



LATEST EVIDENCE
on the National Staff Development Council's
STANDARDS ASSESSMENT INVENTORY

February 2009

Research brief

Michael Vaden-Kiernan, PhD
Debra Hughes Jones, PhD
Erin McCann, PhD

Overview

While districts and schools face critical decisions about how best to support teachers, efforts are underway to refine standards and measurement tools to bring evidence to the definition and impact of high-quality teacher professional development. An underlying and related policy question is whether investments in effective teacher professional development increase the odds that schools will meet high-stakes student achievement goals. The timing of efforts to examine how the quality of professional learning affects student learning is critical, given the current increase in federal and state-level interest and participation in initiatives to improve teacher quality.

The National Staff Development Council (NSDC), a private, nonprofit association, has outlined high standards for educator professional learning. One demonstration of NSDC's commitment to the goal of ensuring all schools support and use high standards for professional learning is the organiza-

tion's investment in developing an instrument to assess the alignment of schools' professional development with NSDC standards. In 2003, SEDL developed The Standards Assessment Inventory (SAI) for NSDC, grounded in NSDC's standards which reflect the theoretical foundations and promising practices in school-based professional development literature. SEDL conducted three pilot studies to refine an initial bank of 100 survey items to 60 and began the process of establishing psychometric properties for the 12 standards represented by the instrument. The SAI was intended first to assess the quality of professional learning occurring at a school and to allow use of the findings to guide schools in improving professional development programs. The longer-term goal was to improve teachers' professional learning and thereby have a positive and sustaining effect on student learning.

NSDC is focused on advancing effective policies at the federal, state/provincial, and local levels that emphasize the importance of high-quality teacher professional development in achieving improved stu-

dent outcomes. The relationship between NSDC's standards and the empirical evidence for fidelity of the factors tapped by the SAI and their tie to student learning is particularly relevant to current and future policy research efforts.

This research brief summarizes SEDL's research and development on the SAI to date, with specific emphasis on its psychometric properties, and presents the latest evidence on how teacher professional development contributes to student achievement. After providing background information on the development of the SAI, this brief presents the current findings on whether identifiable factors emerged from teachers' responses on the SAI, mapped onto the NSDC standards of professional development, and whether teacher assessment of professional development using the SAI was related to student learning.

Brief description of the Standards Assessment Inventory¹

SEDL researchers developed the SAI through multiple iterations of an item refinement process where they reduced the initial bank to its current 60 items. They then began the process of establishing psychometric properties for the 12 standards the instrument represents.² The resulting five-item subscales representing each of the 12 standards are designed to collapse into three overarching categories corresponding to context, process, or content stan-

dards and measure the degree to which school-level professional development adheres to the standards as reported by teachers and other school staff. The items comprising the inventory were confirmed for content validity through several iterations of item endorsement by teachers and four experts NSDC selected to develop the standards. Criterion-related validity also was supported by the results of descriptive analyses. Tests of construct validity for the SAI did not support 12 separate standards as theoretically proposed by NSDC's Standards for Staff Development. Factor analyses of teacher responses on the SAI revealed that five to seven factors were represented among the items. The preliminary findings suggested that overlap exists among the 12 subscales, and further analyses were required to clarify how many standards were represented by the survey.

Follow-up study to advance the validity and reliability of the SAI³

SEDL researchers, contracted by NSDC for a follow-up study, designed a set of studies to extend the validation work conducted on the SAI during its initial development. Given the previous difficulties in finding valid constructs using teacher-level factor analyses, researchers re-examined the SAI factor structure at the school-level using a large database drawn from state-level online administration of the measure. SEDL researchers also used total SAI scores and school-level factor scores to assess the contribu-

¹ A fuller description of the instrument, key constructs measured, and evidence to date of its validity and reliability are presented in the attached appendix.

² A summary report of the instrument development process and psychometric properties is available at www.nsd.org/standards/instrumentreport.pdf.

³ More detail regarding the data and methods used for the SEDL follow-up study are described in the full technical report (Vaden-Kiernan, Jones, and McCann, 2008).

tion of teacher professional development to student achievement. A summary of findings is laid out below in the order of the following major research questions:

- Do identifiable school-level factors emerge from teachers' responses to the SAI?
- Does positive assessment of professional development contribute to positive student outcomes?
- Do identifiable school-level factors emerge from teachers' responses to the SAI?

In follow-up analyses to establish the underlying constructs, factor analyses were conducted on the 60 items composing the SAI, with item scores averaged to the school level on a sample collected in spring 2006 of 429 elementary schools in Georgia. The first exploratory factor analysis revealed four factors, but upon close examination, the fourth factor was found to be unstable and therefore dropped. As a result, the final procedure was run on the basis of three factors and the results indicated that three factors were stable statistically and, taken together, explained a substantial amount of the underlying meaning captured by the items.

Factor 1 was explained by a large number of the items, a majority (i.e., 63% or 27 of the 43 items) of which was originally grouped under the standard category titled "process."

Factor 2 loaded on five out of seven items that corresponded to the standard category "context," interpreted as leadership. The remaining two items corresponded to "content" and "process" standard categories, but their content was subsumed under the leadership label.

Factor 3 loaded strongly on three items exclusively derived from the standard category "content," specifically from the "equity" category. Therefore,

the factors are interpreted by the following labels: Factor 1 is process, factor 2 is leadership, and factor 3 is equity.

Factors emerging from the analyses did not map tightly onto theoretically derived standards of professional development. The five-item subscales representing each of the 12 standards did not retain their original grouping when subjected to factor analyses. The subscales were designed to collapse into three overarching standards categories corresponding to context, process, or content; the findings from the factor analyses identified three broader categories that roughly reflected the meaning represented by the original standards categories.

Does positive teacher assessment of professional development contribute to positive student outcomes?

The predictive validity phase of the project first tested the average total school-level SAI score, controlling for school-level demographic measures, in relation to student achievement in reading and English language arts (ELA) on Georgia's Criterion-Referenced Competency Tests (CRCT) in spring 2006. The average score on the total SAI was a positive predictor of grades 1–5 reading and ELA student achievement, except for fifth-grade reading. Predictive validity analyses using a similar database compiled from 102 elementary schools from another state (Alabama) also indicate that the SAI total scale had a significant positive relationship with student academic achievement. A set of initial regression models was run on both Alabama and Georgia data to test all three school-level standards and all 12 school-level standard categories, with mixed results. The school-level average SAI scores proved to be a more stable and predictive measure than the stan-

dards or standard categories measures.

The next phase of the predictive validity research used the three school-level SAI factors named process, leadership, and equity to test their contribution to student achievement in reading and ELA on Georgia's CRCT for spring 2006 using the same methods employed for the first phase of predictive validity testing. Regression models were tested for the relationship between student achievement scores in grades 1–5, controlling for demographic measures (percent of African-American students, teacher

salary), to examine the predictive capacities of the three factors found in the exploratory factor analyses. Process and equity were significant and positive predictors of reading and ELA achievement across all grades, except for fifth-grade reading, to which only equity contributed significantly. The use of factors in the regression models increased explanatory power, refined meaning, and reduced data, thereby improving the predictive validity of the SAI. The items composing the leadership factor require further study for decisions about their predictive capacities.

Discussion of findings

When aggregated to the school level, the SAI total score, a large (43-item) process factor and a small (three-item) equity factor have demonstrated some important psychometric qualities in terms of reliability and predictive validity of academic achievement outcomes for students. Specifically, the SAI total score at the school level demonstrated good predictive validity with academic achievement outcomes for students across two states (531 schools). The SAI process and equity factors at the school level demonstrated even better predictive validity in terms of academic achievement for students in Georgia (429 schools).

Taken together, these findings are promising and provide researchers and educators more confidence that the SAI measure, when aggregated as a total school-level summary variable, is a reliable and valid measure that has demonstrated significant associations with student academic achievement. The findings to date support a set of two factors as the most

useful elements beyond the total SAI score in terms of predictive validity. The current findings support the notion that schools and staff working on the issues covered by the SAI in the aggregate may have a significant association with effect on academic achievement outcomes for students. Specifically, the findings presented here support teachers working on the issues covered by the SAI more in an aggregate sense than in any specific sense related to the 12 standards areas.

Overall, the findings also indicate that the SAI still needs a good deal of attention. For example, it is unclear how many standards are measured reliably with the current items. The study found significant relationships between the process and equity factors and student achievement in reading and ELA. This raises the question of whether the two factors, which could be interpreted as broad standards categories, are capturing sufficient measures of the quality of teacher professional learning that occurs at the schools. Taken together, these findings are promising and encourage additional analyses to explore the po-

tential of the SAI measuring a set of educationally meaningful constructs at the school level and their relationship to academic achievement of students in predictive models. The current findings were derived from a large database but included elementary schools in one or two states (Georgia and Alabama). Additional analyses are needed to replicate these findings in other states and other school types before making significant changes to the instrument.

LIMITATIONS AND CAVEATS

As with every study, this study has a number of important limitations to keep in mind when interpreting the findings. The study does not support causal interpretations between the SAI and academic achievement. The findings are based on a cross-sectional correlational study and are, therefore, to be interpreted cautiously. In particular, there is no evidence from this study that would support the notion that the issues measured by the SAI lead to direct changes in academic achievement outcomes for students. There is a strong association or relationship between the SAI and academic achievement outcomes collected at approximately the same points in time. There may be other, third variables or alternative hypotheses which cannot be addressed in the current study that would explain these results. For example, it may be that positive traits cluster together and high-quality professional development happens to be offered in the same schools where students have high academic achievement outcomes due to factors other than the SAI, such as resources or characteristics of teachers and students. A series of additional research studies are needed to address the issue of whether the perceptions that the SAI captures lead to changes in student academic achievement.

The current findings were derived from large databases with very large sample sizes, but included only elementary schools in just two states in the Southeast region of the United States (Georgia and Alabama), which may limit the generalizability of the findings. Findings need to be replicated in other states or regions and other school types to address this concern. Finally, reliance on state archival data tends to limit data access, timeliness, and quality. The use of archival data in this study has provided a good first look at the qualities of the SAI, but additional prospective research studies are needed that collect data using multiple methods (i.e., survey, observations, interviews, focus groups) and address multiple dimensions of professional development beyond teacher perceptions (i.e., attitudes, beliefs, behaviors) to attain a broader picture of professional development in the field.

IMPLICATIONS

Study findings indicate that developing and validating an instrument takes time and should be done incrementally. The current findings suggest that continued research, perhaps using responses from ongoing online administrations of the SAI, would lend sufficient amounts of data to confirm and replicate construct and external validity findings. Specific implications for policy, practice, and research are discussed below.

Policy. The study does appear to validate that the SAI (particularly the school-level total score and the process and equity factors) is a reliable measure at the school level and captures important and meaningful aspects of professional development that are related to academic achievement outcomes. Results imply that the overall emphasis of NSDC's Standards

for Staff Development may be more important than any single standard. This work also contributes to other policy research efforts to address the increasing need for solid information and evidence regarding the potential contributions of teacher professional development to student achievement and increased resources and supports for such efforts.

Practice. The study provides further support to the notion that teachers and schools that use the SAI to inform and discuss the professional development they receive are spending time and energy on a potentially worthwhile activity. As discussed above, the study provides only correlational evidence that should not be mistaken for causal, but it provides more evidence than was previously available regarding the SAI's predictive validity related to student academic achievement. This study also supports the increasing need for valid and reliable instruments to guide improvements in school professional development programs. It also supports administrators' need for tools to assist in effective investments in teachers that may relate to improvements in student achievement.

Research. The study provides some helpful insights into the potential uses for the SAI in future research endeavors. While establishing a better sense of the psychometric properties and value of the instrument, the study also raises important implications for using the instrument in future research. The instrument is quite lengthy, and the study has established that potentially the most helpful information for research purposes is at the school level, not the teacher level. Therefore, when researchers assess the instrument's overall value in terms of research, they should consider the effort that the instrument requires of teachers and schools and the burden it places on them. This project also suggests the value of developing a consortium of researchers that would devote resources to sharing data and designing follow-up research studies to improve the empirical basis and use of the SAI. Continued work on the psychometric properties of the SAI would improve its use in more rigorous research designs and the interpretability of findings.

Appendix

MEASURE PROFILE NATIONAL STAFF DEVELOPMENT COUNCIL STANDARDS ASSESSMENT INVENTORY (SAI)

BACKGROUND INFORMATION

Author/source

Source: SEDL Evaluation Services. (2003, December). *National Staff Development Council Standards Assessment Inventory: Summary report of development process and psychometric properties of the instrument.* Austin, TX: Author.

Publisher: National Staff Development Council
www.nsd.org/standards/instrumentreport.pdf

Purpose of measure

The SAI, developed by SEDL for the National Staff Development Council (NSDC), was intended to assess the degree to which schools' professional development programs adhere to standards representing best practices for staff learning. Items were

developed to measure the 12 standards of NSDC's Standards for Staff Development. The resulting 12 subscales, designed to collapse into three overarching categories corresponding to context, process, or content standards, measure the degree to which school-level professional development adheres to the standards as reported by teachers and other school staff. When the instrument is used in conjunction with an Innovation Configuration (Roy & Hord, 2003) for each standard, it assists school leaders in their efforts to measure and align their professional development programs with NSDC's Standards.

Target population for instrument

Teachers and other school staff

Key constructs of measure

The SAI includes 12 standards or subscales, measured with five questions each, which comprise the 60-question survey. The subscales are subsumed under three broad categories: context, process, and content standards.

Context Standards

- *Learning communities.* Items tap teachers meeting as a community to discuss teaching improvements, observing other teachers' classrooms, mentoring new teachers, providing collegial feedback on classroom practices, and examining student work.
- *Leadership.* Items tap principals' beliefs about teacher learning, teachers' influence on principals' decisions, principals' commitment to teachers' opportunities to improve instruction, principals' ability to foster a culture of instructional improvement, and whether the principal is perceived as empowering staff.
- *Resources.* Items tap resources available to implement new instructional practices, opportunities to learn new technologies for instruction, availability of substitutes to cover teachers who are engaged in professional development, creativity used to expand human and material resources, and whether school goals determine resource allocations.

Process Standards

- *Data-driven.* Items tap teachers' knowledge on using student improvement data to assess student needs, evaluating the effectiveness of professional development programs, planning for professional development and curriculum, and analyzing improvements in student learning with other teachers.
- *Evaluation.* Items tap the design of evaluation prior to professional development, the number of sources used to evaluate professional development experiences, the use of professional development outcomes to plan for future

choices, and the use of student performance to evaluate professional development.

- *Research-based.* Items tap the use of educational research to select professional development programs, the use of research on effectiveness of school improvement efforts to decide on strategies, evidence of improvement programs' effectiveness for student achievement gains, and the effectiveness of improvement programs in schools with similar student populations.
- *Design.* Items tap teacher learning through a variety of strategies, the design of improvement strategies based on clear outcomes for teacher and student learning, teacher learning as part of the school improvement plan, consideration of teachers' prior knowledge and experience when designing staff development, and commitment to sufficient time with improvement initiatives to result in changes in instructional practice and student performance.
- *Learning.* Items tap opportunities to practice new skills, support for implementing new skills, promotion of deep understanding of a topic, learning through a variety of methods, and teachers' choice of the type of professional development they receive.
- *Collaboration.* Items tap learning about effective ways to work together, structuring time for teachers to work together to enhance student learning, teaching and learning goals dependence on staff's ability to work together, leaders encouraging sharing responsibility to achieve school goals, and principal modeling effective collaboration.

Content Standards

- *Equity.* Items tap adjusting instruction and as-

assessments to meet the needs of diverse learners, showing respect for all of the student subpopulations, expecting high academic achievement for all students, creating positive relationships between teachers and students, and teachers receiving training on curriculum and instruction for students at different learning levels.

- *Quality teaching.* Items tap teachers having opportunities to gain deep understanding of subjects, professional development models, instructional strategies to be used in classroom, teachers' use of research-based instructional strategies, professional development teaching effective student assessment techniques, and school administrators engaging teachers in conversations about instruction and student learning.
- *Family involvement.* Items tap provision of opportunities to learn how to involve families in children's education, prioritizing the communication of the school's mission and goals to families and community, work done by school leaders with community members to help students achieve academic goals, the principal as a model of building relationships with students' families, and teachers' work with families to help them support student learning at home.

Norming of measure (Criterion or norm-referenced)

No norming information is currently available.

Comments

The SAI is a relatively new measure. Researchers continue to contribute to a growing body of evidence that supports the SAI's integrity as a measure of professional development that maps onto NSDC's

Standards. Findings to date encourage additional analyses to establish the SAI as a measure of valid constructs at the school level and to replicate preliminary findings regarding the relationship to student achievement in predictive models. The findings to date were derived from a large database but included schools in only two states. Additional analyses are needed to replicate these findings with other large datasets before making significant changes to the instrument.

ADMINISTRATION OF MEASURE

Who is the respondent to the measure?

The SAI's primary target is teachers.

Who administers the measure?

The online version of the survey allows all staff members to respond to the 60-question survey. Schools receive an access code that allows staff members to respond to the online survey through access to a web site for a designated period of time. At the conclusion of that time period, a number of summary reports are generated for the school. School staff in a designated setting may self-administer the print version of the inventory.

Setting

The instrument is independent and self-administered.

Time needed and cost

Time: Time estimates vary given that teachers may return to online sessions to complete the survey; print version completion takes 20 to 30 minutes on average.

Cost: Online and print versions of the inventory

may be purchased from NSDC. See www.nsd.org/standards/sai.cfm for more information.

FUNCTIONING OF MEASURE

Reliability information from instrument development pilot studies

Internal consistency reliability

- Cronbach's alphas for overall instrument reliability were consistent and high across all three pilot studies ($\alpha = .98$). Reliability estimates for all 12 subscales ranged from good to strong across tests ($\alpha = .71$ to $.98$).

Validity information from instrument development pilot studies

Content validity

- Expert and school staff ratings supported the content validity for the 12 subscales.

Criterion validity

- Teacher 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

- Tests of construct validity revealed that nine subscales were moderately supported: learning communities, leadership, resources, research-based, design, learning, collaboration, equity, and quality teaching.
- Tests of construct validity for the SAI did not statistically support the 12-factor model theoretically defined by the NSDC standards. Prelimi-

nary factor analyses indicated support for a five-to seven-factor model.

Construct reliability

- Reliability was found to be consistent and high across all three pilot studies for the overall scale, and consistently good for the 12 subscales.

Validity and reliability information from follow-up study⁴

Content validity

- Expert reviewers suggested a number of structural changes as well as more superficial editorial changes to the SAI to improve the content validity of the instrument. To date, SEDL has recommended that NSDC consider adding items to capture demographic information on the SAI.

Construct validity

- Exploratory factor analyses of the SAI at the school level in more than 400 elementary schools in Georgia indicate a three-factor solution capturing high-quality professional development (process), leadership, and equity factors, indicating the SAI may capture some stable school-level constructs worth exploring further.
- The variance explained by the overall three-factor solution was 79.1%; the variance explained by factor 1 was 73.92%; by factor 2, 3.22%; and by factor 3, 1.96%. Cronbach alphas were .992 (43 items) for factor 1; .968 (7 items) for factor 2; and .911 (3 items) for factor 3.

Predictive validity

- Predictive validity analyses using the total SAI found support for the importance of teacher-reported perceptions of professional development

⁴ (Vaden-Kiernan, Jones, and McCann, 2008)

as measured by the school-level average on the SAI in predicting student achievement outcomes in Georgia and Alabama elementary schools.

Total SAI scores were statistically significant predictors of student reading scores, even after controlling for student and teacher characteristics.

- Predictive validity analyses using the SAI factors in the exploratory factor analyses showed even stronger support for school-level factors as significant predictors of student academic achievement outcomes in Georgia elementary schools. Specifically, high-quality professional development (process) and equity were significant positive predictors of student reading scores. The use of SAI factors in predictive models increased the explanatory power of the SAI and the predictive validity of the SAI.
- Overall regression analyses found preliminary

support for the positive contribution of the school-level average on the SAI to student achievement.

- The models tested in the regressions explained a respectable amount of variance in student achievement (19% to 55%).

References

Roy, P. & Hord, S. (2003). *Moving NSDC's staff development standards into practice: Innovation configurations*. National Staff Development Council and Southwest Educational Development Laboratory.

Vaden-Kiernan, M., Jones, D.H., & McCann, E. (2008). National Staff Development Council: *Revisions and additional psychometric data analyses of the standards assessment inventory*. Austin, TX: SEDL Research and Evaluation.