

State Council for Educator Effectiveness (SCEE)
University of Denver
March 18, 2011
9am-5pm

Attendees: Sandra Smyser, Jo Ann Baxter, Amie Baca-Oehlert, Nina Lopez, Matt Smith, Kerrie Dallman, Bill Bregar, Tracy Dorland, Towanna Henderson, Lorrie Shepard

Staff/Others: Ulcca Hansen, Alyssa Whitehead-Bust, Tom Elliott, Toby King, Vanessa Roman

1. Welcome and Announcements

We have three more meetings left. Today, we'll continue to drive toward consensus. We'll spend time on the principal evaluation system, new measures for student growth, final revised language for teacher evaluation system, statewide scoring framework, role of statewide model, cost study, and the appeals process.

The Council requested a summary of where we got in the whole discussion of one panel/two panels, regarding novice vs. experience. At the last meeting, we articulated the things that were important and discussed merits and disadvantages of having to panels. We didn't reach any consensus.

2. Final language revisions (*Attachment 1 & 2*)

The Council took time to read the State Scoring Framework (*Attachment 1*) and the Continuing Role of the State Council (*Attachment 2*) document and look for red flags and omissions. Edits to these documents were requested at the last meeting.

Continuing Role of the State Council:

- KERRIE DALLMAN – What happens after September 2011? Can we eliminate the section of appointments and composition?
- BILL BREGAR – Jo Ann and I are term-limited as school board members, so we no longer represent that group. So, when we're done, do we step down?
- ULCCA HANSEN – There's a balance between not representing that group and having Council members who have been part of this work for some time.
- NINA LOPEZ – The statute talks about task forces and vacancies.
- BILL BREGAR – On the first page, I think we have to have language that covers a single system or not. You could add a bullet that makes a provision for systems that are developed by local school boards.
- NINA LOPEZ – Statute says we'll provide recommendations on the performance evaluation system. Whether there is variability in the components is consistent with what we've talked about, but the statute talks about a system.
- BILL BREGAR – There's another place in the statute that talks about systems, plural.
- ***Staff will work on language to make sure we cover the statewide model system and any locally developed systems. The Council agreed to this.***

- This document has not been called to consensus yet, because we may identify things that should be on here later. For example, there needs to be work done on alignment of current policy and it would be good if the Council has a role in that. Perhaps that can be added to one of the advisory functions.
- MATT SMITH – In terms of what this means as attached to our recommendations, are we presenting this as a council consensus of this idea along with individual commitment to the prescribed activities after April, or are we presenting this as this might be a good thing to do and if we do adopt this sort of approach, we don't know what the Council might look like in May or beyond.
- ALYSSA WHITEHEAD-BUST – This is more at the conceptual level. The notion of this would be challenging if all of you decided you no longer had the time to commit to this work. We have not asked who's in and who's not.
- MATT SMITH – We might consider whether this is a valid approach if we get a majority turnover in the Council. If that's not the intent of our recommendation, what would we consider a minimum body of heritage members that would be able to sustain the work with the experience that the Council has brought?
- NINA LOPEZ – There are things that the Council has to do whether it's comprised of the same individuals or not. There will be a council whether it's us or not. I don't think there's a question of whether there's a role for the Council, it's simply the scope of it.
- AMIE BACA-OEHLERT – There are some things we're supposed to make recommendations on that I don't see in here. Doesn't the statute say that we need to make recommendations about how it relates to pay/alternative compensation?
- NINA LOPEZ – You're right. It's in the policy recommendations.
- AMIE BACA-OEHLERT – If there are things that we're making recommendations on, they should be in these boxes.
- ULCCA HANSEN – It's more of a statutory review in terms of what fits in what is and is not allowed.
- ***Staff will scrub the statute to make sure it includes all statutory requirements.***
- MATT SMITH – It seems that part of this table is continuing statutory responsibility and part is proposed continued involvement in other things. Is that right?
- ALYSSA WHITEHEAD-BUST – Yes.
- SANDRA SMYSER – Does anybody know or care if CDE wants our advice?
- TOM ELLIOTT – Yes, we do want your advice.

3. Role of the Statewide Model System (Attachment3)

The Council reviewed this document for red flags or omissions.

- SANDRA SMYSER – This is a bigger audience than just students. Parents and the general public should also have assurance.
- BILL BREGAR – Regarding the idea that the statewide model system should be as good or better than the locally-developed system, when you say better, better from whose perspective? I think it depends on the specific school district. The assumption here is that a district will either take the state model system in its entirety or develop their own in its entirety. But I think that districts should be able to pick and choose components of the model to adopt.

- NINA LOPEZ – I’m not wedded to the language. But if we’re going to have a model, it ought to be a really, really good one.
- The Council agreed that the state model system should be of a high quality so that districts want to use it.
- MATT SMITH – I think the statewide model should be executable for districts who want to use it. It should be a model that districts should be able to build upon.
- ALYSSA WHITEHEAD-BUST – I think you’re talking about the bottom of page 2, where we’re saying that there’s a system that evolves and improves over the best local system. Then, the system is decided upon in 2015.
- BILL BREGAR – I think the model system should not be a done deal at some point. I think there has to be an evolutionary process where that model system is continually tweaked, improved, revised over time. The work that’s already been done shouldn’t be ignored when that model system is put together. I’m okay with the way this is phrased; I just think there needs to be an emphasis that the statewide model system shouldn’t be a done deal.
- TRACY DORLAND – Maybe what’s not captured here is the dynamic nature that a state system should be. It should adjust based on the pilot and what districts are doing.
- ALYSSA WHITEHEAD-BUST – I think those notions are well-captured in other places.
- MATT SMITH – It might be important to use a word like “objectives” rather than “ideals”. Ideals implies that we might not think we can get there. [no opposition]
- TRACY DORLAND – I think some people will do the state model and others will do different stuff and we’ll draw a Venn diagram and over time you’ll find more things coming to the middle and you’ll be able to tease out what has to come to the middle and what doesn’t need to. The main issue we have to remember is that the needs of the kids in different parts of the state do matter. That’s where the flexibility will need to come from and I think we’ll see that over time.
- JO ANN BAXTER – I like the word ideals. When we talk about “critical elements”, I like Kelly’s use of “mandatory and discretionary elements”.
- ***Kelly will figure out what language we should use and will make sure it’s consistent throughout the report.***
- ALYSSA WHITEHEAD-BUST – Are we ready to call consensus to this document?
- ***The Council reached consensus on this document with modest word tweaks made today.***

4. Principal Evaluation System (Attachment 4)

The Council reviewed this document for red flags or omissions.

- SANDRA SMYSER – There’s no reason to increase the number of teachers who are ineffective. The point is that we want to make sure the system is implemented with fidelity.
- SANDRA SMYSER – Principals have no job protection and this doesn’t give them any. That’s just the way it is. If I have a principal who’s rated 0 on the HR quality standard, I might remove them and the other stuff might be okay. That’s not the way it is with teachers.
- ALYSSA WHITEHEAD-BUST – So, it doesn’t seem like a recommendation, but maybe reminder language.
- ULCCA HANSEN – I think we’ll have to draft something to preface the framework. Principals don’t have job protection. This isn’t changing this.
- ALYSSA WHITEHEAD-BUST – Sounds like there’s a need for a framing paragraph.
- LORRIE SHEPARD – Sandra’s saying that we need to be more cautious about imagining that there’s some standard weighting scheme. It might be more important to emphasize

transparency. The weights might not be evenly distributed nor applied evenly. You're not promising some kind of uniform weighting scheme. Not proposing a compensatory weighting system.

- ULCCA HANSEN – In some ways, I think we're going in that direction, but we wanted to make sure we had that conversation.
- JO ANN BAXTER – does portability apply to principals as well?
- SANDRA SMYSER – No.
- TOWANNA HENDERSON – Section 17, letter B, I have clarification question. Does that mean that they can opt out of using parent and student perception? I just want to make sure that we're not saying that it's an option to pursue parental involvement.
- ALYSSA WHITEHEAD-BUST – I read this to be about the measurement tools vs. the quality standards. Earlier in this document, it talks about the fact that principals should be evaluated on the full set of quality standards. Also, the only required measurement tool is peer and principal observations.
- ULCCA HANSEN – Supervisor feedback is not a shall.
- KERRIE DALLMAN – I'd like to have a conversation about having CDE develop a statewide tool to get teacher feedback on a principal evaluation, so that there's some consistency. But I'd like that tool to have some flexibility so as to allow additions to it.
- SANDRA SMYSER – Then you can add the questions around that person's personal performance plan.
- ALYSSA WHITEHEAD-BUST – So, CDE shall develop the teacher tool that'll be used for principal input that allows districts to add more? It would be a consistent statewide measurement tool that districts could tweak or add on to.
- SANDRA SMYSER – I do think that if we did this, it'd need to specifically request feedback about the quality standards for a principal, against the standards.
- AMIE BACA-OEHLERT – Don't we refer to the TELL survey somewhere?
- KERRIE DALLMAN – I'd like to see that added into 10, incorporated as part of the professional performance plan. It's about time, resources, empowerment and leadership. It's different than the statewide measurement tool that would gather teacher perception data.
- BILL BREGAR – The shall part, then, is that CDE shall develop this survey. What about the use of the survey in the evaluation? Would it be part of the data that's used? It can be given different weight then?
- TRACY DORLAND – So, 17a1 should be a may?
- SANDRA SMYSER – So, at the teacher level, we've said they shall get student feedback, but we haven't said it's a CDE tool. There's also a difference between kid and adult feedback.
- ALYSSA WHITEHEAD-BUST – We didn't say they shall get student feedback. What do people think?
 - AMIE BACA-OEHLERT – I like the idea of CDE developing a tool, but I think we do need to learn through the pilot.
 - NINA LOPEZ – When we talk about a survey, I'd like to expand it to staff, not just teachers.
- ALYSSA WHITEHEAD-BUST – CDE shall develop a tool to collect staff and teacher perception data over the course of the pilot.
- SANDRA SMYSER – If you look at the quality standards evaluation part in the quality standards for the teacher, we're not talking about a rubric or tool, but for the principal we are. I'm struggling with that.

- MATT SMITH – I’m feeling a point of disagreement that doesn’t have to be. I’m hearing Kerrie ask for two shalls. One, develop a tool and hone it through the pilot process. The other shall is use it. The use of it, being part of the weighting and weighing that each district will do. But take the results of a tool however an individual district sees fit. You shall use it as part of your input/consideration in the process.
- ALYSSA WHITEHEAD-BUST – We have congruence that CDE shall develop the tool, but there’s pushback on whether districts shall use it.
- BILL BREGAR – So, this tool becomes a part of this model system. So, they can use it in total, peel out parts of it, or create their own. I’d go along with, districts shall use a measuring tool and offer CDE-developed tool as a resource/may.
- ALYSSA WHITEHEAD-BUST – In the role of the state model system, maybe we need to make more explicit.
- KERRIE DALLMAN – We’re requiring districts to use standards and elements. This survey tool is measuring perception of the principal against the standards. I think it’s reasonable to have a statewide tool that we hone between now and 2015 with the intent that we’re going to use it statewide.
- SANDRA SMYSER – I think it’s pushing against local control so hard.
- KERRIE DALLMAN – Do you think districts will forget to survey their teachers? That’s the worry.
- SANDRA SMYSER – So you want teeth to enforce what we have?
- KERRIE DALLMAN – I don’t see this as a big deal, because it’s already required in law.
- LORRIE SHEPARD – We’re not disagreeing that it should be done.
- BILL BREGAR – Can we say that the state will develop an instrument so good that everyone will want to use it?
- KERRIE DALLMAN – I’d also like to continue the discussion of a statewide rubric for teachers and principals.
- MATT SMITH – If you break down the definition of use, the first part of the definition is that you have to use a tool and collect the prescribed data. That could be the first part. The second part could be whether we collect any data statewide that can be used to analyze the effectiveness of the system and how it’s being implemented.
- TRACY DORLAND – We’re going to want to add something about urban kids, diverse populations, but if the thing is 75 questions long, will I really be able to add to it if it’s that long?
- KERRIE DALLMAN – We’re allowing CDE to create the technical guidelines.
- NINA LOPEZ – We already have clear statements that this needs to be not so big that people can’t use it. We’ve expressed that. Also, we have a unit we’ve never had before. Their primary goal is to support districts. There is an independent goal for CDE to provide support to districts and we need to think about opportunities for this.
- ALYSSA WHITEHEAD-BUST – That statement is captured in the role of the statewide model system document.
- NINA LOPEZ – I just want folks to think about how that plays out.
- ALYSSA WHITEHEAD-BUST – we’ll come back to the notion of a statewide rubric, consistency of tools and whether the idea of a supervisor evaluation should be a shall.
- TRACY DORLAND – Can we mirror the language from the teacher recommendations in 19? [This will be done]
- KERRIE DALLMAN – In 6 at the end, I’d like to add that the tools should also meet CDE technical requirements.

- KERRIE DALLMAN – I’m concerned that we’re going to have a district evaluation committee, another committee and then the SAC or the DAC.
- TRACY DORLAND – We do that. We did that so that we could get broad representation and buy-in. We now have a group looking at this system overall. It’s worked okay. The issue with having it be the 1338 committee is that it’s a limited group of people.
- ALYSSA WHITEHEAD-BUST – This issue will be contemplated over lunch.
- KERRIE DALLMAN – Where did we get the 12%?
- ULCCA HANSEN – We talked about dividing them equally, so that’s the percentage I came up with that. We’ll come back to it.

5. Public comment

6. Principal Framework Continued

The Council contemplated whether teacher perception survey should be a shall.

- AMIE BACA-OEHLERT – In the pilot, we need to be explicitly clear. We’d have two systems. DPS, for example could pilot the shalls and come up with mays. Another district such as Adams 12 would pilot the whole system, which would include tools. Pilot districts like that would have to use all of those tools. At the end of that, we’d compare the two sets of outcomes and discuss what we mean as a Council.
- ALYSSA WHITEHEAD-BUST – I’m hearing that there are a subset of districts and districts participating in the pilot, utilizing the statewide model system shall utilize the CDE developed tools.
- AMIE BACA-OEHLERT – Right.
- TRACY DORLAND – I already think that if we’re piloting the statewide model system, those districts piloting it should use the tools. I think you’ll find that districts piloting the state system will end up in different places with the system. You’ll end up seeing different things happening all over the place. We’re just saying to come back together and look at the data and reassess.
- ALYSSA WHITEHEAD-BUST – Yes. So, if you’re using the statewide model system during the pilot phase, that includes an adherence to a set of tools.
- MATT SMITH – An objective we might be missing is that when the representative set of districts goes through the pilot process, some will use the model system, some will use existing systems and the idea is to maximize learning. One objective is to develop consistent statewide measures by the end of that pilot period that can be use statewide to understand how one approach compares to another.
- ALYSSA WHITEHEAD-BUST – But the tools should be developed at the onset so that they’re getting piloted.
- MATT SMITH – The tools can provide data that will allow us to analyze. That’s a different objective. If we say that it’s a tool that we need districts to populate so that we can collect statewide data.
- ALYSSA WHITEHEAD-BUST – That value is articulated in the role of the statewide model system.
- SANDRA SMYSER – I think maybe what’s not clear is that when we say the statewide system, we’re thinking the shalls and mays. Now we’re saying the shalls, mays, tools and all the way to the detail. That’s something else.

- BILL BREGAR – The only question I have is if we’re comparing districts, on what basis are we comparing them? Is it within our responsibilities to come up with those technical guidelines?
- MATT SMITH – I’m suggesting a recommendation that says we want CDE to develop and implement measurement tools and think about it in terms of a survey. We want to develop them as part of the process so that when we go to full implementation, you don’t get to opt out. The survey is collecting data that will allow the district-unique evaluation systems to be assessed for what’s involved.
- TRACY DORLAND – So you’re talking about a survey about the system?
- MATT SMITH – Yes
- ALYSSA WHITEHEAD-BUST – I don’t think we have time to design all of the criteria.
- MATT SMITH – That’s not what I’m asking. I’m suggesting we may want to have a recommendation that by the end of the process, you have tools like this that can give statewide understanding of what’s going on.
- ALYSSA WHITEHEAD-BUST – We probably need to add that to the continuing role of the state Council. We won’t get to it now.
- ***The Council agreed that the state model system includes consistent statewide tools, rubrics, etc.***
- TRACY DORLAND – If there are things in the pilot that don’t align with the bargaining agreement in my district (peer observation, e.g.), how many districts do we want in the pilot?
- ALYSSA WHITEHEAD-BUST – That’s in the details of the pilot that we’ll get to later

There was a suggestion that we reference the TELL survey specifically in 10. The Council contemplated this and revisited.

- KERRIE DALLMAN – I think it’s a resource that’s used to set goals in the professional performance plan. It was referenced in statute. The legislature, districts and schools have yet to use that data in any substantive way. Including it as part of a principal’s performance plan is important. I’d like to call out that TELL is one of the pieces that should be reflected in the PPP
- TRACY DORLAND – If we’re going to do that, we should maybe reference other surveys.
- ALYSSA WHITEHEAD-BUST – So, resources include...and then list them.
- KERRIE DALLMAN – I’m proposed that TELL be a shall.
- JO ANN BAXTER – If it’s good data, why aren’t people using it?
- KERRIE DALLMAN – Because it’s only been 2 years.
- BILL BREGAR – How the district uses the data from shall would be up to them? Yes
- SANDRA SMYSER – My district would pick other things other than the shall.
- KERRIE DALLMAN – That would be optional.
- JO ANN BAXTER – I agree with Sandra. TELL doesn’t have the experience we’d like for it to have. If you have to have 50% participation in order to get results, how can you require that it be used?
- KERRIE DALLMAN – We can require that it be used for those who have 50% participation rate. I think TELL is valuable. It’s a consistency piece. It’s about using it as a tool to engage your staff. IT’s nothing more than that.
- ALYSSA WHITEHEAD-BUST – A compromise: The Council recognizes that TELL is required by statute and strongly encourages use by districts.

The Council agreed that 15 and 17 should be merged.

The Council discussed whether a reference to supervisor should be added.

- TRACY DORLAND – I think you have to have the supervisor of the principal be a shall in the principal's evaluation.
- ULCCA HANSEN – The question was whether supervisors are really going to observe the principal. They'll monitor the development and the PPP. The reason we didn't put it as a shall is because they said they're involved in the PPP evaluation, but not observing. It's a question of what, exactly, it means to have the supervisor involved.
- ALYSSA WHITEHEAD-BUST – The question is whether or not a shall pertains to supervisor input.
- SANDRA SMYSER – Who's doing the evaluation if it's not the principal?
- KERRIE DALLMAN – It could be the board of ed in small districts.
- SANDRA SMYSER – Size is a huge factor.
- KERRIE DALLMAN – If you have a system that has true pure evaluation, it might be a supervisor.
- ULCCA HANSEN – We could add language to 8 about who will be in charge of aggregating the measures. We can parallel language to teacher framework in 18 and add that districts shall clearly outline the process whereby measures of principal performance are aggregated.

The Council talked about where the 1338 committee fits in and how it fits with DAC and SAC.

- KERRIE DALLMAN – I don't think we should try to sidestep negotiated agreements here. We negotiate with our district about how folks are appointed to agreements, so I think here we need to identify if the 1338 is the right people and not get to who appoints those people. We ought to focus on that we have these entities and who is making the decisions.
- TRACY DORLAND – We don't have an option to not have 1338 be in there because it's in statute.
- KERRIE DALLMAN – I don't want to not have members of the principals or teachers association involved. We have the 1338 committee, the SAC and the DAC. WE have to figure out who's doing this work. That should happen at the district level. I just recommend that legal people look at the SAC, DAC, 1338 language and bring recommendations to the next meeting.
- LORRIE SHEPARD – We don't want proliferation of committees. I need more understanding about levels. What's now a state responsibility, district responsibility? Accountability and personnel evaluation overlap to the extent that the measures start being shared. Especially when we get into issues like climate.
- NINA LOPEZ – I don't want to leave here with the sense that staff is going to create that comprehensive list by next week. I don't want to leave here with the false sense that we'll compile that information. Lorrie's suggestion to figure out the interplay between the different levels in this system might not occur by next Friday
- ALYSSA WHITEHEAD-BUST – Reference in policy recommendations that this is an area where some further scrubbing needs to be done.
- KERRIE DALLMAN – I think this ought to take place before the summer.
- LORRIE SHEPARD – I agree. Don't think it should be put off.
- NINA LOPEZ – I think it goes in the district implementation guidelines. I think a sense of urgency makes sense.

- ***The Council agreed that language be added to the district implementation guidelines that specifically reference looking at the alignment of the 1338 committee, the DAC and the SAC in order to not create redundancy.***
- The Council also agreed to make a note of broad constituent involvement.

Matt raised a concern about student perception data.

- MATT SMITH – IN 17a, we used strong language about student perception data; we don't mention specifically that it includes an actual evaluation of a teacher. In b, it's as if the wording is purposely vague. Are we trying to be vague? It's almost as if we're limiting it.
- ULCCA HANSEN – I think it's an effort to craft that we want to target what we've asked principals to do.
- MATT SMITH – Are we trying to steer away from principal evaluation and steer toward things that are associated with the results of an effective principal?
- ULCCA HANSEN – It feels like everything's captured in "school culture and learning environment"
- TRACY DORLAND – In teacher recommendations, it refers to "student data", we could make that language consistent.
- SHELBY PARKER-GONZALES – I feel more comfortable if we mirror the language that's in the teacher framework document.

The Council discussed the weighting policies recommendation

- ALYSSA WHITEHEAD-BUST – Do we want to set a minimum percentage for the quality standards?
- TRACY DORLAND – I think we should make it flexible and just say that all standards must be measured and districts can decide the percent.
- KERRIE DALLMAN – Would we allow districts to weight one standard nothing? Or .01%?
- TRACY DORLAND – Yes
- NINA LOPEZ – In some respects, it doesn't really matter because they're under contract.
- TRACY DORLAND – I think it plays out in context, what type of school you're in, the history of that school, the type of community.
- LORRIE SHEPARD – I think we should consider not requiring a percent weight. Have all of the standards be addressed. The weighting in a given year should be transparent.
- SANDRA SMYSER – And line up with the PPP.
- LORRIE SHEPARD – There will be certain circumstances where one of the dimensions is so important that you want to weight it more than 50%.
- ALYSSA WHITEHEAD-BUST – So, the counter-proposal: language where districts develop a framework where all quality standards are addressed, in alignment with PPP.
- MATT SMITH – Is there a philosophical incongruity that we only define standards that are fundamental and flexible, yet we're saying that they could be 0 weighted?
- SANDRA SMYSER – The context for a principal's work is slightly different than that of a teacher's work. If you're really bad at one standard, you're job's at stake. With teacher's that's not the case. With principals, there should be less tolerance.
- ALYSSA WHITEHEAD-BUST – If you do the 10%, the discretion is the same as it is with teachers. With that, are you comfortable with the 10% threshold?
- SANDRA SMYSER – Yes, I'm good.

- ***The Council agreed that standards for principals should all be weighted, none less than 10%.***

7. Measuring Student Growth in Principal Evaluations (*Attachment 5*)

Discussion

- KERRIE DALLMAN – In 14a, there’s a reference to the Colorado growth model. We don’t want to incent a principal to take their poorest performer, yet not ineffective teacher and load them into grades 11 and 12.
- ULCCA HANSEN – Is that captured in b?
- KERRIE DALLMAN – But it’s a may in b, which means they wouldn’t have to consider student growth in 11th and 12th grade.
- ULCCA HANSEN – The only thing we have required for principals’ growth measure is school performance framework. Everything else is a may. I think that’s part of what we want to revisit. If we say they have to do it, we have to be clear about how it’s rolled up.
- RICH WENNING – We will measure growth between grade 10 and 11 and have also correlated data between CSAP and ACT.
- ALYSSA WHITEHEAD-BUST – Do we want to signal back to the language about development of new measures?
- KERRIE DALLMAN – I worry. 11th and 12th grade teachers will have student growth measures attached to them. Why not principals?
- LORRIE SHEPARD – I hope we’ve been clear about how problematic SGOs are regarding the non-tested subjects. They’re exacerbated for principals. I think this is the place to put things in about graduation rates, increasing percent of kids doing AP and IB, kids who go to college, dropout rates between grades, calculating it.
- KERRIE DALLMAN – I’m good with it. Just didn’t want us to be silent on paying attention to 11th and 12th grade.
- ULCCA HANSEN – Also, you have an indirect aggregation of student growth measures represented through the number and percentage of teachers piece.
- RICH WENNING – Another measure could be the on-track to graduate piece of data.
- ALYSSA WHITEHEAD-BUST – So, do we want to recommend language about graduation rates, AP, IB, reduction of dropout rates, on-track to graduation rates?
- AMIE BACA-OEHLERT – So we should say that they shall use additional measures and they may include any of the following.
- NINA LOPEZ – The growth measures for principals ought to explicitly encompass things beyond that which teachers are evaluated on.
- ***Staff will draft a statement about the evaluation of principals including measures that go beyond what teachers are measured on.***
- TRACY DORLAND – We should make it explicit that if you’re a principal responsible for students in ECE-5 grades, you should pay attention to numbers 12 and 13.
- NINA LOPEZ – Do our standards apply to preschool teachers? If they’re licensed. Have we paid enough attention to this?
- TRACY DORLAND – I think we need to pay attention to this in the pilot. I think we’re okay in the standard area, but we should look at the tools.
- ALYSSA WHITEHEAD-BUST – So, this applies to licensed teachers who teach 3 and 4 year olds.
- NINA LOPEZ – On the first page, last bullet, the last sentence isn’t clear. How do you act on it?

- ULCCA HANSEN – This was part of a conversation about making clear that the SPF is limited in what it measures, we want principals to focus on a range of things not necessarily in the SPF. Do we want to explicate further?
- LORRIE SHEPARD – We are always worried about standardized measures that are a narrow subset of goals. It doesn't include everything. If it becomes the whole focus, it has a distorting effect. Is this full steam ahead where you want all of the effort to be?
- NINA LOPEZ – I feel like a lot of the recommendations talk about other things. I don't want to leave it in a recommendation.
- LORRIE SHEPARD – We need to re-ask the question is this set of standards adequately represented by the SPF? It's a reflective piece.
- RICH WENNING – The SPF is a parsimonious approximation of the broader set of performance indicating that the Department will do. You might instead reference the performance indicators mandated by statute, including the SPF. You might consider referencing the performance indicators and the SPF. That might help to make the net a little wider.
- **Staff will revise language to broaden beyond the SPF.**

8. Student Growth (*Attachment 6*)

Lorrie gave an introduction to this document. The Council reviewed this document for clarity, red flags and omissions.

- TRACY DORLAND – It might be good to add something about “with anchors”.
- LORRIE SHEPARD – Break up the Recommendation 2 to specify that the state should consider expanding state summative assessments in the areas of social studies and science.
- NINA LOPEZ – Can we reference the Colorado Academic Standards in social studies?
- TRACY DORLAND – The consortia working on common core assessments are working on literacy and math, so we're not duplicating their efforts, right? Right.
- NINA LOPEZ – On page 3, under recommendation 2 think we want outcomes in early grades. I'd hate for us to leave thinking that the only thing we focus on is literacy. So, can we take out “in literacy”? [Agreed]
- LORRIE SHEPARD – Should have broad language about the language of the assessments all being high-quality, higher order thinking, “next generation”, critical thinking, etc. and focus resources on some areas. Further define what a high quality summative assessment means.
- TRACY DORLAND – Should clarify that the consortia should focus on “other”.
- NINA LOPEZ – On the last page, last sentence in b, revise to say “teachers should be broad enough to capture the breadth of what teachers teach.
- MATT SMITH – Recommendation 2 at the bottom, it goes into explanation of a set of order instructional tasks. I don't know if it's emphasized enough or comes across that it's really a strong recommendation the Council has. Given that this is an area that anybody could use, starting now.
- ULCCA HANSEN – It would make sense to pilot it.
- LORRIE SHEPARD – The law doesn't let you set priorities. It says you have to do everything. We know that's not the right way to get things done. Anything we can say, without violating what the law says we have to do, that would encourage people to do something well, learn from it and then do the next thing, I'd love to have that kind of language.
- ULCCA HANSEN – We've started to do that with the implementation guidelines. Next week, we can think about what we want to sequence in what way.

- NINA LOPEZ – I feel like there are a couple of statements like that that we’ve said throughout our work that might go at the beginning of the report.
- MATT SMITH – A lot of the most incredible innovations are just using something you’re already doing in a different way. Everybody’s developing, teaching, reviewing instructional products, so to use them in this way, learn how to use them in an ordered sequence to measure growth is a phenomenal opportunity. It fills this gap of having more statewide assessments. If we believe that, I don’t see a problem in including that as something we recommend people taking advantage of.
- ALYSSA WHITEHEAD-BUST – So, the Council has agreed to three changes: further explicate instructional tasks, further explicate what we mean by high quality, better delineate the difference between state developed and consortium developed assessments, priorities around framing that we’d put elsewhere.
- ***The Council reached consensus on this document knowing that staff will make those changes.***

9. State Scoring Framework

The Council reviewed this document.

- MATT SMITH – We use the word “achievement” through this document, but it doesn’t necessarily mean growth.
- ALYSSA WHITEHEAD-BUST – You’re right, that should say growth. We’ll fix that.
- AMIE BACA-OEHLERT – can we discuss the concept of counting towards non-probationary status?
- LORRIE SHEPARD – My big question is how do we get from 6 categories to 4 on the two dimensions?
- ULCCA HANSEN – This table is not meant to be the final version, just an illustration of what it could look like. There’s a strong argument for it to be 5x5 or 6x6, but we just used this to get an idea of what it looks like.
- SANDRA SMYSER – But no matter how many numbers there are, there are just four colors?
- LORRIE SHEPARD – Yes, that’s fine. But, it will be politically unacceptable to have one red box.
- MATT SMITH – Given that the details need to be discussed, the structure that delineates which elements of the matrix are counted against years of effectiveness leading to non-probationary status or years of ineffectiveness leading to the opposite, I think that structure is important.
- ALYSSA WHITEHEAD-BUST – We were hoping that structure would be the sweet spot.
- TRACY DORLAND – Does this assume no novice panel?
- ALYSSA WHITEHEAD-BUST – Yes. The only reference to novice is how it accrues or doesn’t accrue non-probationary status.
- LORRIE SHEPARD – You could have a different weighting scheme for novices to end up on the 1-6 score.
- NINA LOPEZ – Would you weight it in a different way that the way we talked about it before with the minimums?
- LORRIE SHEPARD – The profile behind the matrix would be different. You would be adjusting the weights. Once you had your score, it has the same effect of what was being discussed about treating the progressing and effective differently. It’s a different way of acknowledging it. That’s more acceptable. He wasn’t as happy with the leadership standard anyway.

10. Cost Study (*Attachment*)

11. State Scoring Framework Cont.

- AMIE BACA-OEHLERT - Even if a non-probationary teacher becomes probationary, they can be non-renewed. Any rating other than ineffective should count toward non-probationary status.
- LORRIE SHEPARD – I think we’re getting political feedback that says no novice, but some opening for considering differential weighting for novice. If that’s the way to handle novice, I’ve got to settle in my mind whether the scoring rubric will be uniformly applied. Why do we want four categories instead of 3?
- ULCCA HANSEN – If you have 3, everybody ends up in the middle. Having 4 allows for more differentiation because you force people to make a choice.
- LORRIE SHEPARD – Then, you really have to make a choice.
- NINA LOPEZ – You could keep 4 categories and create the opportunity for differentiation, without having to have a new label.
- LORRIE SHEPARD – The 4 categories refers to the colors.
- NINA LOPEZ – Each category could have more than one box to it.
- LORRIE SHEPARD – How to handle novice? how many color?, on the two dimensions how many categories should there be? What’s the implication of the colors?

Number of performance categories

- SANDRA SMYSER – They’re not equally distributed groups. I’m comfortable with four categories and progressing effective is one of the ineffective categories. The implications of ineffective is a small percentage in the distribution.
- LORRIE SHEPARD – If you have four colors, somebody that’s above median growth is being called probationary and declared below the effectiveness cutoff. The diagonal is dangerous with so few categories.
- NINA LOPEZ – You could have the same categories.
- LORRIE SHEPARD – Someone who’s getting good growth is being unfairly treated if we do this. I am trying to make the number of scoring categories consistent with the data we see.
- NINA LOPEZ – If we’re not going to issue this box, then, can’t we resolve this?
- AMIE BACA-OEHLERT – How can we talk about implications if we don’t know how they’re going to play out?
- BILL BREGAR – If I’m a teacher who’s scoring a 4 on student growth, but 1 on professional practice, are my students growing that much in spite of the teacher or because of the teacher?
- ALYSSA WHITEHEAD-BUST – I think we could come to congruence around the number of categories.
- MATT SMITH – We’ve taken our best shot at defining quality standards that we believe the research says should make a difference. Our premise is that based on what we know, if you really execute these standards, you should get good results. The other side of the coin is that there may be a parent anomaly in a given situation, where someone’s formative assessment is off the charts and growth is low or vice versa. In the context of this system, you have to start somewhere with the evaluation and that would be an example of what would go into an appeals process. I think the overall rating for that sort of combination of

- things, you'd want to look at more intensively. You don't want to assume that good growth overcomes all or good practice overcomes all. I'm not sure this allocation is necessarily bad, if this system can compensate for it.
- KERRIE DALLMAN – That would argue that the appeals process is not just about process, but about substance as well. We need to consider that when we get to the appeals conversation again.
 - SANDRA SMYSER – If we have the approaching category and that's where novices land, that's okay. I'm waffling on the idea that we have a different standard for them.
 - LORRIE SHEPARD – I'm not willing to call them ineffective if they're doing everything well but leadership. I'd like to allow differential weighting and leave that to the district. What we know about performance of novice is that a lot of them have strength in content, good knowledge about pedagogy, but they're bad at putting it all together all of the time. I think we could leave it up to principals/districts to decide how to weight those things. So, on the third page under b, there's a section that defines novice, makes it clear that when we get to the scoring framework, they'll enter it with number not directly comparable and that districts can make decisions about how to weight the standards for novices.
 - MATT SMITH – If our expectation is that a novice teacher should be capable of achieving 15%...
 - AMIE BACA-OEHLERT – That's the weight, though.
 - MATT SMITH – If they're low in a standard, you could weight them as low as 15% which accommodates a low performance level.
 - ALYSSA WHITEHEAD-BUST – So, would it be okay if they wanted to weight leadership at a 0?
 - MATT SMITH – You're saying that you think it's likely that new teachers need to have 0 in some standards in order to be effective.
 - TRACY DORLAND – We say that this is the standard for performance teaching, but we understand that when you come in new, you might not be there but you have to get there.
 - KERRIE DALLMAN – Do we have any consensus around these values?
 - LORRIE SHEPARD – Bullet 4 is stated in offensive manner. I think that moving away from the current three years to non-probationary status is a huge decision. I come at that not from trying to protect performance, but saying we don't know. To build in something that's going to have that dramatic effect, I will not stand behind this negative language, but we're tinkering with our ability to recruit and retain in the first five years.
 - TRACY DORLAND – If you don't hit effective in your third year, it doesn't mean you have to be non-renewed. The allowance gives us flexibility of letting new teachers grow.
 - SANDRA SMYSER - We keep thinking about non-probationary teachers losing their job.
 - AMIE BACA-OEHLERT – The reality is not this loving support environment. It's a high stakes, cutthroat environment. It is very disincentivizing. I don't know how you create strong language that ensures a supportive reality. The reality I see is that you're expected to be a master teacher. That doesn't happen overnight. I know that it takes time and support to become a master teacher. There's so many loaded assumptions. I think we need to be careful. It implies that people are hiding behind status. I want us to move away from assumption type language. I agree with Lorrie that this is a huge issue.
 - JO ANN BAXTER – We're talking about novice and saying that approaching effective is a path upward, but for the non-probationary teacher is going the wrong way. If we say that progressing towards effective will cause you to lose non-probationary status, I think you'll be disincentivizing teachers who may want to leave the profession. We need to figure out

some way to alleviate that. That concern is more for experienced teachers. I agree with where we're going for novice teachers.

- MATT SMITH – Are you suggesting that the two year in statute is fundamentally unsupportable?
- JO ANN BAXTER – Not if you're ineffective. This middle ground thing is bothering me. Where is the advantage to a teacher?
- LORRIE SHEPARD – It would help to have some experience with data. Has DPS ever done this cross-tabulation? I'm hung up on the different dimensions of practice. If that's controlled by the principal, they answer this question, adjust it mentally. They give a 3 instead of a 2 if they believe they've seen a disproportionate number of special needs kids or whatever. So the cross-tabulation in a grid like that...As you start going up, I don't know what 2s, 3s and 4s on growth are. As soon as you see the people on those categories, you start adjusting the colors. Because we don't have experience with it, it's hard. If DPS had that cross-tabulation, you could get an idea of where to land on it.
- ULCCA HANSEN – I suggest we lay out some of the issues. We've already said there'll be no implications for the ratings. For the pilot, you can say that you'll collect data about how teachers are falling out and where they would be on the panel. It's buying ourselves the time to look at the data.
- JO ANN BAXTER – I would suggest not adding a fourth.
- NINA LOPEZ – We will have some data, but I'm not convinced the data will solve this. Whether you have 4 or 5 or 20, you're going to stick them where they need to be. I would advocate more for, if you're not effective you ought to tell them that. I feel like it's hard for us to get out of where we are. We're talking about changing the system, changing expectations. So, I think the unintended consequences are innumerable. I think you're right to say that this might have a negative impact on people entering the profession. But it feels like it compounds itself and we won't be able to fully contain all of those. I think we need to challenge ourselves to think about it differently. Statute does say you have to look at the school environment. I just feel like we're stuck in where we are.
- BILL BREGAR – We're operating on an assumption that the values of professional performance levels will be exact levels. They won't be. This system takes what we've been doing and adds the student growth piece. Now it's more complicated. I'm thinking, if I could flip the calendar back to when I was a beginning teacher, how would I want to be evaluated, I'd want to know exactly how I'm doing. I'm going back to thinking that the three colors are just fine. By adding a fourth, we're not giving an honest evaluation to teachers.
- ALYSSA WHITEHEAD-BUST – Is there anybody who feels wedded to more than three categories?
- TRACY DORLAND – I do. I think if we're going to raise the standard of what instruction looks like, we have to raise the expectation of people who are in the classroom now. In a time of change, we want to be able to tell you where you are against where we need you be. I think it allows for growth and differentiation against folks who aren't quite there.
- KERRIE DALLMAN – I do.
- ALYSSA WHITEHEAD-BUST – Most of you suggest liking having one panel. There are questions about what happens behind that panel. At this point, we could contemplate recommendations that allow for intense scrutiny in the pilot so that we have more information to determine implications. Or, we could engage Council members to write a couple of different (minority/majority) reports. Are there other paths that folks see?

- ULCCA HANSEN – I know Tracy wants 4 colors, but also differentiation, which gives more room. Is there a possibility to come to consensus on some of these?
- TRACY DORLAND – I know that there will be teachers who come to work every day wanting to meet the needs of their kids, but they don't know how. They're struggling. To say "ineffective high" is different than "progressing". I think the labels matter, especially to non-probationary teachers who are trying really hard.
- ALYSSA WHITEHEAD-BUST – We have one more substantive meeting and a meeting we'd like to reserve. What's the most helpful process at this point?
- MATT SMITH – I think because we have to provide a significant level of flexibility in whatever we roll out, I think we should err on the side of simplifying things rather than being more complex. I think we should do a simple matrix with the right level of flexibility built in. A novice teacher is different than a teacher who's been rated ineffective. We'd never hire someone who we expect to be ineffective. I don't think we'd do that for teachers either. If the expectation is that I'm going to bring in a new teacher and he or she is going to be effective, I don't need a different set of standards. If I'm expecting a significant number of teachers to come in as below what I want as a standard, maybe I should design something to that. For the system we're trying to create, I don't think there's an expectation that we're going to hire ineffective teachers.
- ALYSSA WHITEHEAD-BUST – I don't think we'll get to a place to how we're going to actualize our values. Where are we leaning?
- SANDRA SMYSER – If we haven't reached agreement on something, we should say why. The problem with the questions that we're dealing with is that the answer is often "it depends". I think we just have to say that stuff. If we can't reach agreement, we need to say why. We'll have to punt to the pilot.
- TRACY DORLAND – That makes sense to me. I don't know that we're going to reach agreement.
- LORRIE SHEPARD – I can buy into Tracy's four categories with the yellow category being not effective, so long as we also have language that says that we can't do that if it's 4x4. I think that we need to say that we have the least experience with what it means to cross-tabulate these. It shouldn't have to go all the way to an appeal. I think it could be worked out in the pilot.
- ALYSSA WHITEHEAD-BUST – We'll capture some of these high level dialogues and ask that these be further investigated in the pilot.
- AMIE BACA-OEHLERT – I am opposed to a minority/majority report. I believe that this isn't an issue where we're split. I think we're split because we haven't had the time to dive into it. I am in favor of capturing the high level conversation. I am in favor of using the pilot, because I do believe it. I respect everybody here and their opinion. I just worry about us making decisions that we don't know what the consequence will be. I am in favor of the pilot. But, we don't have any of that set up yet.
- ALYSSA WHITEHEAD-BUST – so, the path is an option of capturing our conversations and some of our values and teasing out some of the questions and ways in which data should be analyzed at the end of the pilot.
- Matt Smith – I think there's a level of structure that we all agree on. Where we start to wonder about our agreement is in the detailed level of structure of content. It seems like we could recommend the way we think this should be structured and applied, but we don't feel like we have the level of data to go to the level of specificity that's needed. We'd need additional data through the pilot process. If we take that approach, we can get to it.