

	Year 1 2011-12 CSAP	Year 2 2012-13 TSAC	Year 3 2013-14 New CSAP ?	Year 4 2014-15 New CSAP ?	Year 5 2015-16
Context and implications	Principal evaluations with new system; teacher instruments used for training and feedback; <i>blue for principal only</i>	Both teachers and principals evaluated using new state model system <i>red for teacher instrument</i>	Bring other licensed personnel on-line First year evaluations count for earning non-probationary status	SCEE provides updated recommendations to SBE First year evaluations count for loss of non-probationary status; appeals for 2013-14	
Strand 1 – Validation	<p>GENERAL</p> <ul style="list-style-type: none"> • <i>Usability and utility of the scores</i> • Content of rubrics (all captured, redundant etc.) • Do teacher and principal rubrics hold the promise of being fair and reliable? • How do differences in combining multiple measures and weighting of standards by district effect overall performance ratings? <p>SPECIFIC</p> <ul style="list-style-type: none"> • develop out statement of the purposes and outcomes of the system • collect baseline data around the system outcomes • Explaining what it means to be a pilot and requiring districts to a priori make decisions around weighting of standards. • COMPOSITE REPORT OF TEACHER SCORES OR PRE-WEIGHTED SCORES BY STANDARD • collect baseline data around issues such as training and other aspects of systems implementation [separate column] • determine whether elements are associated with standards • ensure rubrics contain all necessary elements of professional performance • ensure rubric ratings effectively differentiate levels of performance • determine if evaluation feedback is timely and useful to end users 	<p>GENERAL</p> <ul style="list-style-type: none"> • <i>Usability and utility of the scores</i> • Were modifications adequate to address the issues? • Full-blown validation study for the pilot districts • Data on weighting of standards and how this effects outcomes <p>SPECIFIC</p> <ul style="list-style-type: none"> • ensure rubrics contain all necessary elements of professional performance • ensure rubric ratings effectively differentiate levels of performance • determine if evaluation feedback is timely and useful to end users • examine whether evaluation is driving professional growth • determine whether elements are associated with standards • identify practice descriptors that do not adequately differentiate performance level • assess score reliability • assess score variance by relevant/irrelevant characteristics • begin examining use of the new state testing system for evaluation 	<ul style="list-style-type: none"> • Reliability – are the scores in the pilot districts more reliable than in the new districts coming on-line? More stable? (Do year 2 all over again with sample from entire state) 	<ul style="list-style-type: none"> • If changes based on 13-14, have they addressed the issues arising? • Re-validation if necessary • In-depth studies based on initial validation study findings. 	<ul style="list-style-type: none"> • Impact studies to determine long-term effects of SB 191.

		<ul style="list-style-type: none"> • identify practice descriptors that do not adequately differentiate performance level • assess score reliability (use of videos) – scores from the scoring process • assess score variance by relevant/irrelevant characteristics • instrumentation? What are the instruments? Rubrics, summary sheets • developing process and protocols for feedback to evaluator and evaluate 					
<p>Strand 2 – Process/ Management</p> <p>(Outcomes of this will effect interpretation of S1)</p>		<ul style="list-style-type: none"> • Training – how well does the training capture the issues that need to be addressed (fidelity of implementation) – time involved? What needs to be required for CDE certification of evaluator trainers and trainings? • Change management – asking if it is working or needs to be enhanced. • Is the process working? Too cumbersome? Resonate with people who need to use it? • How much time is required for the whole process? 	<ul style="list-style-type: none"> • Does the evaluation take less time? • Do modifications improve the system? • Development of a checklist/power elements/spot observation processes? 	<ul style="list-style-type: none"> • Stabilization of outcomes over time • Checklist development 			
<p>Strand 3 – Student Growth</p>		<ul style="list-style-type: none"> • What measures are being used by districts and how? • What would the implications be of using these measures for calculating student growth for the purposes of evaluating teachers? • What do districts need to do to measure student growth • Training around SGOs – is it adequate, are the goals being developed rigorous, consistent, process in the districts? <p>SPECIFIC</p> <ul style="list-style-type: none"> • evaluating process of setting goals and determining the quality and the rigor of the goals • districts doing a review of their current 	<ul style="list-style-type: none"> • Shared variance among measures • Piloting of new measures of student growth (validation and these new measures) • Comparison of student growth outcomes on new measures against CSAP? 	<ul style="list-style-type: none"> • Piloting of new measures of student growth (validation of these new measures) • <i>Determining where and how correlations exist between student growth and professional practice scores</i> 	<ul style="list-style-type: none"> • Piloting of new measures of student growth (validation of these new measures) • <i>Determining where and how correlations exist between student growth and professional practice scores</i> 		

	<ul style="list-style-type: none"> assessments and their quality as measured against technical guidelines issued by CDE add content collaboratives collecting feedback from the collaboratives begin process of developing and validating new measures 					
<p>Strand 4 – Model System versus districts mapping to framework</p> <p>What are the anticipated outcomes for educators and students under model system in pilot districts?</p>	<ul style="list-style-type: none"> Collect data on questions re: outcomes Likert scale – to improve the system <p>Qs 4(a)(i)(1)-(3), (11); 4(a)(ii)(1)-(2)</p>	<ul style="list-style-type: none"> Collect data on questions re: outcomes <p>Year 1 PLUS Qs 4(a)(i)(4)-(5), (9)-(11) and 4(a)(ii)(6)</p>	<ul style="list-style-type: none"> Collect data on questions re: outcomes <p>Years 1 and 2 PLUS Qs 4(a)(i)(3),(7 for teaches)(8),(12)</p>	<ul style="list-style-type: none"> Collect data on questions re: outcomes <p>Years 1, 2 and 3 PLUS Qs 4(a)(i)(6)(8)</p>	<ul style="list-style-type: none"> Collect data on questions re: outcomes <p>Years 1, 2, 3 and 4 PLUS Qs 4(a)(i)(7 for all else), 4(a)(ii)(3)-(5)</p>	
Partner districts	<ul style="list-style-type: none"> Collect data on all outcomes as with pilot districts Conduct studies of partner district systems and integration districts similar to those of state supported districts. (on-going) 	<ul style="list-style-type: none"> Conduct studies of partner district systems and integration districts similar to those of state supported districts. (on-going) 	<ul style="list-style-type: none"> Conduct studies of partner district systems and integration districts similar to those of state supported districts. (on-going) 			
Framework as a whole?	<ul style="list-style-type: none"> Which elements of framework align with existing systems? Which don't? Adjustments? 	<ul style="list-style-type: none"> What are the differences in outcomes for pilot versus partner versus integration districts? 	<ul style="list-style-type: none"> What elements of the framework elements are redundant or missing? 			
<p>Strand 5 – Support for Implementation</p>	<ul style="list-style-type: none"> Define standard for what makes the system ready for statewide rollout Collection of data from districts of varying sizes and situations (unique characteristics) to change/improve CDE approach (case studies/qualitative research) 	<ul style="list-style-type: none"> Determine whether system is ready for rollout based on criteria set in 2011-12. Collection of data from districts of varying sizes and situations (unique characteristics) to change/improve CDE approach (case studies/qualitative research, additional supports) identify needed policy alignments determine final delivery platform for the evaluation system 				

			<ul style="list-style-type: none"> • develop a train-the trainer system • identify any additional policy and process supports that are needed 				
Strand 6 – CDE governance/ oversight		<ul style="list-style-type: none"> • Project management charter needs to be created • Data collection platform and processes • Data reporting support • BOCES/regional technical assistance, (regional managers) training • State policies needing to be changed or added to adequately implement this system 	<ul style="list-style-type: none"> • State policies needing to be changed or added to adequately implement this system • Student growth analysis • Data collection platform and processes • Data reporting support 	<ul style="list-style-type: none"> • Student growth analysis 			

Issues to be considered across all strands:

- Implementation
- Relevance, utility, quality
- Fidelity of implementation