

Accuracy is Not Enough: Developing Evaluator Capacity Not Merely Testing for Compliance

ReVision Learning Partnership, LLC

Patrick Flynn, Founder and Executive Director
Martha Bless, Senior Contributing Consultant
Greg McVerry, Senior Contributing Consultant



ReVision
LEARNING PARTNERSHIP

As more and more states begin to implement formal teacher evaluation systems, it is critical for educators to shift their thinking away from compliance models and realize the full potential a growth-minded model for supervision and evaluation (Frase, 200; Stronge, 2010a). Far from mere compliance, the evaluation of educators presents a continuous professional learning opportunity for both teachers and evaluators (An Act Concerning Educational Reform, 2012; Danielson & McGreal, 2000; Feeney, 2007; Frase, 2001; Marshall, 2009; Stronge 2010a, 2010b). As teachers and administrators come to understand the evaluation process from a growth mindset (Dweck, 2006), they will need to develop and utilize protocols that facilitate evidence-based observation practice, and structures to provide both teachers and evaluators with feedback for growth (Anast-May et. al, 2011; Covey, 1991). Through skillful observation, artifact review, and collegial conversation, evaluators have a basis for providing meaningful, actionable feedback to teachers. Research on teacher evaluation confirms that providing “accurate and straight-forward feedback regarding performance... is crucial to improvement” (Frase, 2001, p. 178). Thus, understanding an evaluator’s ability to provide teachers with effective feedback is critical in supporting teacher’s professional growth (Danielson, 2007; Feeney, 2007; Frase, 2001).

Revision Learning Supervisory Continuum

ReVision Learning has established a set of practice standards that describe and promote growth in the discrete skills of an evaluator associated with observing classroom practice and providing feedback and coaching within teacher evaluation models. The ***ReVision Learning Supervisory Continuum™***, a continuous growth rubric, is designed to directly support evaluators in reviewing teacher performance and providing meaningful feedback to teachers.

In the same way that instructional frameworks are used as tools to facilitate teacher growth, the ***ReVision Learning Supervisory Continuum™*** is used to support evaluators’ growth. Supervisors and consultants use the *RVL Continuum* to provide targeted feedback to evaluators as they develop their skills and improve their professional practice. In this way, the *RVL Continuum* is a parallel instrument to the one used to evaluate teacher performance and support their continued growth. This reciprocity in application between the teachers’ and evaluators’ supervision and professional learning generates a level of empathy that not only enables growth to occur but also ensures that it is sustained (Wheatley, 2009).

In addition, the values, criteria, and scales contained within the rubric have been extensively field tested. The ***ReVision Learning Supervisory Continuum™*** was formatively developed through its use in over 2,500 complementary classroom observations and the training and support of over 700 evaluators. This extensive classroom-based development process has not only helped to ensure the validity of the rubric, but it has also led to a set of key guiding principles. These principles build on the belief that teacher evaluation should not be merely an inspection and assessment system, but instead one of continuous cycles of improvement in which both teachers and evaluators advance their professional practices to a level of proficiency that ultimately supports student growth.

The ***ReVision Learning Supervisory Continuum™*** is designed to directly support an evaluator in their work of utilizing instructional frameworks as they observe teachers, collect evidence, rate teacher performance, and write observation reports. The continuum also helps evaluators develop their

skills of providing verbal feedback to teachers. Therefore, the rubric is broken into two domains. Domain One provides standards for evidence-based observation and writing observation reports. Domain Two is designed to support the work of engaging in feedback and coaching conversations with teachers.

Domain One

Domain One has six practice indicators that describe best practices and are designed to support evaluators in collecting and reporting on evidence of teacher performance and practice in alignment with their district instructional rubric. The *RVL Continuum* is used to support initial professional learning of evaluators as well as provide on-going calibration support to districts. Evaluators use the Domain One of the *RVL Continuum* to help them in crafting written feedback reports for teachers.

Domain Two

Domain Two has five practice indicators that describe best practices and are designed to support evaluators as they engage in feedback and coaching conversations about teacher performance and practice in alignment with their district instructional rubric. The *Continuum* is used to support initial professional learning of evaluators as well as provide on-going support to evaluators in developing coaching relationships with teachers. Evaluators use the continuum to help in organizing and delivering feedback to teachers.

Development of the Revision Learning Supervisory Continuum

While the *ReVision Learning Supervisory Continuum™* has two domains, as of now, the cut scores in the Domain One have been tested and validated. It is critical to focus our efforts on building up evaluator capacity in the observation process. Coaching for growth is impossible without evidence that is tied to student achievement.

As a diagnostic tool, the *ReVision Learning Supervisory Continuum™* rates evaluators on six scales within four performance levels: 1- basic, 2- developing, 3-proficient, and 4- exemplary.

Cut Scores are set as a sum of scores on each indicator:

TABLE 1

Basic	Developing	Proficient	Exemplary
6-11	12-16	17-21	22-24

Domain One on the *ReVision Learning Supervisory Continuum™* underwent four phases of development. In the first phase, six guiding principles were developed and operationalized into indicators. In the second phase of development, the instrument self-report data were gathered to help hone definitions of proficiency. In phase three of development, the descriptions of scales were expanded. Efforts focused on developing key levers between each descriptor of scale. This instrument was then piloted as a self-report measure. Through each design phase, the instrument was tested and revised in a field-based, iterative process. In the fourth phase, which is ongoing, the *ReVision Learning* tool is being used in our collaborative, team-based professional learning model, *Collegial Calibrations™*.

Phase I

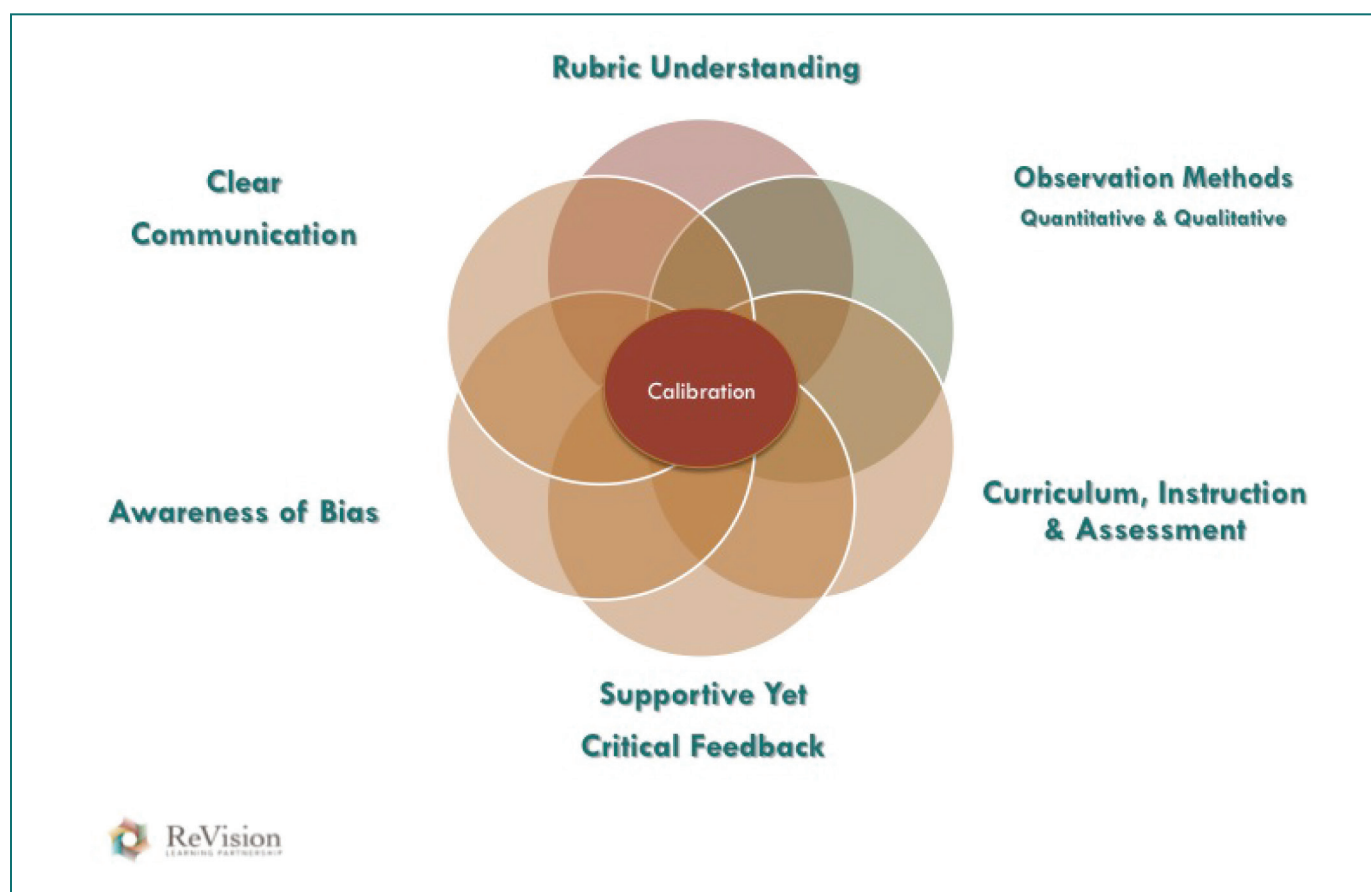
In the first phase, a literature review was conducted (McNamara, 1996) to identify the characteristics of effective evaluators. Data for the review included research literature and observation rubrics that were being used in hundreds of school districts. A synthesis of these sources, plus insights developed during the process of conducting 2,500 complementary observations, led to the identification of six overall traits: rubric understanding; observation methods; curriculum; instruction and assessment; critical feedback, awareness of bias, and clear communication (Figure A). Then operational definitions for the *ReVision Learning Supervisory Continuum™* were developed from these constructs. The focus on the first iteration was to understand and delineate the characteristics of effective evaluators and operationalize them through a lens of evidence based observations.

Rubric Understanding. The greatest step in establishing the credibility and validity of teacher evaluation is ensuring evaluators have a clear understanding of the rubric (Glickman, 2008). In order for evaluators to reach a level of proficiency, they need to understand the organization, criteria, and scales of difference. This includes the identification of particular behaviors that can become key levers in coaching teachers toward growth in their professional practice. Rubric understanding was operationalized and described in the *RVL Continuum* as: *evidence provided is directly tied to the expected indicators of performance.*

Observation Methods. Evaluators must be skilled in collecting and presenting data (Covey, 1991; Danielson & McGreal, 2000; Fullan, 2007a, 2007b; Glickman, 2008). In order to support the ratings assigned to teachers, proficient evaluators gather a preponderance of diverse evidence (Saphier, 1993). This includes more than the scripting of teaching. Direct observations, qualitative and quantitative data, artifacts, and collegial observation methods was operationalized and described as: *evidence cited includes a balance of qualitative and quantitative data. al conversations can all should be present to support the rating assigned to a teacher..*

Curriculum, Instruction, and Assessment. Evaluators need a strong command of curriculum, instruction, and assessment in order to be proficient evaluators (Bambrick-Santoyo, 2012; Danielson, 2007, Frase, 2001). Curriculum, instruction and assessment was operationalized in the *RVL Continuum* as: *evidence cited includes feedback associated with student interaction and learning and has been directly tied back to the implementation of chosen teaching strategies used during lesson.*

FIGURE A



Critical Feedback. Feedback is the heart of learning (Feeney, 2007; Hattie & Temperley, 2007; Hattie, 2009). As Frase (2001) pointed out, “evaluation and supervision can and should be a means of providing feedback and direction for improvement” (p. 178). Proficient evaluators give critical, evidence-based feedback that teachers can use for growth (Bill and Melinda Gates Foundation, 2013). Critical feedback was operationalized and described in the RVL continuum as: *evidence cited includes areas of strengths as well as areas of growth*.

Awareness of Bias. Proficient evaluators need to be self-aware of any potential bias or perceived bias in the teacher evaluation process (Patton, 2002). Respect for cultural and developmental diversity are essential in the supervision and evaluation process. Yet, professional biases evaluators may hold can sometimes appear subtly in evaluation reports. For example, does the report reveal an evaluator favors specific pedagogical methods or content? Awareness of bias was operationalized and described as: *evidence cited is objectively stated and without opinion*.

Clear Communication. Clear communication in written feedback and during coaching sessions is critical to teacher evaluation (Platt et al., 2000; Stronge, 2010a). Proficient evaluators support claims with evidence and use text structure to minimize any perceived inferences required on part of the teacher being evaluated. During coaching sessions, proficient evaluators use the language of the rubric and evidence observed to suggest a plan of growth. Clear communication was operationalized and described as: *evidence provided is effectively communicated through the demonstration of strong written communication skills*.

Phase II

The goal of Phase II was to set basic proficiency levels for each criterion. Two different scales were tested: a five point scale and a four point scale. There were performance levels descriptors at the extremities: beginning and proficient. These were the only descriptors of scale.

Phase II: Content Validity. Initial descriptor language for both basic and proficiency were first reviewed and edited by a team of experts. The panel first reviewed the indicators and wrote initial feedback. Then a focus group with semi-structured interview questions was conducted. Questions included: *How well does this indicator align to the six characteristics of effective observation? What characteristic do you believe this indicator aligns best with? How important is this indicator in the overall quality of an evaluator?*

The results of the focus group led to revisions around increasing the focus on evidence of student learning and outcomes. It was decided to start each indicator with evidence to drive home the concept or evidence-based practice during observations.

Phase II: Piloting. The four point instrument was then piloted with 143 evaluators. The five point instrument was piloted with 46 evaluators. Each instrument was given and administered as a self-report. Average score on the four point instruments was 16.9 with a standard deviation of 4.4. These scores put self-report measures, which are subject to bias of the participants, just below a 17.

Phase II: Initial results revealed that the instrument was not sensitive enough between beginning and proficient, though agreement was high all on six indicators. Facilitator feedback indicated that some criterion were not sensitive enough to the characteristics of effective feedback. The panel felt this was a common key lever indicator in most observation rubrics in use and needed to be reflected in the *RVL Continuum*.

Phase III

Phase III Content Validity. The first decision of Phase III was to revise the indicators. After an initial revision of the *RVL Continuum*, it was brought back to an expanded expert panel. Criteria to join the panel included experience as a district level administrator or having experience completing over fifty formal teacher observations and their corresponding reports. The panel was asked the same set of questions. Final revisions of the descriptors emerged from this focus group. This effort was undertaken to further establish content validity. See Table 2 for a list of revisions.

TABLE 2

A. Evidence provided is directly tied to the expected indicators of performance.	Evidence provided is directly tied to the expected indicators of performance.
B. Evidence cited includes a balance of qualitative and quantitative data.	Qualitative and quantitative evidence cited in feedback is aligned, appropriate and facilitates targeted growth and improvement.
C. Evidence cited includes feedback associated with student interaction and learning and has been directly tied back to the implementation of chosen teaching strategies used during lesson.	Evidence cited is focused on meaningful student engagement and learning and is directly connected to a specific teaching strategy and/or teacher action and impact on students.
D. Evidence cited includes areas of strengths as well as areas of growth.	Feedback contains areas of strengths and areas of growth explicitly connected to the indicator and observed practices/ evidence and are developed based on indicator language and the key levers between ratings.
E. Evidence cited is objectively stated and without opinion.	Evidence cited is objectively stated and without opinion.
F. Evidence provided is effectively communicated through the demonstration of strong written communication skills.	Feedback report serves as a comprehensive learning tool containing clearly articulated evidence-based feedback and explicit connections.

Indicator B was revised to increase the difference between proficiency and beginning. In fact, the item had almost 100% agreement due to the operational definition. The facilitators felt it was a dichotomous item: the report did or did not include both quantitative and qualitative evidence. Thus, a qualitative value was attached to how well this evidence aligned to teacher growth.

An analysis of the written *RVL Continuum* reports and feedback from facilitators indicated a need for greater qualitative difference in the scale of Indicator C. Indicator C was also the only item to consistently be scored with less than 80% agreement after two training sessions. The facilitators wanted evidence of student engagement and learning to have an even greater role in determining the proficiency level of an evaluator.

Evaluators also felt there was not enough qualitative difference in the scale of Indicator D. It too felt like an either/or proposition. Hence, it was revised to add additional variance between beginning and proficiency.

The number of scales was increased to four. The tool was expanded to include four performance level statements per indicator (Beginning, Developing, Proficient, and Exemplary) and language was added to the indicators to increase clarity for both the user and the review.

Phase III Self-Report Testing. The first round of testing involved administering the *Revision Learning Supervisory Continuum™* as a self-assessment before beginning professional development sessions on best practices for teacher evaluation to support growth.

Data were collected through the completion of a self-report Likert scale. Participants also completed two qualitative questions to assess their strengths and weaknesses. The tool was administered using an electronic survey. Each indicator was displayed in table format above a five point Likert-scale. Participants also had paper copies of the entire *RVL Continuum*.

The *Revision Learning Supervisory Continuum™* tool was administered as a self-report tool to 403 evaluators from twelve districts. Half of these districts were included in the analysis of scores. Afterward, data cleaning 139 surveys were collected for analysis.

Phase III Results. Based on the survey instrument, the majority of participants rated themselves below a seventeen ($x = 16.83$ $sd = 2.308$). This averages to 2.83, just below the proficiency scale value of three.

The mean score from this self-report data was used to establish a hypothetical cut score of 17. The data were then arranged by these hypothetical scales to examine the frequencies and distributions.

TABLE 3

School		Evidence cited is directly tied to the expected indicators of performance.	Evidence cited includes a balance of qualitative and quantitative data.	Evidence cited is associated with student interaction and learning and has been directly tied back to the implementation of chosen teaching strategies used during lesson.	Evidence cited includes areas of strengths as well as areas of growth.	Evidence cited is objectively stated and without opinion.	Evidence cited is effectively communicated through the demonstration of strong written communication skills.
branford	Mean	2.64	2.55	2.27	2.82	3.00	2.64
	N	11	11	11	11	11	11
	Std. Deviation	.674	.522	.467	.603	.447	.674
hartford	Mean	3.20	2.96	2.95	3.30	3.32	3.20
	N	74	74	74	74	74	74
	Std. Deviation	.548	.730	.639	.613	.576	.619
norwich	Mean	3.25	2.88	3.00	2.94	3.25	3.00
	N	16	16	16	16	16	16
	Std. Deviation	.577	.500	.632	.680	.577	.516
rs6	Mean	2.73	2.18	2.45	2.64	3.18	2.36
	N	11	11	11	11	11	11
	Std. Deviation	.467	.405	.522	.674	.405	.505
rsd19	Mean	3.00	2.60	2.20	2.80	3.00	3.00
	N	5	5	5	5	5	5
	Std. Deviation	.707	.548	.447	.447	1.000	.707
wolcott	Mean	2.91	2.77	2.86	2.55	3.05	2.73
	N	22	22	22	22	22	22
	Std. Deviation	.750	.685	.710	.596	.653	.767
Total	Mean	3.07	2.81	2.82	3.03	3.22	2.99
	N	139	139	139	139	139	139
	Std. Deviation	.621	.687	.662	.680	.590	.681

The majority of evaluators rated themselves as either in the developing or proficient range. A score of 16 was the most frequent (n=21), followed by a score of 17 (n=20). This distribution places the cut score in the range of evaluators self-reported beliefs about effective teacher evaluation. There were few scores at the extremes of the scale: one basic and two exemplary.

TABLE 4

	Basic 6-11	Developing 12-21					Proficient 17-21					Exemplary 22-24		
Scale	1	48					48					2		
Score	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Frequency	1	2	3	6	16	21	20	12	8	6	2	2	0	0

The final step in Phase III analysis was an examination of the qualitative feedback statements left by participants. They were first examined and coded by the indicator they would best describe. The posts were then examined in terms of being a strength and weakness and compared to the self-reported value. In an overwhelming majority of instances, the evaluators' qualitative reflective statements matched their indicated level of performance. This adds further credence to using 17 as a cut score for proficiency.

Phase IV

The goal of the ongoing phase four is to test the rigor and credibility of the rigor. While the tool is meant as an holistic instrument focused on growth, it can provide reliable data that districts can use as part of their supervision and evaluation of administrators' observation skills.

Phase IV Population. The *Revision Learning Supervisory Continuum™* was next administered as part of an ongoing calibration process in the State of Connecticut. In this ongoing phase, evaluators from 28 districts are being scored by trained facilitators. These facilitators are using the *Revision Learning Supervisory Continuum™* to score administrators. In Phase IV, the tool moved from being a self-report instrument to an evaluation instrument. At the time of analysis, 151 evaluators had been scored against the *Revision Learning Supervisory Continuum™*. Of these 151 participants, 37 have been scored twice, seven have been scored three times, and two evaluators have been scored four times.

Phase IV Procedures. The *Revision Learning Supervisory Continuum™* was administered as part of the ReVision Learning Collegial Calibrations™ process. A group of administrators would complete four formal observations over the course of a school year. Evaluators would attend a professional development session. The Collegial Calibrations™ process includes team-based classroom observations with a focus on continued calibration to the teacher behaviors described in the framework/rubric selected by the district as a part of its teacher evaluation plan. Participants develop proficiency and are given written feedback on their evidence collection methods and techniques and in delivering actionable feedback to teachers. These data are collected using the ReFLECT™ system. This online portal is a ***formative learning educator capacity tool*** designed to support evaluators in learning the key skills and competencies of evidence-based observation practice and providing quality feedback. The entirety of the system is designed to ensure that evaluators are provided with the resources and feedback on their practice to help them become effective teachers of teachers.

All sessions begin with a morning pre-brief, focused on targeted areas of evaluator performance relative to the *RVL Continuum*. Evaluators then participate in 2 to 3 authentic observations or walkthroughs. The team of evaluators meets to formally debrief and discuss their collected evidence from the observation and align it to specific indicators on the district rubric. Evaluators end each session completing an independent written report of at least one classroom observation. These reports are then reviewed by a trained facilitator, and data are recorded in the ReFLECT™ system.

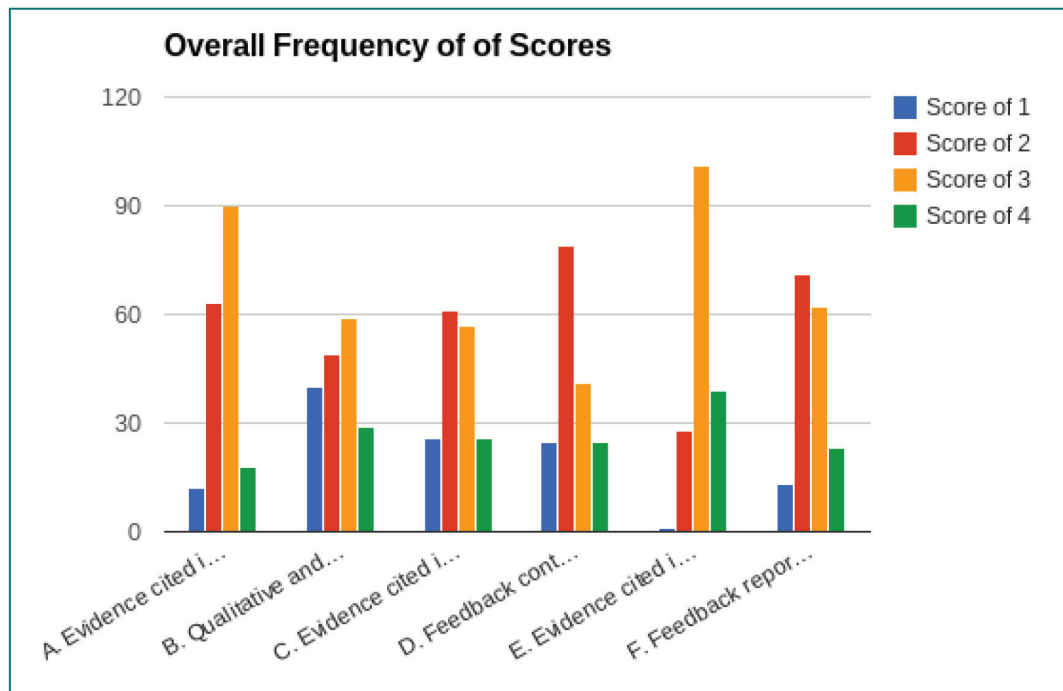
Phase IV Results. It was hypothesized that a score of three on all indicators would mark a clear cut score. First an examination of descriptive statistics was used to determine the average rating and the standard deviation at two different timepoints. Given the low number of respondents who had three evaluations completed, they were not included in the analysis. The overall mean score for the instrument was a score of developing ($x=16$ $SD=4.35$). Table 5 displays the mean scores for all indicators in the *RVL Continuum*.

TABLE 5

Domain One Indicators	X	SD
A. Evidence cited is directly tied to the appropriate indicators of practice and accurately represents the levels of performance.	2.61	0.750
B. Qualitative and quantitative evidence cited in feedback is aligned, appropriate and facilitates targeted growth and improvement.	2.41	1.03
C. Evidence cited is focused on meaningful student engagement and learning and is directly connected to a specific teaching strategy and/or teacher action and impact on students.	2.49	0.93
D. Feedback contains areas of strengths and areas of growth explicitly connected to the indicator and observed practices/ evidence and are developed based on indicator language and the key levers between ratings.	2.38	0.91
E. Evidence cited is objectively stated and without opinion.	3.05	0.65
F. Feedback report serves as a comprehensive learning tool containing clearly articulated evidence-based feedback and explicit connections.	2.57	0.82
Total	16	4.35

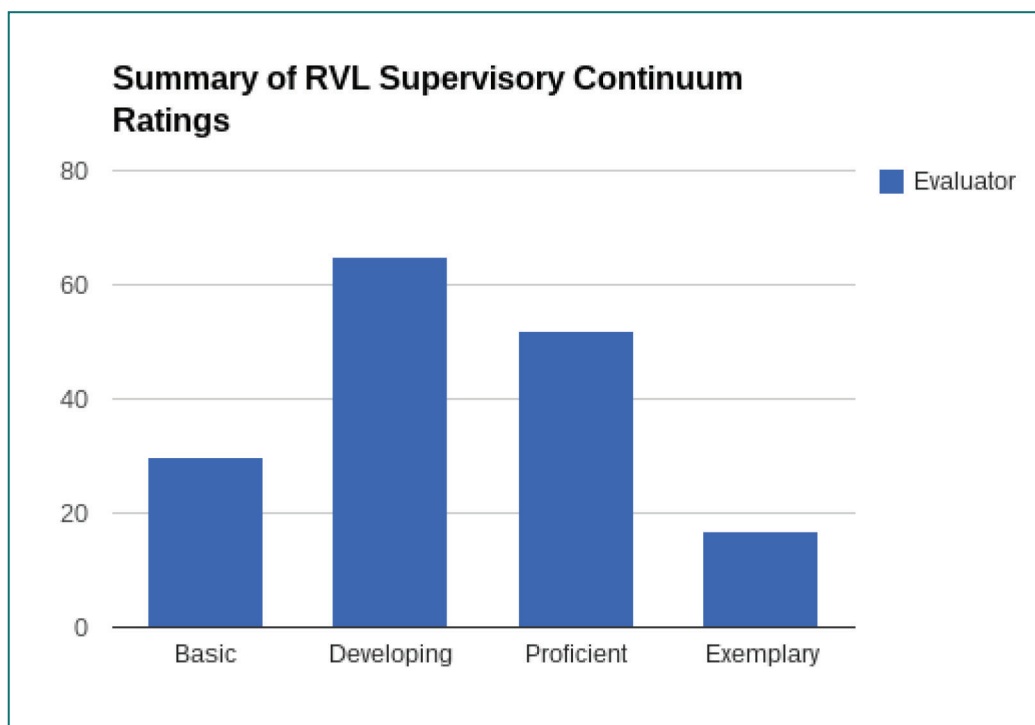
An examination of the frequency of scale ratings mirrors the descriptive statistics. Figure B finds a score of developing to be most frequent on three indicators and a score of proficient to be the most frequent on three other scale. Once again, a cut score of 17 places a level of proficiency just above the current noted capacity of evaluators.

FIGURE B



Using the proposed cut scores from Phase III, the majority of evaluators, when sessions one and two are combined, are rated as developing (see Figure C). Sixty-five (65) of the evaluators were rated as developing. Using a cut score of 17, fifty-two (52) evaluators were rated as proficient.

FIGURE C



An examination of scores on each indicator and an overall score show strong growth on the *RVL Continuum*. This helps to establish the validity of 17 as a cut score. Evaluators score on the average as developing prior to training, and after two professional development sessions, the average evaluator is scored as proficient. Overall, the initial results of Phase IV support the hypothetical scales.

TABLE 6

	A.	B.	C.	D.	E.	F.	Score
Session One	2.54	2.25	2.36	2.243	2.97	2.43	14.80
Session Two	3	3.26	3.33	3.13	3.53	3.26	19.53

Discussion

This report detailed a four-phase process to develop, validate, and test the *ReVision Learning Supervisory Continuum™*. In the first phase, criteria were operationalized from theoretical beliefs about teacher evaluation. In the second phase, levels of proficiency were set. In the third phase, descriptors were written for each level of scale. Then in the third phase, self-report data were collected and a hypothetical cut score of 17 for proficient was determined. In the final and fourth phase, the performance scales were tested. Initial results and the rigorous field tested development of the *ReVision Learning Supervisory Continuum™* establish the validity and reliability of the instrument. This process has led to the creation of a rubric that is sensitive to various qualities of written feedback reports.

The *ReVision Learning Supervisory Continuum™* established content validity throughout the process. Facilitators with the experience of supervising over 700 evaluators and working in 2,500 classrooms consistently iterated on the design through the first three phases. This approach of embedded expertise, combined with design-based research ensured that all aspects of using feedback for teacher growth were included.

Criterion validity was established by the expert panel as well as content validity. Specific attention was taken to ensure that key levers in the descriptors of scale were connected to evidence-based practice. These can be used to coach and measure future progress of evaluators.

An examination of descriptive statistics also helped to establish the reliability of a cut score of 17 for proficient. At this stage in most districts, teacher evaluation is still new. If the *ReVision Learning Supervisory Continuum™* cut score was set lower than 17, then a majority of evaluators would be rated as proficient. Such a cut score would not match evaluators perceived self-efficacy, nor the realities of the classroom. An overall mean score of 16 and a mean score after one round of training of 14.80 supports the reliability of the *ReVision Learning Supervisory Continuum™*.

While the four-phase process established the validity and reliability of the *ReVision Learning Supervisory Continuum™*, work development efforts should continue. To establish inter-rater agreement, an anchor set of feedback reports should be selected from each performance level. These should then be given to outside experts to score.

Criterion validity can also be established by testing how well the rubric correlates with real-life practice. For example, correlations between agreements of individual groups and scores on the *RVL Continuum* could be investigated. If the instrument has high criterion validity, higher scores on the rubric should correlate with fewer disagreements.

The greatest source of established validity for the the *ReVision Learning Supervisory Continuum™* is strong ecological validity. Every aspect of the rubric, the criteria, scales, and descriptors, have undergone extensive field testing. The *ReVision Learning Supervisory Continuum™* was developed and refined in the process of supporting evaluators as they coach teachers for growth. The rubric emerged after 2,500 evaluations with over 700 evaluators. This development process helped to ensure the validity of the rubric.

Conclusion

Districts can feel confident in using the *ReVision Learning Supervisory Continuum™*. The rubric measures the qualities of effective evaluators with established consistency. The validity and reliability of the of the rubric was established through an extensive four-phase development. This process was constantly reviewed by industry experts and extensively field tested. Overall, the *ReVision Learning Supervisory Continuum™* will be a tool to support a growth mindset (Dweck, 2006) in the field of teacher evaluation.



References

- Anast-May, L., Penick, D., Schroyer, R., & Howell, A. (2011). Teacher conferencing and feedback: Necessary but missing. *International Journal of Educational Leadership Preparation* 6(2), 1-7.
- An Act Concerning Educational Reform. CT HB 458. Feb 2012. Retrieved from <http://www.cga.ct.gov/2012/TOB/S/2012SB-00458-R00-SB.htm>References
- Bambrick-Santoyo, P. (2012). *Leverage leadership: A practical guide to building exceptional schools*. San Francisco, CA: Jossey-Bass.
- Bill and Melinda Gates Foundation (2013). *Culminating findings from the MET project's three-year study*. Retrieved from <http://www.metproject.org/reports.php>
- Covey, S. R. (1991). *Principle-centered leadership*. New York, NY: Simon and Schuster.
- Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* 2nd ed. Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C. & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association of Curriculum and Supervision; Princeton, NJ: Educational Testing Service.
- Dweck, C. (2006). *Mindset: The new psychology of success*. New York: Random House.
- Feeney, E. J. (2007, April). Quality feedback: The essential ingredient for teacher success. *The Clearing House*, 191-197.
- Frase, L. E. (2001). Constructive feedback on teaching is missing. *Education* 113, 176-181.
- Fullan, M. (2007a). *The new meaning of educational change* 4th ed.. New York, NY: Teachers College Press.
- Fullan, M. (2007b). *The six secrets of change: What the best leaders do to help their organizations survive and thrive*. San Francisco, CA: Jossey-Bass.
- Glickman, C. (2008). *Leadership for learning: How to help teachers succeed*. Alexandria, VA: ASCD.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research* 77(1), 81-112. doi:10.3102/003465430298487
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.
- James-Ward, C., Fisher, D., Frey, N., & Lapp, D. (2013). *Using data to focus instructional improvement*. Alexandria, VA: ASCD.
- Marshall, K. (2009). *Rethinking teacher supervision and evaluation: How to work smart, build collaboration, and close the achievement gap*. San Francisco, CA: Josey-Bass.
- Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for instruction*. Alexandria, VA: ASCD.
- Marzano, R. J. (2005). *School leadership that works: From research to results*. Alexandria, VA: ASCD.
- McNamara (1996).
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* 3rd ed., pp. 260-326. Thousand Oaks, CA: Sage Publications.
- Platt, A. D., Tripp, C. E., Ogden, W. R., & Fraser, R. G. (2000). *The skillful leader: Confronting mediocre teaching*. Acton, MA: Ready About Press.
- Saphire, J., & Gower, R. (1979/1997). *The skillful teacher: Building your teaching skills*. Acton, MA: Research for Better Teaching.
- Stronge, J. H. (2010a). *Evaluating what good teachers do: Eight research-based standards for assessing teacher excellence*. Larchmont, NY: Eye on Education.
- Stronge, J. H. (2010b). *Effective teachers = student achievement: What the research says*. Larchmont, NY: Eye on Education.
- Wheatley, M. (2009). *Turning to one another: Simple conversations to restore hope to the future*. San Francisco, CA: Berrett-Koehler.

ReVision Learning's mission is to develop services designed to transform the hearts and minds of educational professionals towards superior service of our children and our future world.