

Growth Model October 23 Meeting Summary

(Next meeting is December 1, 2010)

Present: Dr. DeStefano, Chair; Kathlene Shank, summarizing; Amy Alsop; Ellen Cwick; Susy Woods; Daniel Brown and Jim Palmer, ISBE Assessment; Debbie Meisner-Bertauski; Laura Cresap; Joseph Matula (also, PEAC); Mark Doan; Wilma Vanscyoc, ISBE; Connie Wise, ISBE; Robin Ehrhart; Michael Afolayan (left at noon); Amy Nowell; Carmen Acevedo; Sue Morrison; Mark Mitovich; John Byrne; Robert Grimm; and Elizabeth Freeman, ISBE.

Agenda Item # 1: Charge of the Workgroup

Connie Wise shared with us relative to our charge and a December goal. She also announced that Joyce Zurkowski is leaving Illinois; Daniel Brown and Jim Palmer will be joining us.

Superintendent Koch has asked us to focus on state and district accountability.

Dr. DeStefano suggested we look at short term relative to Illinois' current assessment system and suggestions to PARCC consortium. She shared she will continue to talk to the Chair of PEAC and that Joseph Matula will continue to keep us informed.

Amy Alsop shared her concern that it was important to be mindful that the work can not be separated with a "wall in between". She shared that we need to keep in mind the relationship between the two. Dr. DeStefano says she shares Amy's thoughts.

Kathlene shared her concern with a December timeline given we only have this meeting and the next one is December 1, 2010. Dr. DeStefano sees that we could possibly agree on the characteristics of a growth model and short term recommendations.

Agenda Item #2: PARCC

The purpose of looking at this is to understand purposes and what the Consortium is proceeding to do. Dr. DeStefano sees what PARCC is doing as being "flexible".

Daniel Brown shared information on PARCC; he shared the web sites and said that Jim Palmer would post these for us. We can GOOGLE PARCC. PARCC stands for "Partnership for Assessment of Readiness for College and Careers". Dr. DeStefano shared that both consortiums share the same goals because the RFP guidelines for funding were prescriptive. PARCC has

more of the larger states and states with more diversity. Dr. DeStefano is on the national "TAC" which is the doc advisory committee that will advise the secretary of education and see the work of both consortiums.

PARCC will look at an assessment system (28 states are in the PARCC consortium) that will use college and career readiness (CCR) as an anchor; measure rigorous content and students' ability to apply that content; measure learning and provide information throughout the school year; and the plan is to leverage use of technology. SMART is looking at computer based assessments. Dr. DeStefano commented the differences between the two consortiums at this time are "subtle".

Dr. DeStefano reflected on where she perceives Illinois is relative to its Illinois Longitudinal Data System and our state capabilities and current limitations of data management.

PARCC will cover grades 3-11 and will look at three points of assessment (25%, 50%, and 75%); the assessments will be based on the current adopted common core in reading/English/language arts and math. There will be "pilot phase" and the states in the Consortium will be involved. There will contractors hired by next September. The goal is for assessments to be ready by 2014. Schools will need the scope and sequence in Jim Palmer's opinion as soon as possible.

Dr. DeStefano brought up that teachers will need to be given technical assistance and technology support if this is going to work.

The money at this point is going to development of assessments; very little to states and none to districts. Amy Alsop asked where the districts are to get the funds to do what needs to be done relative to Common Core Standards and district related curriculum scope and sequence work. She also suggested we as a group need to be sure part of our recommendations have to do with funding and support for local districts. Dr. DeStefano said she agreed this should be one of our recommendations.

Kathlene shared her concerns that we have to think through how to assure the system is coherent as PERA (Performance Evaluation Reform Act) also has a timelines across 2014-2016.

Susy Wood asked if learners with significant disabilities are included in the thinking of development of assessments and that answer was "not at this time".

One goal is that teachers will have an assessment system that provides as much for them as it asks from them. Another goal is for students to know their areas that need improvement as they move through not just at exit. The common assessment system will help leaders and policymakers make the case for support of education. Amy Nowell made the point that this is a "tool" of improvement; it will not in and of itself "improve education". There will be a summative statement made at the end of 11th grade relative to college and career readiness of each student.

The high school plan is not clearly specified at this time. The tests by 2014 will be in English/reading/language arts; course does not mean "course".

Illinois has made commitments related to PARCC to fully use the assessments by 2014-2015; to use the assessments in their state accountability systems; and to provide staff to support "partnerships" activities.

Kathlene Shank asked if these developed "assessments" are planned to be "piloted", and assured to be "reliable" and "valid" by 2014. Mr. Brown said this is the plan and he recognizes this is an aggressive timeline.

The Theory of Action is: "Building collective capacity to dramatically increase the rates at which students graduate from high school prepared for success in college and the workplace."

Carmen Acevedo brought up issues related to English Language Learners that will need to be considered.

Joe Matula brought up that teacher evaluations for non-tenured teachers occur in March and the timelines for these assessments do not seem consistent. He suggested this needs to be considered.

Agenda Item #3: State Accountability between now and 2014-Our Recommendations

The focus of our recommendations in Dr. DeStefano's thinking is on school accountability not on teacher and principal accountability.

Dr. DeStefano shared the importance of identifying the purposes. CCSSO has summarized the four common purposes for using growth: provide additional information to parents on their student's progress; enhance school accountability; improve teaching and learning; program evaluation; and evaluate teachers or principals.

Amy Nowell said she thought we were looking at school accountability. Dr. DeStefano responded affirmatively to this statement.

Joe Matula suggested that we look at a model that included same students from one month to the next, e.g. March to March; two years or more in a given district; and students that have attendance rates of 95% or more. Colleen Legge brought up the mobility piece. Kathlene brought up her concerns with what we have learned about the Tennessee and Colorado models and the need for at least 3 years of data. Colleen brought up need for fairness across districts/schools and diversity of learners, including disabilities. Mark Mitovich brought up our lack of sophistication in our current assessments and the lack of a level playing field across districts. Robin Ehrhart asked for clarification about the year to year comparison. Dr. DeStefano said we would use past performance to build the comparisons. Amy Alsop said that "it is critical to use the same measures across the state" and we need to compare "apples with apples".

Amy Alsop asked for clarification of “what can ISBE do?” Capacity is an issue. Dr. DeStefano responded that it is important to know what Illinois has. ISBE has four years of data on ISAT, IAA and the prairie state test. Carmen brought up for ELL we have 3 years of data.

Mark Doan asked about direction from ISBE relative to getting ISBE to 2015. Dr. DeStefano is talking about short term recommendation for an interim measure of growth. We currently have a status model. AYP is a status model.

Kathlene brought up the need for caution given that the states that have been using growth models have not been reporting data to districts for multiple years and the issues with ELL, special education, and other aspects of diversity.

Amy asked the relationship between using the model and making decisions relative low performing schools. Colleen asked about the growth model and the tests we are using. Colleen asked would the state be refining a method of using the AYP and growth model data.

Carmen brought up that being able to look at growth versus status could be helpful.

Dr. DeStefano said “it is additional information”.

Kathlene brought up that there is no doubt that AYP, a status model, will go away with reauthorization, and replaced by growth models but we needed to be sure that we address appropriate use.

We talked about the importance of our recommendations including recommendations about use.

Dr. DeStefano explained that one question we need to answer is how much growth is enough. We also need to consider what differences we are going to control for. Another question is “should it be the same for every student” rather than discriminating by group”. Also, are we talking about typical or normal growth versus expected rate of growth?

Statistical controls can be used in value added which can not be accomplished through a growth model that is not value added. For value added there is little evidence it is useful at the teacher level.

Illinois does not have a vertical scale in the test data we have so we must use a model that is compatible (e.g. Delaware). Delaware has a “transition matrix”. This model uses “value tables” so cut scores are set and become very important. In place of “value tables” quartiles” can be used. Kathlene asked who used “quartiles”? Dr. DeStefano says that Colorado uses a “quartile regression growth model”. The test is normed and then uses quartiles and can be used for sub groups.

Dr. DeStefano again said “we must use AYP” (is still in NCLB which will be reauthorized most likely in the year ahead).

Laura Cresap said she would want to see school data that does three things: district, school, and individual.

Attendance could be put into a “regression model”. Mark Mitovich brought up that here is technology out there that will disaggregate data.

John Byrne brought up teachers need for data. Amy Alsop brought up that there is no money to get this down to the teacher level of understanding and use and we need to focus on use at the district level.

A “difference gain score” requires a common scale. This model asks: “is the gain for a group higher or lower than average”? Validity relies on a strong vertical scale in Dr. DeStefano’s professional judgment this does not match the data we have in Illinois. This model does not do well with high and low performing schools; it does better with middle performing schools.

“Growth in relation to performance standards” (a linear equating model) if system is not well aligned this model is not a good one. Illinois has disconnects between ISAT and PSE. Colorado and Maryland have element of this. This one has problems for Illinois due to the tests we are currently using. The question asked is “did students stay at the same quartile”? Dr. DeStefano has not seen this used at the grade 3-8 level. Given we do not have multiple measures at the high school level this may not be a good one for us. This model is the one that helps look at whether that student is making progress.

Amy Nowell brought up that primary goal is school accountability relative to the decisions we are trying to make.

“Residual gain score” model does not require a common scale. It does better at both ends of performing schools versus “difference gain scores. Here we are looking at a group to see if it is higher or lower than average. This model does not require a common scale. These scores are easier to interpret.

Amy Alsop brought up we need to decide if we want something predictive.

Our current tests were not developed with growth in mind, according to Dr. DeStefano. We do not technically have a vertical scale. Question we have to answer is “do we want to go with a growth model that requires a vertical scale?”

John Byrne brought up “Explore”; it is not given to all students across the state or by all districts.

Dr. DeStefano shared information on “confidence intervals”. With some measures the band of error is so great the data is not useable. If we want to be sure that the data has confidence we need to use a “wide band” confidence interval. Given our instruments were not designed to be used we will want to recommend a “wide band”.

Missing data is an issue; for this reason we need to make rules on missing data. Precision is very important when you look at subgroups. We need to make a recommendation on the importance of transparency with missing data especially relative to sub groups.

We make want to make specific recommendations on student mobility. Some models keep students in that stay within a district; an adjustment can be made for number of months student is in the school. Focus can be on inclusion or exclusion depending on the recommendation.

Alternate Assessments are not included in most state growth models. The “Alternate Assessment” students in Illinois are currently included in state results. In Illinois we do not have the “Alternate Assessment” on a common scale. We could recommend they not be included in the “Growth Model” in this interim period. Kathlene brought up that the CCSSO documents say that one significant issue with the pilot state use of growth models was the unresolved issues with inclusion of students being assessed on “alternate assessments”.

Levels of reporting is another area where recommendations needs to be made. We discussed level of use. Opinions were expressed that districts use data at the student level and the model needs to do this. A dissenting opinion was expressed that in the interim this is not necessarily a good idea as the tests we have were not designed to be used for these purposes. We could recommend that local districts can use data in varying ways but the purpose of the state growth model is not to “drill down to student level data”.

Another decision point “Is how to use growth data in accountability decisions?” In addition to status is necessary until ESEA is reauthorized and AYP is no longer required. “Safe harbor” is federally defined so this would also have to continue so growth model would also be in addition to “safe harbor”.

Data analysis is another recommendation point. Average scores by time by grade assesses linearity of growth. False positives and false negatives are something we may want TAC to look at. Stability of scores over time of the ISAT scores is an issue that would effect useability of the data in a growth model.

We are recommending what the State needs to consider in choosing a growth model. TAC will run some models and then that group will recommend which one to use.

Amy Alsop brought up that "Residual Gain Scores" predict and we may want to have TAC look at "Residual" versus "Transition".

Agenda Item # 4 "Draft"

Dr. DeStefano will do a "draft" that we can read and think through prior to our December 1 meeting. Dr. DeStefano would like to see us have recommendations done to share with TAC which meets December 16 and 17.

Kathlene Shank shared her concern about this date and some of our members needs to share with the groups they represent.

Meeting adjourned at 1:40 p.m..