsequent achievement for a student as if that student would be attending the average school in the state. Since the purpose of the growth model is to make prediction about the likely progress of students in each particular school, we maintain that school specific effects should be included in the model. Absent such effects, we do not see how the model can pass the prediction validity test described in the specific weaknesses.

Additional panel comments:

1. See also general guidance for future proposals below developed by the panel.

An additional panel comment:

Since the idea of growth models in the context of AYP decisions under NCLB is relatively new, and often the procedures proposed are complex and may be untested, the panel is concerned about the prediction validity of the system proposed. We suggest that in any future new proposal you do the following:

1. Please provide in one place in the proposal all of the necessary rules, procedures, statistical models and estimation procedures (if you are employing a statistical model), AMO information and so on required to operationalize your proposed system. The detail should be sufficient such that an independent third party could, in principle, build a system that reproduces your AYP growth results. It should describe how all children will be handled including: those that change schools; change LEAs; are retained in grade; transition for one school type to another (e.g. elementary to high school); and who participate in alternate assessments or with the use of accommodations. A resubmitted proposal that does not meet this standard will be considered unacceptable.
2. Provide multiple illustrations/simulations of how individual students beginning at different test levels and grades, and progressing at different rates over time would be judged under this system. These examples should be chosen to illustrate a diverse range of the data patterns that might occur and should follow these same students over time.
3. Additionally, the simulations should include students whose scores may be treated differently under the proposed model (i.e., retained students, students with missing scores, and students who participated in an alternate assessment). Please detail if there are cases where scores for some groups of students might be treated differently in the model or in an alternative growth model.

4. Provide an assurance to the panel that you have carefully considered the overall prediction validity of your system. That is, the growth accountability option involves making predictions, based on children's past academic growth, about the likelihood that children will in fact achieve proficiency at some future time point. We are particularly concerned about the possibility that the system might over predict subsequent proficiency rates in very low achieving schools. We would appreciate any information that you can provide that addresses this concern. If some extant prior years' data permit, you could apply the system detailed under 1 above to student results, say in 2004, to make predictions about status attainment, to say in 2006, and then compare the prediction to the actual status attained. An unbiased system would not result in systematic discrepancies between predictions and actual attainments.
5. Some members of the panel were particularly concerned about growth model formulas that do not take into account the possibility that some students who may be judged proficient in the current year may regress below that threshold in subsequent years. If growth projections are not made for "currently proficient students" how will your system take this consideration into account? (This consideration is especially relevant to the issue of prediction validity noted under 3 above.)
6. Please indicate if there are any sub-groups of students for which you will be unable to apply your prediction model. Note an acceptable growth proposal must detail a strategy that includes predictions for all students tested.
7. In reporting on matching rates, it would be helpful to understand the sources of non-matching (e.g. students dropped out, moved out of state, no longer participate in the testing program etc.) This would help the panel to better evaluate the quality of the underlying data system. To the extent that some data is available or can be gathered we encourage you to include this in any future resubmission.

(Use additional space as necessary)

### Specific Strengths in the Proposal

Using your notes from the Peer Review Guidance, please note areas where the proposal was especially strong, ingenious, high quality, or exceeded the Peer Review criteria. Please cite specific aspects of the proposal and include references to the Peer Review Guidance criteria (e.g., B.1.2.1) and the proposal (e.g., page numbers).

1. The proposal includes the projections for all students in the growth model in the computation of AYP.

Dissenting comments:

(Use additional space as necessary)

## Specific Weaknesses in the Proposal

Using your notes from the Peer Review Guidance, please note areas where the proposal was unclear, incomplete, or did not meet the Peer Review criteria. Please cite specific aspects of the proposal and include references to the Peer Review Guidance criteria (e.g., B.1.2.1) and the proposal (e.g., page numbers).

1. It's not accurate to characterize students who are on track to being proficient as proficient. They're on track, not proficient. It appears that they're reporting both groups to parents in the same way.
2. Because they don't have enough waves of data, they can't implement this right now (because you need three years of data). It is our understanding that at least one cohort must have gone through each of the grade spans to which the prediction model will be applied.
3. Their match rates are low. The panel was not sure that their data system is mature enough to be able to implement this program. There was a question of whether they will be able to track students well enough.
4. Their assessment proposal is approval pending.
5. Taking 8<sup>th</sup> grade data to predict 11<sup>th</sup> grade scores is quite a stretch and may prove unreliable. There is no data on the accuracy of the predictions in the model. Validity tests of the accuracy of prediction need to be conducted and reported.
6. The panel is concerned about the prediction method proposed. In essence the method is predicting subsequent achievement for a student as if that student would be attending the average school in the state. Since the purpose of the growth model is to make prediction about the likely progress of students in each particular school, we maintain that school specific effects should be included in the model. Absent such effects, we do not see how the model can pass the prediction validity test described above.

Dissenting comments:

(Use additional space as necessary)