Macrothink Institute™

Journal of Educational Issues
ISSN 2377-2263
2022, Vol. 8, No. 1

Teacher Engagement Scale for Professional Development

Cara L. Kelly (Corresponding author)

Department of Human Development and Family Sciences, University of Delaware

111 Alison Hall West, Newark, Delaware 19716

E-mail: clkelly@udel.edu

Laura L. Brock

Department of Teacher Education, College of Charleston

66 George Street, Charleston, South Carolina 29424

E-mail: brockll@cofc.edu

Julie Dingle Swanson

Department of Teacher Education, College of Charleston

66 George Street, Charleston, South Carolina 29424

E-mail: SwansonJ@cofc.edu

Lara Walker Russell

Charleston County School District

5109 W. Enterprise Street, North Charleston, South Carolina 29405

E-mail: lara_russell@charleston.k12.sc.us

Received: March 9, 2022 Accepted: April 2, 2022 Published: April 21, 2022

doi:10.5296/jei.v8i1.19636 URL: https://doi.org/10.5296/jei.v8i1.19636

Abstract

Teacher professional development aims to bolster effective instructional practices that enhance

student learning. The development of the Sustained Investment and Teacher Engagement Scale for Professional Development (SITES-PD) grew out of a need to understand how teachers respond to opportunities for professional learning in ways that contribute to skill improvement. One hundred and eight elementary teachers participated in a sustained professional development coaching intervention. Prior to and following the intervention, data were collected from multiple sources (*e.g.*, teacher report, blind classroom observations). For the development of this instrument, coaches rated teacher engagement with opportunities for professional learning. Exploratory factor analysis reveals a one factor solution and that a summed composite of eight items is appropriate. Reliability and validity results suggest the SITES-PD instrument may be a useful tool for investigating the underlying mechanisms that mediate the efficacy of teacher professional development interventions.

Keywords: professional development, teachers, validity, reliability

1. Introduction

Teacher professional development is critical for school improvement and student achievement (Desimone, 2009; Guskey, 2002; Jones, 2017). Significant resources, both time and money, are spent on professional development each year (Desimone, 2009). While researchers have identified certain features of effective professional development models affecting teacher change (Darling-Hammond et al., 2017), little research has examined the measures used to assess professional development interventions. Further, many professional development efforts lack a robust assessment component (Earley & Porritt, 2014). Ensuring the effectiveness of professional development is imperative. In addition, engaging teachers in professional development is a crucial component of professional development (Garet et al., 2001). Additional research and resources are needed to understand the components of professional development that impact teachers' willingness to engage in professional development. The current study fills a critical gap in the literature by developing an assessment of teacher engagement for professional development interventions. The development of the proposed measure will enable researchers and practitioners to identify teacher engagement behaviors that contribute to more effective classroom practice.

Desimone's (2009) seminal work outlined five critical features of effective professional development: content focus, active learning, coherence, sustained attention, and collective participation. Additional studies have highlighted the importance of sustained attention in professional development interventions (Corcoran, 1995; Garet et al., 2001; Whitworth & Chiu, 2015). Professional development programs that are longer in duration can result in positive changes in teacher practice compared to shorter professional development programs or workshops (Whitworth & Chiu, 2015). Research suggests it is imperative to implement professional development that is school-based and job embedded to promote engagement among teachers (Fairman et al., 2020). While opportunities for teachers to engage in active learning are a hallmark of effective professional development (Garet et al., 2001), there is a lack of research about the assessment of the effectiveness of professional development. The few studies that have assessed specific aspects of professional development have limitations, such as focus on short-term changes in classroom practice, lack of follow-up, and reliance on

teacher-report data (Palermo & Thomson, 2019; Palmer & Noltemeyer, 2019).

In order to support new behaviors that foster increased student learning, effective professional development must be ongoing and sustained over time (Yoon et al., 2007). The impact of professional development has largely focused on teachers' satisfaction; teachers' learning, use of new practices in the classroom environment, and student outcomes have not been the focus of extant literature (Rhodes et al., 2004; O'Sullivan, 2011). Recent studies have examined the impacts of professional development on teacher learning (Liang et al., 2020; Palermo & Thomson, 2019). Yet, additional research is needed to assess teacher engagement in professional development in order to understand the components of professional development that are most beneficial to teachers.

Desimone and Garet (2015) found that changing teachers' classroom behavior was easier than improving content knowledge or inquiry-oriented instruction techniques. In addition, their results suggest that teachers vary in response to the same professional development, and leadership plays a key role in supporting and encouraging teachers to implement strategies and ideas they learned during professional development. A recent study suggests there is an association between perception of usefulness of professional development and positive changes to classroom instruction (Palermo & Thomson, 2019). Taken together, these studies highlight the need for professional development that impacts changes in classroom practices. However, assessing the effectiveness of professional development continues to be an understudied topic.

1.1 The Current Study

The aim of the current study was to develop an instrument that could help researchers identify teacher engagement behaviors with sustained professional development that lead to the greatest improvements in effective classroom practice and teacher beliefs about teaching and learning. The present study attempts to develop and validate a measure, Sustained Investment and Teacher Engagement Scale for Professional Development (SITES-PD). The goals of this study were to (1) assess content validity, (2) assess construct validity, and (3) assess predictive validity of the SITE-PD measure. We hypothesize that teacher engagement will serve as a stronger predictor of professional development outcomes than mere number of hours of participation in professional development.

2. Method

2.1 Sample

Across six elementary schools, 108 teachers were enrolled in the study. Teachers self-reported teacher and classroom demographic characteristics. Descriptive characteristics of teachers and students are presented in Table 1. Teachers in the study were mostly female (98%). The majority of teachers identified as White (64%) or Black/African-American (20%). Many teachers in the study had a master's degree (41%); all teachers had at least a bachelor's degree (43%). Teachers' teaching experience ranged from zero to 45 years.

Table 1. Descriptive characteristics of teachers and students

	N	%	*M*	*SD*	*Min*	*Max.*
Teacher Demographic Variables						
Age	87		33.76	9.54	22	64
Gender	88					
Female	86	98%				
Male	2	2%				
Ethnicity	88					
White/Caucasian	66	75%				
Black/African-American	20	23%				
Other/multiple ethnicity	2	2%				
Education	88					
Bachelor's degree	46	52%				
Master of Arts in Teaching	9	10%				
Master of Education	31	35%				
Certification	88					
Early childhood	55	62%				
Elementary	49	55%				
Literacy	9	10%				
English Language Learner	2	2%				
Number of years teaching Grade level	88		4.99	4.64	0	21
Number of years teaching current school	88		5.45	5.57	0	44
Classroom Characteristics						
Number of students per class	87		19.79	4.04	8	42
Students receiving special education services	84		1.81	1.39	0	7
Students identified as gifted and talented	78		1.23	1.66	0	7
English Language Learner students	84		2.37	2.16	0	12
White/Caucasian students	86		4.88	3.32	0	12
Black/African-American students	86		10.66	3.92	0	27
Hispanic/Latino students	82		2.93	2.62	0	13
Multiracial	64		0.84	0.98		4
Other	51		0.54	0.95	0	6

Note. Several teachers held more than one certification, hence frequency count outweighs number of teachers.

2.2 Procedures

Researchers collaborated with school partners to create "Talent Development Academies" in rural Title I elementary schools with the aim of developing and identifying talent in traditionally underrepresented culturally and linguistically diverse populations. Highly trained teacher coaches held professional development (PD) sessions at specific intervals throughout the year; the teacher coaches also offered voluntary individualized coaching for teachers at each of the elementary schools. Coaches offered demonstrations, co-planning, and co-teaching as a part of their coaching repertoire.

Initial data collection (pre-test) occurred at the end of the academic year prior to initial professional development sessions and served as baseline data. The same data were collected at the end of the academic year when the intervention had been fully implemented (posttest). Data were collected from three sources. First, coaches recorded teachers' participation in mandatory and voluntary PD by number of hours and type of interaction. Coaches were additionally asked to rate teachers' engagement with professional development at the end of each academic year, using the SITES-PD form found in Appendix A. Second, teachers completed a series of questionnaires that gauged teacher perceptions of *Teacher Self-Efficacy* (Tschannen-Moran, & Woolfolk Hoy, 2001), *Teacher Mindset* (Gutshall, 2014), and *Teacher's Observation of Potential in Students* (Coleman et al., 2010). Third, trained and certified observers, blind to the conditions and objectives of the study, visited classrooms to assess the effectiveness of instructional interactions that occurred within the classroom using the *Classroom Assessment Scoring System* (CLASS; Pianta et al., 2008) and the *Classroom Observation Scale-Revised* (VanTassel-Baska et al., 2007).

2.3 Measures

2.3.1 Sustained Investment and Teacher Engagement with Professional Development

Items on the SITES-PD measure assessed purposefulness of teachers' communication, teachers' level of comfort with learning, self-reliance, flexible thinking, perception of students as learners, teachers' views of rules and authority, teachers' views of themselves as professionals, and teachers' assumption of a leadership role. The SITES-PD measure assessed change over time on each of these items by comparing the pre-test scores (assessed at baseline) to posttest scores (assessed after the professional development intervention). The measure was specifically created to assess the effectiveness and usefulness of the professional development intervention.

Throughout the year, two highly trained PD coaches documented all face-to-face and virtual teacher-coach interactions and assigned a value to capture the depth of interaction on a scale of 1 (rote) to 3 (reflective). An example of a rote interaction would be "Can you get me materials for X lesson?" and an example of reflective communication would be "I tried X lesson. I noticed a group of students seemed apprehensive. How can I adapt the next lesson to elicit more brainstorming?" If teachers did not seek out any additional coaching support throughout the academic year, they were assigned a 0. Researchers developed an 8-item global rating survey that implementers completed for each teacher at the end of the academic year. The

SITES-PD also had a scale of 1 (rote) to 3 (reflective). Each teacher was rated once at the end of each academic year. See Appendix A for the full version of the SITES-PD measure.

2.3.2 Classroom Assessment Scoring System (CLASS)

The Classroom Assessment Scoring System (CLASS; Pianta et al., 2008) was used to measure the effectiveness of teachers' interactions with children. The CLASS, a standardized observational instrument, measures the effectiveness of classroom social interactions along 10 dimensions. Dimensions are scored on a Likert scale from 1 (low) to 7 (high). The 10 dimensions are aggregated into three domains: Emotional Support, Classroom Organization, and Instructional Support. The Emotional Support domain, which includes positive climate, negative climate (reverse coded), teacher sensitivity, and regard for student perspective dimensions, measures the degree of warmth present in the classroom and the degree to which the teacher is sensitive to the needs and perspectives of students ($\alpha = .70$). The Classroom Organization domain, which includes behavior management, productivity, and instructional learning formats, measures managerial behaviors and interactions that provide structure and consistency to the learning environment ($\alpha = .74$). The Instructional Support domain, which includes concept development, quality of feedback, and language modeling, measures the degree to which higher-order learning occurs and the presence of teacher behaviors that promote student understanding and learning through reciprocal interactions ($\alpha = .88$).

Observers who were CLASS certified and blind to the goals of the study conducted classroom observations. The observers coordinated with teachers to conduct classroom observations at a time that was convenient for the teacher. Blind observers conducted two observation cycles in each classroom; the observation cycles lasted for twenty minutes with ten minutes of coding following the observation cycle. Thus, each teacher was observed for a total of 40 minutes. To ensure inter-rater reliability throughout the study, twenty percent of observations were dual-coded. During dual-coded observations, two observers independently coded the CLASS protocol following the observation cycles and later compared their scores to ensure reliability between coders. Research suggests there is moderate stability in CLASS observation ratings after two cycles (Curby et al., 2011).

2.3.3 Classroom Observation Scale-Revised (COS-R)

To gauge the effectiveness of general and differentiated teaching behaviors, certified and blind observers used the COS-R, a 50-item observation instrument, to assess classroom teaching. This instrument has strong technical adequacy and is highly reliable ($\alpha = .91$ to $.93$), with the subscale reliability for all of the clusters averaged above .70 (VanTassel-Baska et al., 2007). Observers used the COS-R to examine teaching in teacher instructional behaviors (23 items; $\alpha = .89$) and student engagement with teacher instruction (23 items; $\alpha = .89$) across a range of sub-categories, including general teaching/student behaviors (curriculum planning and delivery) and differentiated teaching behaviors (accommodations for individual differences, problem solving, critical thinking strategies, creative thinking strategies, and research strategies). A rating scale of effectiveness for each item is used in the COS-R, with 3 being effective, 2 being somewhat effective, and 1 being ineffective.

The Principal Investigator trained two COS-R observers, and 20% of observations were dual-coded to ensure inter-rater reliability. The COS-R provides evidence of the degree of differentiated strategies teachers are employing and helps project staff see where growth in desired teacher behaviors is occurring. Teachers were rated once per assessment window after scheduled observations that varied in length from 30-50 minutes (average 45 minutes). The COS-R and CLASS observations occurred simultaneously.

2.3.4 Teacher Self-efficacy

Self-efficacy describes a belief in one's ability to perform at an expected level. Teacher self-efficacy more specifically addresses teachers' beliefs in their ability to reach and teach a variety of learners in their classroom. Teachers offered their perspective of their own teaching efficacy across three domains: (a) Efficacy in Student Engagement (8 items; $\alpha = .88$), (b) Efficacy in Instructional Practices (8 items; $\alpha = .93$), and (c) Efficacy in Classroom Management (8 items; $\alpha = .88$) using the long form of the Teacher Sense of Efficacy Scale (Tschannen-Moran, & Woolfolk Hoy, 2001). Teachers responded to items, such as "How much can you do to get through to the most difficult students?" on a 9-point Likert scale where 1 = "nothing" and 9 = "a great deal."

2.3.5 Teacher Mindset Scale

Mindset refers to the degree to which a person believes talents, skills, or intelligence more broadly are inherent or unchangeable on the 'fixed' end of the spectrum or can be developed and strengthened with practice and effort on the 'growth' end of the spectrum. The Teacher Mindset scale assesses the degree to which teachers believe students can grow their talents and skills with effort across six prompts, for example "You can learn new things but you cannot change your basic intelligence." Teachers were asked to rate each prompt on a 6-point Likert scale where 1 = "strongly disagree" and 6 = "strongly agree" (Gutshall, 2014; $\alpha = .82$).

2.3.6 Teacher's Observation of Potential in Students (TOPS)

Teachers were asked to provide a rating across nine indicators of talent for each student in their classroom. The nine indicators, including Learns Easily, Shows Advanced Skills, Displays Curiosity and Creativity, Has Strong Interests, Shows Advanced Reasoning and Problem Solving, Displays Spatial Skills, Shows Motivation, Shows Social Perceptiveness, and Displays Leadership, were assessed using a 4-point Likert scale (1 = never to 4 = always). A composite talent score was created using all nine indicators with strong internal consistency ($\alpha = .94$; Coleman et al., 2010). For the purposes of this study, individual student ratings were aggregated at the teacher level.

2.4 Design

Exploratory factor analysis (EFA) rather than confirmatory factor analysis (CFA) was chosen to identify integral constructs underlying the SITES-PD. EFA was employed because of the novelty of the topic (previously unexplored themes of engagement with professional development) created theoretical uncertainty surrounding the underlying structure of the SITES-PD (Browne, 2001) and the potential for stronger structural evidence to emerge during

later CFA replications (Goldberg & Velicer, 2006). Principal component factor analysis was employed using Varimax rotation with Kaiser Normalization, given its relative tolerance of multivariate nonnormality and its superior recovery of weak factors (Briggs & MacCallum, 2003; Cudeck, 2000; Fabrigar et al., 1999). Communalities were estimated through squared multiple correlations and were iterated to produce final communality estimates (Gorsuch, 2003).

One of the more critical decisions in an EFA is to determine the correct number of factors to retain and rotate (Fabrigar et al., 1999; Tabachnick & Fidell, 1996). The most common rule is to retain factors when eigenvalues are $\geq$ 1.0. This solitary criterion is the default procedure in most statistical packages. The shortcoming is that implementation of solitary criteria tends to under- or overestimate the number of true latent dimensions (Gorsuch, 1983; Velicer et al., 2000; Zwick & Velicer, 1986). Accordingly, each model was evaluated against the following five rules: (a) eigenvalues greater than 1.0 (Kaiser, 1960); (b) scree (Cattell, 1966), (c) Glorfeld's (1995) extension of parallel analysis (PA; Horn, 1965), (d) minimum average parcels (MAP; Velicer, 1976), and (e) interpretability (Fabrigar et al., 1999; Gorsuch, 1983). Results from several investigations demonstrated that MAP and PA are the two best methods for determining the correct number of factors to accept and that the scree test is a useful adjunct (Buja & Eyuboglu, 1992; Glorfeld, 1995; Verlicer et al., 2000; Zwick & Velicer, 1986).

3. Results

3.1 Content Validity

Table 2 provides descriptive information for teacher participation in PD as well as SITES-PD indicators. The last column reveals that hours of PD and the mean quality ratings of all face-to-face and virtual teacher-coach interactions were correlated with the SITES-PD measure. As well, the SITES-PD indicators showed robust correlations with the summed composite (r = .70-.90). Correlations suggest that the SITES-PD measure reflects a teacher's sustained investment and engagement with professional development across the school year, rather than capturing a recency effect or other unforeseen subjective temporal influences.

Table 2. Descriptive statistics and correlations for Professional Development (PD) variables

	N	*M*	*SD*	*Min*	*Max.*	*r* SITES-PD
Total number of PD hours	108	22.04	12.82	0	59.50	0.29 ***
Mandatory PD hours	108	15.87	9.85	0	34.50	0.23 ***
Voluntary PD hours	108	5.95	5.01	0	25.00	0.28 ***
Voluntary PD F2F Quality Rating	108	1.56	0.94	0	3	0.66 ***
Voluntary PD Online Quality Rating	108	0.60	0.96	0	3	0.41 ***
SITES-PD						
Purposefulness of communication	108	1.85	0.67	1	3	0.79 ***
Teacher as learner	108	1.90	0.66	1	3	0.89 ***
Self-reliance	108	2.16	0.65	1	3	0.70 ***
Thinking flexibly	108	1.78	0.64	1	3	0.88 ***
Students as learners	103	1.85	0.70	1	3	0.90 ***
Rules, authority, group membership	105	1.85	0.69	1	3	0.82 ***
View of self as professional	105	2.04	0.64	1	3	0.87 ***
Leadership	105	1.65	0.65	1	3	0.82 ***

Note. F2F = face to face.

3.2 Construct Validity

Given the high correlations amongst SITES-PD indicators revealed in Table 2, it was anticipated that a one factor solution might emerge from exploratory factor analyses. Indeed, Table 3 indicates a one factor solution. One component was extracted using Principal Component Analysis. The first component explained 69.74% of variance. In keeping with the primary tenet of the American Psychological Association's Task Force on Statistical Inference (Wilkinson & Task Force on Statistical Inference, 1999), which is to employ minimally sufficient statistics when complex analyses are not necessary, we then compared the predictive validity of the factor score against a summed composite for SITES-PD with high internal consistency (α = .94). We ran our regression models for predictive validity twice; first we employed the extracted one factor solution and then we replaced it with the summed composite. Both solutions arrived at identical findings when regressions were run to assess predictive validity, so the composite is presented below and recommended for future research.

Table 3. Exploratory factor analysis

	1 Principal Component
Purposefulness of communication	0.79
Teacher as learner	0.89
Self-reliance	0.68
Thinking flexibly	0.89
Students as learners	0.90
Rules, authority, group membership	0.82
View of self as professional	0.88
Leadership	0.82

Note. Principal Component Analysis.

3.3 Predictive Validity

The aim of the study was to develop an instrument that could help researchers identify teacher engagement behaviors with sustained professional development that lead to the greatest improvements in effective classroom practice and teacher beliefs about teaching and learning. A series of regressions were run with baseline scores covaried to assess whether (a) the total number of structured professional development hours or (b) the quality of sustained engagement (SITES-PD) was able to predict improvement across the academic year. Differentiating between mandatory and voluntary PD participation did not elucidate any interesting patterns (both forms of PD contributed to gains); thus, all PD participation hours were summed. In this set of analyses, baseline scores were compared to outcomes after the first year (but did not include subsequent years of intervention).

Table 4 includes findings from two data sources in order to enhance predictive validity assumptions. First, observers blind to the aims of the study conducted classroom observations using two measures that assess the effectiveness of teacher-student instructional interactions (CLASS and COS-R). Results indicate that the SITES-PD accounted for all variance that could have been attributed to hours of PD participation and provided some additional variance in terms of predicting improvement in teacher beliefs and behaviors across the academic year. Second, teachers rated their own perceptions of Teacher Self-Efficacy, Teacher Mindset, and Teachers' Observation of Potential in Students (TOPS). SITES-PD predicted more variance in score improvements than numbers of hours of PD for all measures with the exception of TOPS.

Table 4. Predictive validity regressions

	Baseline		# Hours PD		SITES-PD		Full Model	
	Beta	**SE**	**Beta**	**SE**	**Beta**	**SE**	***df***	**F**
Classroom Observation								
CLASS Emotional Support	0.128	0.163	0.214	0.007	0.345*	0.167	55	7.928***
CLASS Organizational Support	0.357**	0.171	0.134	0.009	0.267*t*	0.211	55	8.367***
CLASS Instructional Support	0.272*	0.159	0.090	0.013	0.344*t*	0.309	55	8.245***
COS-R Teacher Instructional Behaviors	0.184	0.127	-0.429	0.007	0.423**	0.160	54	6.568***
COS-R Student Engagement	0.155	0.138	-0.259	0.012	0.209*t*	0.274	54	2.028
Teacher Perceptions								
Self-Efficacy Student Engagement	0.535***	0.090	-0.251*	0.008	0.305**	0.187	68	14.960***
Self-Efficacy Instructional Strategies	0.511***	0.079	-0.157	0.008	0.327**	0.180	69	13.940***
Self-Efficacy Classroom Management	0.565***	0.094	-0.051	0.008	0.297**	0.196	68	19.619***
Teacher Observation of Potential in Students	0.456***	0.134	0.041	0.071	0.101	0.003	57	5.502**
Mindset	0.46***	0.094	0.068	0.011	0.342**	0.200	71	7.178***

Note. t < .10; * p < .05; ** p < .01; *** p < .001.

4. Discussion

Findings suggest that SITES-PD is a useful instrument for predicting whether the intervention has a positive impact on teachers' beliefs and observed classroom practices, above and beyond the total number of hours teachers participate in PD. The instrument holds promise for offering a deeper understanding of teacher engagement behaviors that may mediate the relation between professional development and classroom outcomes (operationalized here as observed effective instructional practices and teacher perceptions of the learning environment). When analyzed alongside total hours of participation in PD, SITES-PD accounted for variance in six of ten outcome variables and positive trends for additional three outcomes. Whereas, total hours of participation in PD only added unique variance for one of ten outcomes.

In the current study, using professional learning opportunities that combined both mandatory and voluntary sessions and support from coaches, as well as involving teachers in decision-making about content were explicitly built into the sustained professional development model to promote teacher buy-in and engagement. The development of the SITES-PD was inspired by the hypothesis that increased teacher engagement could improve the efficacy of the professional development intervention. Findings reveal teacher engagement, as assessed by SITES-PD, predicted the magnitude of desired outcomes, above and beyond PD

quantity. Specifically, teachers who were engaged in the PD intervention throughout the year rated themselves more positively on self-efficacy and mindset, and blind observers noted significant increases in effective instructional practices. Our findings are important, because previous research has indicated an association between substantial professional development and student learning (Yoon et al., 2007). Additional studies have highlighted the need to support and strengthen teacher professional development in order to bolster student academic performance (De Simone, 2020; Gupta & Lee, 2020; Kinnucan-Welsch et al., 2006).

The findings from the current study support previous research that suggests teacher engagement in PD is crucial for changes in teachers' classroom practices (Darling-Hammond et al., 2017). The present study adds evidence that sustained professional development can improve the effectiveness of teacher instruction, and that the mechanism for change is teacher engagement with professional development, as assessed by SITES-PD. Findings reported here suggest the SITES-PD tool demonstrates adequate content, construct, and predictive validity.

4.1 Limitations and Future Directions

The present study has a few limitations and offers pathways for future research. First, although the SITES-PD was an effective tool for the current study, the instrument may not work in the same manner with other interventions that have different structures and goals. SITES-PD may need to be adapted for interventions that differ greatly from the intervention described in this study (for example, year-long coaching with teacher-driven content), which may in turn affect the reliability and validity of the measure. Second, the ratings of the SITES-PD were conducted by coaches, which may introduce bias or subjectivity. For the current study, the measure was implemented by observers who were invested in the outcomes of the project. The observers built a rapport with teachers over the course of the study, so remaining objective could have been difficult when conducting the SITES-PD. For our purposes, we collected data from two additional sources (teacher report and blind observers) to assess the reliability of the measure. To bolster objectivity, coaches were asked to tally all teacher-coach interactions throughout the year and rank the depth of the interaction in the moment. Those moment-to-moment ratings averaged across the year were moderately correlated with the SITES-PD, adding some degree of comfort that the SIES-PD was not prone to recency effects. We further weighed the predictive validity of the SITES-PD against the quantity of PD received. In the context of this study, we were satisfied with the results but other projects that cannot collect data from multiple sources may not be able to ignore the possibility of rater bias.

4.2 Conclusion

The SITES-PD tool holds promise for informing future PD assessments and evaluations. SITES-PD can be implemented by other professional development interventions to measure the effectiveness of the professional development experience. While the indicators of the SITES-PD may need to be modified to fit the needs of specific interventions, the tool can assist coaches in identifying educators who could gain the most from professional development. The current tool may also help identify teachers who have not bought in early on, and coaches can work to target and strengthen those teacher-coach relationships. SITES-PD may be a useful formative evaluation tool that provides feedback on opportunities for PD implementation

improvement in real time.

Acknowledgements and Sponsoring Information

Funding for this work comes from the US Jacob Javits Gifted and Talented Education Act, Grant Award Number S206A140029.

We thank the principals, teachers, and other school personnel without whom this project would not be possible.

References

Briggs, N. E., & MacCallum, R. C. (2003). Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. *Multivariate Behavioral Research, 38*, 25-56. https://doi.org/10.1207/s15327906mbr3801_2

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36*, 111-150. https://doi.org/10.1207/s15327906mbr3601_05

Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research, 27*, 509-540. https://doi.org/10.1207/s15327906mbr2704_2

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276. https://doi.org/10.1207/s15327906mbr0102_10

Coleman, M. R., Shah-Coltrane, S., & Harrison, A. (2010). *Teacher's observation of potential in students: Whole class form*. Council for Exceptional Children.

Corcoran, T. B. (1995). Helping teachers teach well: Transforming professional development. *CPRE Policy Briefs*. https://doi.org/10.1037/e383832004-001

Cudeck, R. (2000). Exploratory factor analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 265-296). Academic Press. https://doi.org/10.1016/b978-012691360-6/50011-2

Curby, T. W., Stuhlman, M., Grimm, K., Mashburn, A., Chomat-Mooney, L., Downer, J., … Pianta, R. C. (2011). Within-day variability in the quality of classroom interactions during third and fifth grade: Implications for children's experiences and conducting classroom observations. *Elementary School Journal*, *112*(1), 16-37. https://doi.org/10.1086/660682

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective Teacher Professional Development*. Learning Policy Institute. https://doi.org/10.54300/122.311

De Simone, J. J. (2020). The roles of collaborative professional development, self-efficacy, and positive affect in encouraging educator data use to aid student learning. *Teacher Development, 24*(4), 443-465. https://doi.org/10.1080/13664530.2020.1780302

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181-199. https://doi.org/10.3102/0013189X08331140

Desimone, L. M., & Garet, M. S. (2015). Best practices in teacher's professional development in the United States. *Psychology, Society, and Education, 7*(3), 252-263. https://doi.org/10.25115/psye.v7i3.515

Earley, P., & Porritt, V. (2014). Evaluating the impact of professional development: The need for a student-focused approach. *Professional Development in Education, 40*(1), 112-129. https://doi.org/10.1080/19415257.2013.798741

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299. https://doi.org/10.1037/1082-989x.4.3.272

Fairman, J. C., Smith, D. J., Pullen, P. C., & Lebel, S. J. (2020). The challenge of keeping teacher professional development relevant. *Professional Development in Education*, 1-13. https://doi.org/10.1080/19415257.2020.1827010

Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*, 915-945. https://doi.org/10.3102/00028312038004915

Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement, 55*, 377-393. https://doi.org/10.1177/0013164495055003002

Goldberg, L. R., & Velicer, W. F. (2006). Principles of exploratory factor analysis. In S. Strack (Ed.), *Differentiating normal and abnormal personality* (2nd ed., pp. 209-337). Springer.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Lawrence Erlbaum Associates.

Gorsuch, R. L. (2003). Factor analysis. In J. A. Shinka & F. Velicer (Eds.), *Handbook of psychology* (Vol. 2). *Research methods in psychology* (pp. 143-164). John Wiley. https://doi.org/10.1002/0471264385.wei0206

Gupta, A., & Lee, G. L. (2020). The effects of a site-based teacher professional development program on student learning. *International Electronic Journal of Elementary Education, 12*(5), 417-428. https://doi.org/10.26822/iejee.2020562132

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching: Theory and Practice, 8*(3), 381-391. https://doi.org/10.1080/135406002100000512

Gutshall, C. A. (2014). Pre-service teachers' mindset beliefs about student ability. *Electronic Journal of Research in Educational Psychology, 12*(3), 785-802. https://doi.org/10.25115/ejrep.34.14030

Hampton, G., Rhodes, C., & Stokes, M. (2004). *A practical guide to mentoring, coaching and peer-networking: Teacher professional development in schools and colleges*. Routledge. https://doi.org/10.4324/9780203591949

Horn, J. (1965). A rational and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185. https://doi.org/10.1007/BF02289447

Jones, K. (2017). Collaboration, creativity and capital in professional learning contexts. *Professional Development in Education, 43*, 1-5. https://doi.org/10.1080/19415257.2017.1254371

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151. https://doi.org/10.1177/001316446002000116

Kinnucan-Welsch, K., Rosemary, C. A., & Grogan, P. R. (2006). Accountability by design in literacy professional development. *The Reading Teacher, 59*(5), 426-435. https://doi.org/10.1598/rt.59.5.2

Liang, X., Collins, L. J., Lenhart, L., & Ressa, V. (2020). Instructional change following formative instructional practices professional development. *Teacher Development, 24*(1), 108-125. https://doi.org/10.1080/13664530.2019.1705886

O'Sullivan, H. (2011). Leading and managing professional learning in schools. In H. O'Sullivan & J. West-Burnham (Eds.), *Leading and managing schools* (pp. 111-125). Sage. https://doi.org/10.4135/9781446288870.n8

Palermo, C., & Thomson, M. M. (2019). Large-scale assessment as professional development: Teachers' motivations, ability beliefs, and values. *Teacher Development, 23*(2), 192-212. https://doi.org/10.1080/13664530.2018.1536612

Palmer, K., & Noltemeyer, A. (2019). Professional development in schools: Predictors of effectiveness and implications for statewide PBIS trainings. *Teacher Development, 23*(5), 511-528. https://doi.org/10.1080/13664530.2019.1660211

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system (CLASS) manual, K-3*. Paul H. Brookes Publishing Company.

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (6th ed.). Pearson.

Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17*, 783-805. https://doi.org/10.1016/s0742-051x(01)00036-1

VanTassel-Baska, J., Quek, C., & Feng, A. X. (2007). The development and use of a structured teacher observation scale to assess differentiated best practice. *Roeper Review, 29*(2), 84-92. https://doi.org/10.1080/02783190709554391

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41*, 321-327. https://doi.org/10.1007/bf02293557

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helms (Eds.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy* (pp. 41-71). Guilford. https://doi.org/10.1007/978-1-4615-4397-8_3

Whitworth, B. A., & Chiu, J. L. (2015). Professional development and teacher change: The

missing leadership link. *Journal of Science Teacher Education, 26*(2), 121-137. https://doi.org/10.1007/s10972-014-9411-2

Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594-604. https://doi.org/10.1037/0003-066x.54.8.594

Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. L. (2007). Reviewing the evidence on how teacher professional development affects student achievement. *Issues & Answers* (REL 2007-No. 033). IES, U.S. Department of Education.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442. https://doi.org/10.1037/0033-2909.99.3.432

Appendix A

Sustained Investment and Teacher Engagement with Professional Development Scale (SITES-PD)

Reflecting on the last PD cycle, rate teacher-initiated interactions on a scale of 1 to 3. Use descriptors and examples to aid score assignment.

1. Was communication purposeful in terms of teaching and learning?

1	2	3
Perfunctory; requests materials or standards alignment.	Enabling; asks for help without attempting or demonstrating initiative	Genuine; asks for feedback or advice after applying new skill. Reflective; aims to improve and meet student needs.

2. What was the teacher's level of comfort with learning?

1	2	3
Fear of losing control or doing something wrong. "I'm not good at this"	Comfortable trying new approaches within comfort zone; small steps forward.	Comfortable with trying new approaches and making mistakes; able to take risks and step outside of comfort zone.

3. Did the teacher demonstrate self-reliance?

1	2	3
Dependent; "How do I do this?" "How do I manage this logistically/instructionally?"	Some barriers to independence; "I'm swamped; can you help me find materials for a lesson I have planned?"	Autonomous; interactions involve information sharing.

4. Did the teacher demonstrate flexible thinking?

1	2	3
Can't apply new concepts to existing framework. Stuck inside current routines/structure.	Thinking inside of the box; can apply some new strategies or concepts within a familiar instructional approach or content area.	Big-picture thinking; able to see how models and strategies apply across contexts; holistic approach to planning.

5. How was the teacher's perception of students as learners?

1	2	3
Students as doers; focus on discrete skills and factual knowledge; "what standards should be accomplished?"	Students as individuals on different points along the same path; all children receive same instruction but pacing is individualized; sees challenging students' thinking and learning as developmentally inappropriate.	Students as creators; focus on understanding, application, synthesis, concept development. "What can students accomplish?"

6.What was the teacher's view of rules and authority?

1	2	3
Rules bound; confined by authority; external locus of control.	Cognizant of and perhaps overwhelmed by competing demands; can see what is best for students but expresses trepidation.	Responsive; relies on one's own judgment and expertise within parameters; internal locus of control.

7. What was the teachers' view of self as a professional?

1	2	3
Sees teaching as a job. Work parameters are defined by school hours.	Defines self as teacher but does not seek continuous improvement beyond administrative mandates.	Sees teaching as a profession: views continuous improvement and student learning as a personal and professional responsibility.

8. Did the teacher assume a leadership role?

1	2	3
Hesitant to share in group setting; participates in PD but uncomfortable or uninterested in leading others	Agrees to participate in demonstrations; shares experiences with colleagues	Actively seeks out opportunities to model and coach; voluntarily mentors colleagues. Seen by others as a resource.