

National Staff Development Council Standards Assessment Inventory

Summary Report of Development Process and Psychometric Properties of the Instrument

The National Staff Development Council (NSDC), a non-profit association of over 10,000 members, is deeply committed to ensuring success for all students through the application of high standards for educators' professional learning. The National Staff Development Council's Standards for Professional Development are grounded in research that documents the connection between staff development and student learning. There are twelve standards that are categorized as either context, process or content standards. Each standard is a statement of the staff development expectation and establishes the level of performance to which all organizations can aspire.

NSDC contracted with the Southwest Educational Development Laboratory (SEDL) to develop a new assessment instrument based on the twelve national staff development standards. The instrument would be completed by teachers and other school staff to measure the degree to which a school's professional development program adheres to the standards. This instrument is intended to be used in conjunction with an Innovation Configuration (IC) for each of the standards (Roy & Hord, 2003), thereby providing schools across the nation with a mechanism for diagnosing and aligning their professional development programs with the NSDC standards.

This report discusses the instrument development process and the results from three pilot studies.

Item Construction

The initial set of items for the instrument was developed by SEDL's Evaluation Services (ES) staff members. They generated an initial pool of 360 items (30 items per standard). Items were reviewed, edited, and 10 per standard were selected for the first draft of the assessment instrument. The instrument was next reviewed by four experts, selected by NSDC, who had contributed to the development of the standards. They provided feedback on the clarity of item wording and appropriateness or relevance in representing a particular standard. Based on their feedback the instrument was refined. A focus group of teachers provided further feedback as to whether the items were relevant to their experiences, were stated clearly, and item wording made sense. Again the items were refined. Instrument review and refinement also occurred after each pilot of the assessment instrument resulting in the current 60-item scale (see Appendix A).

Pilot Study Procedure

NSDC staff contacted schools nationwide to solicit volunteers to participate in one of three pilot studies of the assessment instrument. The pilot school samples were to consist of schools ranging from low to high in the adherence of their school's professional

development program to the NSDC standards. Twenty schools for each of the three pilot studies were selected and consented to participate.

NSDC staff also asked individuals with extensive knowledge of the NSDC standards, and who were also knowledgeable about the professional development programs at the various pilot schools, to complete a rating form about the schools. These individuals rated the degree to which they perceived a particular pilot school's professional development program as demonstrating characteristics of the NSDC standards. The rating form (see Appendix B) contained three to four items per standard representing essential aspects of each of the 12 NSDC standards. Items were rated on a three-point scale (1=Low – not present; 2=Medium – inconsistently present; 3=High – consistently present). Scores were added in all categories to obtain a total school score representing either high, medium, or low adherence of the school's professional development program to the NSDC standards from the expert's perspective.

The Standards Assessment Inventory (SAI) was completed by the participating pilot school staffs. Participants responded to the items on a seven-point scale consisting of 0=Never, 1=Almost Never, 2=Seldom, 3=Sometimes, 4=Frequently, 5=Almost Always, and 6=Always. Response categories (seven-point scale) remained the same for all three pilot studies. After the third pilot study was completed, the response scale was changed from a seven- to a five-point scale. The change results from feedback that suggested distinctions among the response categories were too fine-grained. The response categories for the current instrument are 0=Never, 1=Seldom, 2=Sometimes, 3=Frequently, and 4=Always.

In the first pilot, the SAI contained 100 items. The final instrument was to have 60 items. The process of reducing the bank of items from 100 to 60 began with the first pilot study. After being rated by the pilot schools, items were edited, rewritten or discarded based on statistical analyses and experts' feedback. Using the same process, the instrument was reduced to 63 items for the second pilot study, and finally to 60 for the third pilot study.

A packet was sent to each participating school and distributed to the faculty. Each faculty member completed the assessment instrument and placed it in a sealed envelope addressed to SEDL. SEDL's Evaluation Services entered and organized the data for analysis, and conducted various statistical procedures that are described in detail in this report.

Pilot Study Results

The psychometric properties of the SAI were examined in the three pilot studies to investigate the soundness of this instrument for assessing the degree to which schools' professional development programs adhere to the NSDC standards. The process for determining both the reliability, (consistency of measurement) and validity (a true or valid measure of a construct) of the SAI are presented below.

Reliability

Reliability refers to the consistency of measurement. With respect to measuring the degree of adherence of a school’s professional development program to the NSDC standards, we examined the reliability (or consistency) of the SAI for measuring the various components that characterize the standards. Reliability was investigated using Cronbach’s alpha (Cronbach, 1971), which is a measure of the internal consistency of an instrument. Internal consistency assesses the extent that all items in a scale (or all items within subscales) correlate with each other. An alpha coefficient ranges from 0 to 1. Higher coefficients indicate higher levels of instrument consistency. Both overall reliability and subscale reliability were assessed on the SAI. Overall instrument reliability was consistent and high across all three pilot studies achieving an alpha coefficient of .98 in each study (see Table 1). This analysis also showed stability, or consistency for the SAI as a measurement tool across the three pilot studies.

Table 1
Overall Instrument Reliability

	α	N/Items	N/Cases
Pilot #1	.98	100	411
Pilot #2	.98	63	444
Pilot #3	.98	60	297

NOTE: α = coefficient alpha; N = number.

Subscale reliability was also analyzed to examine how well the items in each subscale grouped together and differed from items in other subscales. Alpha coefficients ranged from .71 to .92, signifying good to strong subscale reliability across the three pilot studies (see Table 2). While smaller coefficients are seen for the third pilot study, they still indicate good reliability and may be an effect of the smaller sample size for that pilot study. The analyses also indicate stability, or consistency in measurement across the three pilot studies.

Table 2
Subscale Reliability

	Pilot #1	Pilot #2	Pilot #3
Learning Communities	.89	.84	.79
Leadership	.84	.89	.85
Resources	.76	.83	.71
Data-Driven	.84	.92	.84
Evaluation	.77	.84	.81
Research-Based	.86	.90	.84
Design	.86	.90	.83
Learning	.87	.88	.80
Collaboration	.87	.91	.83
Equity	.88	.86	.77
Quality Teaching	.86	.88	.81
Family Involvement	.88	.85	.76

NOTE: Reported scores = alpha coefficients.

Validity

Examining the validity of an instrument answers the question of whether the instrument is a true measure of what it claims to be measuring. Several types of validity are appropriate to investigating the soundness of the SAI for measuring the adherence of schools' professional development programs to NSDC standards. The three types of validity discussed for this instrument are content, criterion-related, and construct validity.

Content Validity

Content validity refers to how well the items on the SAI represent the practices of good professional development programs as outlined in the NSDC standards. According to Allen & Yen, "Content validity is established through a rational analysis of the content of a test, and its determination is based on individual, subjective judgment" (1979, p.95). As discussed in the above sections on item construction and procedures, the process for refining the SAI item content included rewording and clarifying items to reflect the most accurate description of the NSDC standards according to NSDC experts. Discussions were held between experts and ES instrument developers to ensure that the developers clearly understood the intent of the standards. Teacher-reviewers provided input on how teachers would perceive and interpret the items, and suggested wording and other changes. Continued input of this nature was solicited by the developers during each pilot study of the instrument. Content validity for the SAI was achieved through this process.

Criterion-Related Validity

A second type of validity that was examined was criterion-related validity, which is an appropriate assessment "when scores can be related to a criterion" (Allen & Yen, 1979, p. 97). This analysis assessed the degree to which the SAI responses by school staff compared to an external measure or criterion (expert raters) of the extent to which schools' professional development programs adhered to the NSDC standards. A discriminant function analysis was performed on each set of pilot data to examine evidence of criterion-related validity. Discriminant function analysis is a process that results in the creation of groups – in this case, high or low with respect to adherence of a school's professional development program to the NSDC standards. It then reveals how well a measure categorizes the variable. In this analysis teacher school ratings were categorized and compared to the expert school ratings.

As noted earlier, individuals with knowledge about various pilot schools' professional development programs and the NSDC standards were asked to rate the schools (high, medium, low) on the extent to which the programs demonstrated an alignment with various components of the standards. School scores, as rated by experts, were divided into two groups (high/low) because few experts scored school professional development programs as low. Since most were rated medium or high, the two naturally occurring categories were used and were dichotomized as "high" and "low" for analyses. Ratings were totaled for each school and then schools were classified into two groups divided at the 50th percentile. Expert scores that fell below the 50th percentile were categorized as "low," and those above the 50th percentile as "high."

For purposes of the discriminant function analysis, only cases with a complete set of ratings were used. In pilot study #1, the original sample size was 535. However, 166 cases were excluded from the analysis because of missing data resulting in a total sample for the analysis of 369. For pilot study #2, the original sample size was 444 with 68 cases being excluded due to missing data. In pilot study #3 the original sample size was 364 and 75 cases were dropped due to missing data. Missing data were randomly scattered throughout the groups (high/low) and showed no discernable patterns in non-responses. Table 3 displays the range of expert ratings and the number in the high and low groups for each pilot study.

Table 3

	Range: Expert Rating Total Score	N – High Group	N – Low Group
Pilot #1	69-118	130	239
Pilot #2	42-126	185	191
Pilot #3	41-126	150	139

NOTE: N = number.

Pilot Study #1. The discriminant function analysis showed statistically significant mean differences between the high and low groups (Chi-square: $X^{2(12)} = 53.36$, $p < .001$). Subsequently, the correlation matrix was analyzed to determine which subscales significantly discriminated between the groups.¹ Correlations on nine of the instrument subscales achieved statistical significance at a moderate level, reliably separating the high and low groups: Learning Communities, Leadership, Resources, Research-Based, Design, Learning, Collaboration, Equity, and Quality Teaching. In other words, expert ratings and school staff ratings were in line with each other on these nine subscales as evidenced by statistically significant correlation coefficients. Table 4 displays these findings.

The discriminant function analysis also provided a classification index; that is, it indicated how well group membership (high/low – as rated by experts) was predicted by teachers’ school ratings. Each teacher rating was compared to the expert’s rating for a given school. As shown in Table 5, approximately 88% of the time (210 of 239), teacher ratings correctly classified their school, which is 38% over what would occur merely by chance. However, high group membership was correctly classified only 40% of the time (52 of 130), which is 10% lower than what would occur by chance. This finding may have resulted from a lack of clarity or relevance of some of the items (which had limited refinement at this point), or from the nature of the sample (which was positively skewed, having more low expert ratings than high). However, as will be seen in the analyses of

¹ Correlations range between +1.00 and –1.00. The closer a correlation coefficient is to absolute 1, the stronger the association. Cohen (1988) provided a scale for interpreting correlation coefficients as follows: .01-.03 = small, .03-.05 = moderate, greater than .05 = large. Hopkins (2003) extends the interpretation to include .05-.07 = large, .07-.09 = very large, .09-1 = nearly perfect to perfect.

pilot studies 2 and 3, the classification rate improves. The stability of these findings was checked through a cross-validation sample. This was a random sample of the data that was rerun to replicate/validate the results. The cross-validation confirmed stability at $\pm 4\%$.

Table 4
Group Means and
Correlation Coefficients of Expert Rating and School Staff Ratings

Instrument Subscales	High Group Means	Low Group Means	Correlation Coefficients
<i>Learning Communities</i>	67.13	62.34	.50*
<i>Leadership</i>	43.96	42.04	.32*
<i>Resources</i>	44.34	41.70	.44*
<i>Data-Driven</i>	40.74	39.55	.20
<i>Evaluation</i>	19.97	19.12	.22
<i>Research-Based</i>	44.96	42.04	.43*
<i>Design</i>	32.84	31.00	.38*
<i>Learning</i>	35.34	32.73	.45*
<i>Collaboration</i>	48.43	46.21	.33*
<i>Equity</i>	69.37	66.74	.36*
<i>Quality Teaching</i>	33.31	31.42	.42*
<i>Family Involvement</i>	44.09	44.34	-.03

NOTE: * = $p < .01$.

Table 5
Classification Index

		Predicted Group Membership		Total
		Low	High	
Original Sample		Low	High	
	Low	210 (87.9%)	29 (12.1%)	239 (100%)
	High	78 (60.0%)	52 (40.0%)	130 (100%)
Cross-validated Sample		Low	High	
	Low	206 (86.2%)	33 (13.8%)	239 (100%)
	High	82 (63.1%)	48 (36.9%)	130 (100%)

Pilot Study #2. The discriminant function analysis showed statistically significant mean differences between the high and low groups (Chi-square: $X^{2(12)} = 57.33$, $p < .001$). The correlation matrix was subsequently analyzed to determine which subscales significantly discriminated between the groups. All of the subscales achieved statistical significance indicating that expert ratings and school staff ratings were comparable in designating high or low status on the schools' professional development programs. The three subscales that emerged as the best predictors for distinguishing between high and low

adherence to the standards were Data Driven, Design, and Learning Communities. These results are shown in Table 6.

Table 6
Group Means and
Correlation Coefficients of Expert Rating and School Staff Ratings

Instrument Subscales	High Group Means	Low Group Means	Correlation Coefficients
<i>Learning Communities</i>	30.68	26.09	.80*
<i>Leadership</i>	29.38	25.96	.64*
<i>Resources</i>	31.94	27.95	.73*
<i>Data-Driven</i>	28.08	23.52	.84*
<i>Evaluation</i>	20.69	15.25	.57*
<i>Research-Based</i>	27.61	24.07	.72*
<i>Design</i>	33.64	28.96	.81*
<i>Learning</i>	27.41	23.86	.71*
<i>Collaboration</i>	34.84	30.95	.63*
<i>Equity</i>	30.16	27.07	.70*
<i>Quality Teaching</i>	28.34	24.63	.76*
<i>Family Involvement</i>	25.64	22.52	.62*

NOTE: * = $p < .001$.

The classification index (of group membership as rated by experts) indicated that membership in the low group was correctly classified by teacher ratings approximately 59% of the time (112 of 191), which is 9% over what would occur merely by chance. High group membership was correctly classified approximately 70% of the time (129 of 185), which is 20% higher than what would be expected by chance. The stability of these findings was checked through cross-validation and was confirmed at $\pm 3\%$. These results are displayed in Table 7.

Table 7
Classification Index

		Predicted Group Membership			
		Low	High	Total	
Original Sample	Low	112 (58.6%)	79 (41.4%)	191	(100%)
	High	56 (30.3%)	129 (69.7%)	185	(100%)
Cross-validated Sample	Low	106 (55.5%)	85 (44.5%)	191	(100%)
	High	60 (32.4%)	125 (67.6%)	185	(100%)

Pilot Study #3. The discriminant function analysis showed statistically significant mean differences between the high and low groups (Chi-square: $X^{2(12)} = 43.25$, $p < .001$). The correlation matrix was subsequently analyzed to determine which subscales significantly

discriminated between the groups. Only two of the subscales achieved statistical significance indicating that expert ratings and school staff ratings were comparable in designating high or low status on the schools' professional development programs for those two subscales. The two subscales that emerged as the best predictors for distinguishing between high and low adherence to the standards were Equity and Resources. These results are shown in Table 8.

Table 8
Group Means and
Correlation Coefficients of Expert Rating and School Staff Ratings

Instrument Subscales	High Group Means	Low Group Means	Correlation Coefficients
<i>Learning Communities</i>	18.77	18.47	-.03
<i>Leadership</i>	24.72	24.83	.10
<i>Resources</i>	22.95	23.47	.31*
<i>Data-Driven</i>	20.38	19.03	-.11
<i>Evaluation</i>	18.57	17.38	-.09
<i>Research-Based</i>	20.02	19.49	.03
<i>Design</i>	20.99	20.26	.03
<i>Learning</i>	19.73	18.89	.11
<i>Collaboration</i>	23.00	22.75	-.002
<i>Equity</i>	22.93	24.31	.43**
<i>Quality Teaching</i>	20.13	19.77	.02
<i>Family Involvement</i>	18.44	19.40	.26

NOTE: * = $p < .05$; ** = $p < .01$.

The classification index (of group membership as rated by experts) indicated that membership in the low group was correctly classified by teacher ratings approximately 68% of the time (95 of 139), which is 18% over what would occur merely by chance. High group membership was correctly classified approximately 68% of the time (102 of 150), which is also 18% higher than what would be expected by chance. The stability of these findings was checked through cross-validation and confirmed stability at $\pm 6\%$. These results are displayed in Table 9.

Table 9
Classification Index

		Predicted Group Membership		Total
		Low	High	
Original Sample	Low	95 (68.3%)	44 (31.7%)	139 (100%)
	High	48 (32.0%)	102 (68.0%)	150 (100%)
Cross-validated Sample	Low	86 (61.9%)	53 (38.1%)	139 (100%)
	High	52 (34.7%)	98 (65.3%)	150 (100%)

Summary of Pilot Studies #1, #2, and #3. As seen from the above discussion and tables, the discriminant function analyses detected statistically significant differences between high and low groups in all three pilot studies. These analyses also indicated where school staff ratings and expert ratings were comparable (i.e., which subscales) for both high and low groupings of schools’ demonstration of components of the NSDC standards in their professional development programs. Subscale correlations differed, as did the strength of correlations for subscales across pilot studies (Table 10 displays a comparison of the subscale correlations across the three pilot studies). Since the twelve subscales represent each of the twelve NSDC standards, this finding is not unexpected. Schools often differ in their emphasis on particular aspects of their professional development programs, perhaps due to school/administrator focus and/or differing levels of available resources. It is also noteworthy that most of the correlations in Pilot Study #3 were small with only a few reaching a moderate level. These findings may be a result of the small sample size in this pilot study. However, a 4th pilot study is planned and the results of the analyses will aid in clarifying previous pilot study findings. Nevertheless, the discriminant function analyses on the three sets of pilot data show acceptable support for the criterion-related validity of the SAI.

Table 10
Comparison of Subscale Correlations
for Pilot Studies #1, #2, and #3

Pilot #1		Pilot #2		Pilot #3	
Learning Communities	.50	Data-Driven	.84	Equity	.43
Learning	.45	Design	.81	Resources	.31
Resources	.44	Learning Communities	.80	Family Involvement	.26
Research-Based	.43	Quality Teaching	.76	Learning	.11
Quality Teaching	.42	Resources	.73	Data-Driven	-.11
Design	.38	Research-Based	.72	Leadership	.10
Equity	.36	Learning	.71	Evaluation	-.09
Collaboration	.33	Equity	.70	Learning Communities	-.03
Leadership	.32	Leadership	.64	Design	.03
Evaluation	.22	Collaboration	.63	Research-Based	.03
Data-Driven	.20	Family Involvement	.62	Quality Teaching	.02
Family Involvement	-.03	Evaluation	.57	Collaboration	-.002

Construct Validity

Allen and Yen define construct validity as “the degree to which [a test] measures the theoretical construct or trait that it was designed to measure” (1979, p. 108). The SAI was developed to measure the extent to which schools’ professional development programs adhere to the NSDC standards. The NSDC standards are asserted to be “best practices” for school professional development programs and consist of twelve areas of focus.

The construct validity of the SAI was examined by performing a factor analysis on each set of pilot data to determine if the items separated into twelve distinct “factors,” or areas

of focus. This would be expected if the items well-characterized the standards and if there are indeed twelve independent standards. Using a principal components analysis and varimax rotation procedures, eigenvalues of one or greater were used as the criteria for factor extraction. In the first pilot data set, a seven-factor structure accounted for approximately 54% of the total variance. For the second pilot data set, a six-factor structure emerged accounting for approximately 67% of the total variance. A five-factor structure, accounting for about 59% of the total variance was found in the third sample. These findings suggest that there are only five to seven distinct categories that are represented by the SAI items.

Ideally, the reproduction of a structure pattern of twelve factors would have indicated that there were twelve separate categories/areas of focus that the SAI items represented. However, the data analyses suggest that only five to seven areas exist. These results may indicate that the items need to be revised to better reflect the focus of each of the standards. However, it is more likely that several of the NSDC standards overlap with one another. To investigate this explanation, a careful examination of the factor loading patterns is required. In addition, the wording of the standards should be examined for overlapping descriptions, and a solution might perhaps rest in the consolidation of some of the standards to more succinctly embody the critical elements of “best practices” for school professional development programs.

Conclusion

This report discussed the instrument development process and the results from tests of reliability and validity in three pilot studies. Reliability was investigated using Cronbach’s alpha (Cronbach, 1971) and found to be consistent and high across all three pilot studies for the overall scale, and consistently good for the 12 subscales. These findings indicated that the SAI is a reliable measurement tool.

Several types of validity were examined to assess the soundness of the SAI as a measure of the degree to which schools’ professional development programs demonstrate an alignment with components of the NSDC standards. The SAI demonstrates good content and criterion-related validity. Expert advice during the development process and refinement of item content was solicited to ensure that the instrument would clearly reflect various actions or activities relevant to each standard and the experiences of school staffs.

Criterion-related validity is supported by the results of discriminant function analyses. Teacher ratings and expert ratings of the degree that the components of schools’ professional development programs reflected the NSDC standards were comparable for schools grouped as both low and high in adhering to the standards.

Construct validity for the SAI was not supported by the twelve-factor model suggested by the NSDC standards. Factor analyses indicated a five to seven factor model as most appropriate. These findings suggest that overlap exists within the twelve subscales of the

SAI and that a further examination of the model of the NSDC standards should be undertaken.

While issues regarding construct validity need further investigation, the analyses of the psychometric soundness of the SAI indicate that it is a reliable and valid measure of the degree that schools' professional development programs reflect the actions/activities set out in the NSDC standards.

References

- Allen, M.J., Yen, W. M. (1979). *Introduction to measurement theory*. California Wadsworth.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). New Jersey: Lawrence Erlbaum.
- Cronbach, L.J. (1971). Test validation. In R.L. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington DC: American Council on Education.
- Hopkins, W.G. (2000). *A new view of statistics*. Internet Society for Sport Science: <http://www.sportsci.org/resource/stats/>.
- Roy, P., & Hord, S. (2003). *Moving NSDC's staff development standards into practice: Innovation configurations*. National Staff Development Council and Southwest Educational Development Laboratory.

Appendix A

Standards Assessment Inventory (SAI)

<i>Please mark the responses that most accurately reflect your experiences at your school</i>	<i>Never</i>	<i>Seldom</i>	<i>Sometimes</i>	<i>Frequently</i>	<i>Always</i>
1. Our principal believes teacher learning is essential for achieving our school goals.	0	1	2	3	4
2. Fellow teachers, trainers, facilitators, and/or consultants are available to help us implement new instructional practices at our school.	0	1	2	3	4
3. We design evaluations of our professional development activities prior to the professional development program or set of activities.	0	1	2	3	4
4. Our school uses educational research to select programs.	0	1	2	3	4
5. We have opportunities to practice new skills gained during staff development.	0	1	2	3	4
6. Our faculty learns about effective ways to work together..	0	1	2	3	4
7. Teachers are provided opportunities to gain deep understanding of the subjects they teach.	0	1	2	3	4
8. Teachers are provided opportunities to learn how to involve families in their children's education.	0	1	2	3	4
9. The teachers in my school meet as a whole staff to discuss ways to improve teaching and learning.	0	1	2	3	4
10. Our principal's decisions on school-wide issues and practices are influenced by faculty input.	0	1	2	3	4
11. Teachers at our school have opportunities to learn how to use technology to enhance instruction.	0	1	2	3	4
12. Teachers at our school learn how to use data to assess student learning needs.	0	1	2	3	4
13. We use several sources to evaluate the effectiveness of our professional development on student learning (e.g., classroom observations, teacher surveys, conversations with principals or coaches).	0	1	2	3	4
14. We make decisions about professional development based on research that shows evidence of improved student performance.	0	1	2	3	4
15. At our school teacher learning is supported through a combination of strategies (e.g., workshops, peer coaching, study groups, joint planning of lessons, and examination of student work).	0	1	2	3	4
16. We receive support implementing new skills until they become a natural part of instruction.	0	1	2	3	4
17. The professional development that I participate in models instructional strategies that I will use in my classroom.	0	1	2	3	4
18. Our principal is committed to providing teachers with opportunities to improve instruction (e.g., observations, feedback, collaborating with colleagues).	0	1	2	3	4

<i>Please mark the responses that most accurately reflect your experiences at your school</i>	<i>Never</i>	<i>Seldom</i>	<i>Sometimes</i>	<i>Frequently</i>	<i>Always</i>
19. Substitutes are available to cover our classes when we observe each others' classes or engage in other professional development opportunities.	0	1	2	3	4
20. We set aside time to discuss what we learned from our professional development experiences.	0	1	2	3	4
21. When deciding which school improvement efforts to adopt, we look at evidence of effectiveness of programs in other schools.	0	1	2	3	4
22. We design improvement strategies based on clearly stated outcomes for teacher and student learning.	0	1	2	3	4
23. My school structures time for teachers to work together to enhance student learning.	0	1	2	3	4
24. At our school, we adjust instruction and assessment to meet the needs of diverse learners.	0	1	2	3	4
25. We use research-based instructional strategies.	0	1	2	3	4
26. Teachers at our school determine the effectiveness of our professional development by using data on student improvement.	0	1	2	3	4
27. Our professional development promotes deep understanding of a topic.	0	1	2	3	4
28. Our school's teaching and learning goals depend on staff's ability to work well together.	0	1	2	3	4
29. We observe each other's classroom instruction as one way to improve our teaching.	0	1	2	3	4
30. At our school, evaluations of professional development outcomes are used to plan for professional development choices.	0	1	2	3	4
31. Communicating our school mission and goals to families and community members is a priority.	0	1	2	3	4
32. Beginning teachers have opportunities to work with more experienced teachers at our school.	0	1	2	3	4
33. Teachers show respect for all of the student sub-populations in our school (e.g., poor, minority).	0	1	2	3	4
34. We receive feedback from our colleagues about classroom practices.	0	1	2	3	4
35. In our school we find creative ways to expand human and material resources.	0	1	2	3	4
36. When considering school improvement programs we ask whether the program has resulted in student achievement gains.	0	1	2	3	4
37. Teachers at our school expect high academic achievement for all of our students.	0	1	2	3	4
38. Teacher professional development is part of our school improvement plan.	0	1	2	3	4
39. Teachers use student data to plan professional development programs.	0	1	2	3	4

<i>Please mark the responses that most accurately reflect your experiences at your school</i>	<i>Never</i>	<i>Seldom</i>	<i>Sometimes</i>	<i>Frequently</i>	<i>Always</i>
40. School leaders work with community members to help students achieve academic goals.	0	1	2	3	4
41. The school improvement programs we adopt have been effective with student populations similar to ours.	0	1	2	3	4
42. At my school, teachers learn through a variety of methods (e.g., hands-on activities, discussion, dialogue, writing, demonstrations, practice with feedback, group problem solving).	0	1	2	3	4
43. Our school leaders encourage sharing responsibility to achieve school goals.	0	1	2	3	4
44. We are focused on creating positive relationships between teachers and students.	0	1	2	3	4
45. Our principal fosters a school culture that is focused on instructional improvement.	0	1	2	3	4
46. Teachers use student data when discussing instruction and curriculum.	0	1	2	3	4
47. Our principal models how to build relationships with students' families.	0	1	2	3	4
48. I would use the word, empowering, to describe my principal.	0	1	2	3	4
49. School goals determine how resources are allocated.	0	1	2	3	4
50. Teachers analyze classroom data with each other to improve student learning.	0	1	2	3	4
51. We use students' classroom performance to assess the success of teachers' professional development experiences.	0	1	2	3	4
52. Teachers' prior knowledge and experience are taken into consideration when designing staff development at our school.	0	1	2	3	4
53. At our school, teachers can choose the types of professional development they receive (e.g., study group, action research, observations).	0	1	2	3	4
54. Our school's professional development helps me learn about effective student assessment techniques.	0	1	2	3	4
55. Teachers work with families to help them support students' learning at home.	0	1	2	3	4
56. Teachers examine student work with each other.	0	1	2	3	4
57. When we adopt school improvement initiatives we stay with them long enough to see if changes in instructional practice and student performance occur.	0	1	2	3	4
58. Our principal models effective collaboration.	0	1	2	3	4
59. Teachers receive training on curriculum and instruction for students at different levels of learning.	0	1	2	3	4
60. Our administrators engage teachers in conversations about instruction and student learning.	0	1	2	3	4

Appendix B

**Expert Rating Form
For Pilot Test of NSDC Instrument**

School: _____

District: _____

State: _____

Please use this form to rate the school you are nominating for inclusion in the pilot study according to the degree to which its professional development program adheres to the NSDC standards for professional development. Please place a “3,” “2,” or “1” in the box for each of the standard components. See end of form for scoring procedure.

	3	2	1
NSDC Standards – Staff development that improves the learning of all students:	High (Consistently present)	Medium (Inconsistently present)	Low (not present)
Learning Communities – Organizes adults into learning communities whose goals are aligned with those of the school			
• Staff teams operate with a commitment to continuous improvement and experimentation.			
• Staff teams meet regularly for planning, problem-solving, learning.			
• Staff teams focus on practical ways to improve teaching and learning.			
• Staff observe each other in classrooms.			
Leadership – Requires skillful school and district leaders who guide continuous instructional improvement			
• Leaders establish structures to support ongoing professional development.			
• Leadership responsibilities - distributed among school staff.			
• Leadership provides quality professional development to school staff.			
• Link between staff development and student learning articulated by school leaders.			
Resources – Requires resources to support adult learning and collaboration			
• Professional learning includes the collaboration of colleagues at school (at least 25%).			
• Trainers, coaches, consultants provided to assist teachers.			
• At least 30% of school’s technology budget devoted to teachers’ development in learning to use technology.			
• The school funds substitutes for teachers attending professional development.			
Data-Driven – Uses disaggregated student data to determine adult learning priorities, monitor progress, and help sustain continuous improvement			
• Data on student learning used to select and assess school improvement goals.			
• Staff development efforts are guided by student learning data.			
• Effects of staff professional development assessed by student data.			
• Staff receive opportunities to learn how to use assessment data.			
Evaluation – Uses multiple sources of information to guide improvement and demonstrate its impact			
• School designs evaluation plans for professional development efforts to assess intended outcomes (that uses information in addition to teacher perceptions).			
• Observations of classroom instruction is used to assess teacher professional development efforts.			

	3	2	1
NSDC Standards – Staff development that improves the learning of all students:	High (Consistently present)	Medium (Inconsistently present)	Low (not present)
• Student learning is used to assess teacher professional development efforts.			
• Teacher professional development efforts are assessed by looking at their impact on school culture and other organizational structures.			
Research-Based – Prepares educators to apply research to decision making			
• Staff development is provided to prepare staff to be skillful users of educational research.			
• Staff teams meet to study current research on school improvement efforts.			
• School engages in pilot studies prior to whole-school implementation of new improvement approaches when research evidence lacking.			
Design – Uses learning strategies appropriate to the intended goal			
• School uses a variety of learning strategies for staff development (e.g., formal training workshops, coaches, study groups, professional networks, action research).			
• A variety of follow-up activities throughout the school year are used to supplement major school improvement initiatives.			
• Staff are encouraged to utilize technology (e.g., email, list-servs, distance learning processes) to access and support professional learning.			
Learning – Applies knowledge about human learning and change			
• Learning methods used in school’s professional development mirror methods teachers are expected to use with their students.			
• School provides opportunities for staff to practice new skills and receive feedback on their performance.			
• School provides professional development that addresses diverse learning styles (e.g., hands-on, visuals, group and individual activities).			
• School taps strengths and talents of staff for peer learning opportunities.			
Collaboration – Provides educators with the knowledge and skills to collaborate			
• School’s staff development prepares educators to understand group processes to effectively achieve their goals as members of teams, committees, and departments at their school.			
• The school’s staff development provides opportunities for learning strategies to address and manage group member conflicts.			
• Peer collaboration at the school incorporates the use of technology to encourage and strengthen teaching and leadership practices.			
Equity – Prepares educators to understand and appreciate all students, create safe, orderly, and supportive learning environments, and hold high expectations for their academic achievement			
• School’s staff development prepares educators to adjust instruction and assessment to address diverse learning requirements of their students.			
• School establishes learning environments that communicate high expectations for academic achievement for all students.			
• School prepares staff in practices that convey respect for students, their families, and diverse cultural backgrounds.			

	3	2	1
NSDC Standards – Staff development that improves the learning of all students:	High (Consistently present)	Medium (Inconsistently present)	Low (not present)
Quality Teaching – Deepens educators’ content knowledge, provides them with research-based instructional strategies to assist students in meeting rigorous academic standards, and prepares them to use various types of classroom assessments appropriately			
• School provides staff with opportunities to develop a deep understanding of the content they teach.			
• School provides sustained staff development that incorporates research findings, strategies for teaching, and student progress assessments.			
• School provides staff with the skills to regularly monitor student learning.			
Family Involvement – Provides educators with knowledge and skills to involve families and other stakeholders appropriately			
• School provides educators with knowledge about family and community involvement that would meaningfully benefit students.			
• School’s professional development prepares staff to understand diverse perspectives of community/families and identify common interests.			
• School’s professional development prepares staff to convey an understanding of the challenges of various cultural groups and appropriately involve caregivers/community members in appropriate school activities.			

Column Totals _____ _____ _____

Scoring Procedure:

Total the points given in each of the columns.
Total all points (high + medium + low).

Categorize school as “high,” “medium,” or “low” based on the following scale:

- 0 - 42 points = low
- 43 - 84 points = medium
- 85 - 126 points = high

Overall School Total: _____
(High + Medium + Low)

School Categorization: _____