

HB 2680 WORK GROUP MEETING

June 27, 2016

Facilitator's Summary

Facilitated by DS Consulting

The following facilitator's summary is intended to capture basic discussion, decisions and actions, as well as point out future actions or issues that may need further discussion at upcoming meetings. These notes are not intended to be the "record" of the meeting, only a reminder for work group participants.

Work Group members present for all or part of the meeting: Morgan Allen (OSBA), Parasa Chanramy (Stand for Children), Laurie Dougherty (Seaside SD), Diane Mattison-Nottage (OEA-Springfield Public Schools), Susan McLain (Legislature), Colleen Mileham (OEA), Jim Popham (UCLA), Chad Putman (Coos Bay SD), Andrea Shunk (OEA), Dev Sinha (University of Oregon), Chuck Tomac (State Improvement Network), Dave Vanloo (Bend La Pine SD) and Maureen Wolf (Tigard-Tualatin SD School Board).

ODE Staff and other partners present for all or part of the meeting: Mary Anderson (ODE-OAA), Derek Brown (ODE-OAA), Holly Carter (ODE-OAA), Lisa Damold (ODE-OSS), Mark Freed (ODE-OTL), Renée LeDoux (ODE-OAA), Brad Lenhardt (ODE-OSS), Vicki Nishioka (Education Northwest-NWCC), Susie Strangfield (OIT), Steve Slater (ODE-OAA), and Bryan Toller (ODE-OAA), and Jon Wiens (ODE-OAA).

Participants on the phone for all or part of the time: Hella Bel-Hadj Amor (Education Northwest), Sandi McClary (Instructional Coach).

Facilitation Team: Facilitator: Donna Silverberg, Facilitation Support: Tory Hines; DS Consulting

Opening Remarks and Introductions

DS Consulting Facilitator, Donna Silverberg, welcomed the work group to the fourth and final House Bill 2680 Work Group meeting. Donna provided an overview of the group's progress, noting that, at the January 15th meeting (Meeting 1), the group: explored technical evidence regarding Match to Standards; discussed the descriptive study concept; and laid the ground work for future evaluations and recommendations. At the March 18th meeting (Meeting 2) the group: explored technical evidence regarding Match to Students; defined student learning gaps; looked at how districts are using assessment data in decision-making; and reviewed the proposed descriptive study methodology. At the June 10th meeting (Meeting 3) the group: established a shared understanding of the purpose of Oregon's Statewide Summative Assessments; evaluated the statewide results of 2014-2015 administration of Oregon's Statewide Summative Assessments in relation to student learning gaps; discussed descriptive study findings; and began to draw connections in preparation of forming recommendations.

Donna reviewed the day's agenda, noting that the purpose of the day's session was to draw connections between the evidence presented at Meetings 1-3 and the charges in HB 2680 and formulate conclusions and recommendations to be included in the final report to the Legislature.

Holly Carter, Oregon Department of Education, thanked work group members and staff for their solid engagement; she noted that work group questions have helped inform ODE's thinking. Derek Brown, Oregon Department of Education, echoed Holly's sentiments and thanked the group for their ongoing commitment throughout the process.

Recap of House Bill 2680 Charge and Process

Donna provided a recap of the HB 2680 charge, process, and the work group's progress relative to that charge. She noted that the work group is comprised of a variety of stakeholders who represent aspects of Oregon's education system and each member brings a unique perspective that is critical to this work. Donna reviewed the work group's charge and what the group had reviewed and discussed relative to each charge:

Charge 1: "Evaluate whether the assessment accurately measures student learning":

- a) At Meeting 1, ODE described its proposed approach to evaluate accuracy of the assessment by looking at the assessment's match to standards and match to students, respectively. The work group discussed the alignment between the summative assessments, the adopted standards and the Common Core, including how the assessments cover the depth and breadth of the Common Core (match to standards). Additionally, at Meeting 2, the work group discussed what features or qualities of the summative assessments had been employed to maximize accuracy of results for all students (match to students).

Charge 2: "Analyze student learning gaps":

- b) At Meeting 2, the work group defined "student learning gaps" as both gaps in learning for groups of students, as well as achievement gaps between student groups. At Meeting 3 the group explored how the summative assessment might be used to evaluate both types of learning gaps. In addition, the group discussed the statewide 2014-2015 assessment results in regards to learning gaps between student groups and discussed the results of the descriptive study in terms of how schools look at gaps in learning for groups of students.

Charge 3: "Identify adjustments in instruction necessary to address student learning gaps".

- c) At Meeting 3, the group also reviewed the descriptive study results for what they could learn about how schools use summative assessment data to identify adjustments in instruction to address student learning gaps. At Meeting 4, the work group continued to reflect on the information presented at Meetings 1, 2, and 3 to explore the extent to which summative assessment data can and should be used to identify adjustments in instruction to address student learning gaps.

Drawing Connections

Work group members were asked to reflect on the evidence presented at Meetings 1-3 and to draw connections between the evidence and each of the charges in HB 2680:

Charge 1: Evaluate whether the assessment accurately measures student learning

- A) Are the summative assessments clearly aligned to the adopted standards, the common core?
- B) How fully do the summative assessments cover the depth and breadth of the Common Core?
- C) What features or qualities of the summative assessments have been employed to maximize accuracy of results for all students?

The work group began by discussing these questions in small groups and then shared the highlights of their conversations with the larger group. The following points were captured on chart sheets by each small group and shared with the larger group:

- The Smarter Balanced assessments have the potential to accurately measure grade-level proficiency of the Common Core State Standards of student groups in the areas tested, if used in conjunction with other data points.
 - One year of data is not enough for the group, or the state, to draw any clear conclusions about system-wide student proficiency.
- The Smarter Balanced Assessments have the potential to accurately measure student proficiency in grade level standards if a set of optimal conditions are in place:
 - 1) technology: equal connectivity, hardware, level playing field, familiarity, well-developed application;
 - 2) test administration: accessible and on a level playing field; all test administrators proficient in the test administration manual; all students given equitable access to all accommodations; and
 - 3) uncontrollable variables such as student social/emotional well-being, absences, school schedules, etc., should be controlled to the greatest extent possible.
- 1.) Yes, the summative assessments are clearly aligned to Common Core State Standards (CCSS) (see evidence from meeting 1);
- 2.) Yes, the summative assessments cover the depth and breadth of the CCSS. The Smarter Balanced assessments help measure depth in a way that prior summative assessments did not. In the report, we recommend adding a definition of ‘breadth and depth’, background on OAKS and Smarter Balanced, and more information on assessment accessibility and accommodations; and
- 3.) Features or qualities of the summative assessments that maximize accuracy of results for all students include accommodations and accountability improvements. For example, customizable tests, test supports, and acknowledgement of limitations in technical alignment, test administration, reporting results, implementation efforts, and communication.

Charge 2: Analyze Student Learning Gaps

For charge 2, work group members reflected on evidence pertaining to student learning gaps that they had reviewed and discussed in prior meetings. Work group members shared the following comments:

- A better definition of gaps is needed. Gaps between proficiency in standards and student performance related to proficiency, as well as a difference between distinctive student groups (i.e. traditional achievement gaps and opportunity gaps).
- The summative assessments should be paired with additional diagnostic tools to identify gaps in learning for individual students as part of a balanced assessment system. It should be noted that a good summative assessment, including Smarter Balanced, can identify achievement gaps between student groups but the test alone cannot eliminate or address these gaps
- Clarification of the phrase “gaps in learning for groups of students” is needed. It should indicate that a “gap” is the space between where a student is and where we want them to be for their grade. Moreover, we should look at achievement gaps between student groups: girls and boys, race and ethnicity, ELL and non-ELL, low SES and high SES,

and SPED. For a full assessment program, formative and interim assessments that are aligned to the standards and the summative assessment are also needed.

Charge 3: Identifying adjustments in instruction necessary to address student learning gaps

Work group members reflected on the evidence pertaining to charge 3 and shared the following comments:

- For charge 3 as written, the answer is ‘no, the test cannot reduce gaps’. The test is a tool. Taking the test does not reduce gaps and the test does not identify gaps at the instructional level. The test can be a useful tool for resource allocation, enhancing professional learning, aligning standards and curriculum, and help us distinguish information gleaned through interim and formative assessment tools.
 - The test can serve as a useful point of departure to investigate how to improve system level adjustments in teaching and learning conditions. With that said, it must be used in conjunction with other pieces of data.
- It is important to integrate the data and, over time, address changes to instruction. There is a need for deeper instruction, teaching, and systems to support that professional learning. Educators should have more meaningful roles. As educators enter the profession, what signals do we send them about assessments? If we are hoping to change instruction, we need to ask how teaching is measured.
- Summative assessments can play a role, but more points of data and tools are needed to make instructional changes.
 - Some additional needs include: high-quality PD at the local level; continuing to impact instruction through assessment literacy professional learning (and measuring that impact); professional learning on quality assessment, analyzing data and using data to inform teaching; strong educator preparation, induction and ongoing training; lessons learned from the ‘Beat-the-Odds’ Schools; leadership opportunities; formative and interim assessments that are aligned to standards and the summative assessment; and professional learning around assessment accommodation and accessibility supports.
- A good summative assessment test is able to identify if system adjustments are needed to meet student needs. What additional diagnostic tools and data are needed to identify adjustments in instruction to address student learning gaps? Also consider the time and cost benefit based on the comprehensive nature of the assessment. Looking at Smarter Balanced, what are the benefits when measuring system performance compared to the time it takes to complete the assessment, the cost of the assessment and the cost of the impact to core instruction?
- There are many additional factors that make the SBA assessment insufficient for supporting adjustments to instruction. (e.g. Margin of error and differences in infrastructure). A glossary of terms might be helpful as an addendum to the final report.

Formulating Conclusions

Following group discussion, Donna asked work group members to elaborate on their dialogue and begin formulating conclusions. Laurie Dougherty, Seaside School District, cautioned the group to look for ways to streamline the assessments in order to lessen the impact on instructional time for students. The summative test is lengthy and impacts students greatly. Laurie explained that her students spent thirty hours participating in the assessment, which otherwise would have been instructional time. Laurie asked the group to phrase recommendations carefully to prevent unintended consequences and mandates on student groups.

Chuck Tomac, State Improvement Network, asked that the report be revised on page 5 to say that the state should support implementation by the school districts. Diane Mattison-Nottage, Oregon Education Association, stated that it is not the State's responsibility to dictate what the Districts should do. Yet, in her experience, the interim assessment used in her district is not aligned to the targets of Smarter Balanced. Others agreed that the summative, formative and interim assessments should be aligned to the state-adopted content standards. Susan McLain, Legislator, stated that the State is not buying another test: instead, she agreed that alignment should occur at the local level.

Andrea Shunk, Oregon Education Association, stated that the State could help students, educators, and parents become thoughtful consumers of existing assessment products. The way individuals think about assessment needs to change from viewing each as a single event to an ongoing occurrence in classrooms. Dev Sinha, University of Oregon, suggested educators have a more integrated role in the interim system to observe how students answer questions. Derek Brown noted that educators (teachers, test administrators, principals, etc.) often describe the tests as two distinctly separate tests or events (CAT and PT), when in fact it is one assessment in each content area. The perception of segmentation is an artifact of how the test is organized and administered. ODE certainly does not view the test(s) as segmented in the context of how the evidence is used to describe student learning. He also noted that it could be beneficial if the work group were to recommend that the State take a more proactive approach to supporting alignment at the local level.

Developing Recommendations

Donna asked work group members to pair off, take a walk and come back with suggested recommendations for the Legislature and State Board of Education that stem from the Work Group's analysis of the statewide summative assessment. Work group members were asked to chart their ideas prior to larger group discussion.

- A. Dev Sinha and Diane Mattison-Nottage recommended: (1) convene a group to make elaborated recommendations to districts about formative, interim, and summative assessments. Specifically, report on approaches, both in Oregon and other states, with records of success and report on commercially available assessment tools, including those available as part of the curriculum. (2) Devote resources to statewide cross-community learning about assessment literacy and deeper learning and teaching. (3) Report on potential efficiencies in assessment systems; for example, immediate reporting of some summative data and interim assessment of CCSSLA-CCSSM-NGSS through projects.
- B. Andrea Shunk and Laurie Dougherty recommended: (1) change the message about what Smarter Balanced is and is not. Move away from "test-punish-label" and one test serving all purposes. Instead, transition the message to note that Smarter Balanced provides a piece of data to support systems and ask: could that information be used differently? Smarter Balanced should be viewed as a starting point into broader discussions. (2) Decouple high-stakes decisions from summative assessment test results. What are more appropriate data to inform high stakes decisions such as educator and school evaluations? (3) Invest in high-quality, ongoing job-embedded professional learning for educators. Additionally, invest in assessment literacy professional development for other education partners like families, school boards, legislators and community partners. (4) Consider other ways to meet our summative data needs. Is Smarter Balanced the best fit for Oregon's needs (I think this better reflects our thinking)?

- C. Chuck Tomac and Parasa Chanramy recommended that ODE proactively support an objective evaluation of current tools (such as the alignment of the standards) and conduct an ongoing analysis of the Common Core State Standards. Moreover, ODE should provide ongoing professional learning in assessment literacy, standards-based instruction, effective professional learning communities (PLCs)/data teams, and accommodations in assessment and instruction. Districts and schools should commit to effective PLC/data team practices in order to gain the potential instructional benefit from a comprehensive and balanced assessment system.
- D. Colleen Mileham and Susan McLain emphasized that (1) the summative assessment is about learning (system learning, not student learning), not just accountability. Does the summative assessment need to be as long and administered as often in order to achieve the same quality of data so the educational system can learn from the results? (2) Clarify and communicate the purpose of the summative assessment. (3) The State has a contract to continue with the Smarter Balanced Assessment Consortium for one more year and there are a lot of questions regarding the three years of data already collected. (4) The State has a supportive role in the interim assessment, not prescriptive, and it is important that the State continue to support assessment alignment.
- E. Chad Putman and Dave Vanloo stated that these discussions are 80% about the summative side of the assessment and 20% about the formative side. There should be a shift of resources and time toward the formative side of the discussion. Once the conversation shifts, the group can explore what the State is able to mandate or strongly suggest and create an offer that is attractive to districts.

Following the recommendations posed above, work group members provided the following feedback:

- There were concerns over ‘replacing Smarter Balanced possibilities to include more sampling of students’ (see above, B-4). Some work group members felt that replacing Smarter Balanced was reactive; rather than recommending Smarter Balanced be replaced, it should be examined in its proper role as a summative assessment.
 - It was noted that the intent behind B-4 was to get at the summative data accurately. If Smarter Balanced is the right way to proceed, then the group should continue down that path and have deeper conversations about appropriate measures.
- One work group member expressed concern over ‘decoupling high stake decisions from test results’ (see above, B-2). In order to measure student growth, it is crucial to see what impact educators have on student growth. The summative assessment is an important data point and it can provide one measure of educator evaluation. The assessment can play a pivotal accountability role in measuring how students do on the assessment and how well they are mastering the standards (school to school, district to district, and state to state).
 - Response: The assessment is not the only measure of student success, and it leaves out other important areas of test scores, like ELA proficiency. Decoupling the summative test from educator evaluations is needed because the summative test was not developed to measure teacher effectiveness. As such, scores should not be used in teacher evaluations; teacher evaluation is a separate entity and process separate from student system evaluations.

- Throughout this process it has been indicated there is a difference between formative, interim, and summative assessments. The group has discussed why all three are needed and the different roles that each assessment plays. How, when, and where test results are applied should be differentiated by role: teacher, school and student.
- The clear message that has been addressed through litigation is that, if a test has a purpose, then there must be evidence supporting that purpose. Smarter Balanced was never developed to evaluate teachers. Instead, a new measure should be devised that has evidence to support its purpose.
 - Oregon's evaluation system does have multiple measures of evidence, including student growth. The recommendation provided in B-2 is to remove the mandate for teachers who teach tested subjects in grades 3-8 to use Category 1 statewide summative assessment as part of their evaluation, because of the lack of research, and move the framework to Category 2 (classroom, school or district level) assessments.
 - Others noted that Category 2 assessments must be aligned to Common Core.
- Holly Carter noted that the Every Student Succeeds Act (ESSA) work groups are grappling with this issue and forming an official recommendation. It is an ongoing conversation that will not be lost if it is not addressed in this work group's recommendations.
 - A work group member requested that the recommendation reword "decoupling" and that group members review ESSA findings on how summative assessments are used, or not used.
 - Rep. McLain commented that she thought the application of the assessment to educator evaluations might be outside the scope of the workgroup's charge.
 - Derek Brown suggested refining the statement to clarify that the summative assessment was not designed to evaluate teachers and that this statement would best be made early in the report.
 - Another noted that there is an opportunity with ESSA to get a deeper understanding of learning through projects, portfolios and learning assessments (e.g. see the New Hampshire model).
- It is important to note that many teachers teach more than the Common Core State Standards. Alignment is significant, but as a work group we also can suggest teachers that can do more than what is required by the Common Core State Standards.
- The work group might want to prioritize recommendations based on available resources and effectiveness.
- The group suggested exploring adjustments to streamline the summative assessment. The adjustments, when integrated with the formative and interim assessments, might provide accurate data for system evaluations and student performance, which may be useful for adjusting instruction to meet standards. Examples of possible adjustments would be how the summative assessment is organized and administered, and the amount of time taken out of classroom instruction.

The facilitator wrote tentative recommendations suggested by the work group on the white board. It was suggested that overall assessment literacy for staff and families be included in the recommendations. Morgan Allen, COSA, noted that ODE has used the available funding well; however, the Legislature needs to provide additional funding so ODE can communicate what the assessment is and how its information is used. He explained that a bigger issue is student technology literacy across the state: due to the varying access to technology across districts, this

work group should incorporate bridging the technology literacy gap in their recommendations. The group also discussed adding a recommendation about knowledge of available accessibility supports. Derek Brown suggested clarifying assessment literacy and distinguishing it from data literacy.

Work group members provided additional comments:

- Chuck Tomac stated that he did not believe streamlining the entire assessment approach is possible, but integrating the assessment with other tools is feasible.
- Laurie Dougherty noted that formative and interim student data is needed to adjust instruction, and something must be removed from teachers' plates before more can be added (such as formative and interim assessments). Moreover, there is a need for summative assessments for system level evaluations.
- Dev Sinha suggested that ODE and various work groups should come together to understand and learn from what other states are doing in the areas of formative, interim, and summative assessments.

Group members continued to refine the recommendations posed by the facilitator and reached consensus on the following list as its recommendations for next steps with the report:

1. The work group's report should be seen as an opportunity to communicate a complete message about the assessment and what it can and cannot do. For example:
 - a. The test was not designed (and should not be used) to make high-stakes decisions about individual teacher effectiveness from SSA. Legislators and the Board of Education should instead look to the recommendations of Oregon's ESSA workgroups.
2. The group recommends that policy makers and ODE should use the statewide summative assessment for what it was designed to assess. (Note: what it is and what it is not should be described in the report—Jim P and Dev agreed to help with this description)
3. The state should devote resources to statewide, cross-community learning focused on high quality assessment literacy, professional learning, and instructional enhancements.
 - a. In addition, focus on communicating and enhancing assessment understanding for communities, families, staff and educators
4. ODE should facilitate an objective evaluation (with districts large and small) of current practices and tools: formative, interim, and summative assessments and their alignments with learning standards.
5. ODE should convene a group to find and report on the efficiencies in assessment systems in order to explore a more streamlined approach. The groups should explore what other states are doing with regards to test length and testing calendar.
6. The state should provide resources to support equitable technology access and skills for students statewide, so they can take the test and be college/citizen ready.
 - a. Include information about accessibility support tools available to accommodate students and the variety of their needs.
7. The state could encourage continued review /exploration of new ways to assess students at the project/contextualized level that more matches teaching tools and aligns to learning standards.
8. To the extent possible, ODE should explore and identify adjustments to streamline the summative assessment such that, when integrated with formative and interim assessments, it could still provide accurate data for systems evaluation along with student performance data useful for adjusting instructions to meet standards. Adjustments should

be considered with regards to organizing and administering to tests to minimize time of classroom instruction.

Next Steps

Donna Silverberg thanked work group members for their participation in today's session. She outlined the process steps moving forward.

- A refined report and draft summary of the meeting will be sent to work group members on **July 11, 2016**.
- Work group members will be asked to submit edits on both the report and the summary by end of day **July 18, 2016**.
- A third draft of the report will be provided to work group members on **July 25, 2016**.
- Work group members will be able to submit additional refinements on the third draft of the report until **July 27, 2016**.
- On **July 28, 2016** the facilitator will ask work group members for their final consensus check on the report and the 6/27 summary. The question will be: can you actively support or live with the summary and report, as written (using the finger method: are you a 1-4)?

The objective is to provide ODE with a final report and summary reflecting the discussions and recommendations of the HB2680 work group by **August 2, 2016**.

Holly Carter thanked work group members for staying engaged throughout this process and maintaining a high level of commitment and passion. Work group members expressed their appreciation to one another and ODE staff for the depth and breadth of knowledge each individual brought to the discussions. Donna Silverberg expressed gratitude towards each work group member and ODE for working diligently on this process over the last few months.

The meeting was adjourned.