

MEASURES OF CLASSROOM INSTRUCTIONAL PRACTICES IN MIDDLE GRADES

Minsuk K. Shim, Robert D. Felner, Eunjae Shim, & Stephen Brand, University of Rhode Island
Minsuk K. Shim, Univ. of Rhode Island, 707 Chafee Bldg., Kingston, RI 02881

Recent national and state efforts for educational reform have brought our attention to changes in instructional practices. Educators and policymakers are interested in identifying the instructional practices that “work” in improving student performance. This has led to current enthusiasm for educational standards in several curriculum areas (National Council of Teachers of Mathematics, 1989; National Council of Teachers of English, 1996).

To monitor the impact of such efforts, we need accurate data on instructional practices. Much of the data on instructional practices are self-reported by teachers, and traditionally of questionable quality (Burstein et al, 1995). For self-reported survey data to be valid, they must portray the accurate description of what is actually happening in the classroom. Teaching and learning are complex procedure, which involves continuous interactions among teachers and students. As Burstein et al (1995) argued, little effort has been made to validate whether the national survey data collected from teachers measure the complex procedure of classroom instruction. This explains why many studies on instructional practices have depended on in-depth case studies from a handful of classrooms. It is hardly possible to generalize the findings to other classrooms. The limited generalizability of case studies becomes more problematic as policymakers need to understand the impact of reforms in our educational system. Survey data, a cost-effective way to include a large number of classrooms are very appealing in that sense. Few studies have examined the validity of the self-reported data on instructional practices although they have often been used to determine the impact of educational reforms. Mayer (1999) called for more research on the issues of survey reliability and validity.

The purpose of this study is to obtain evidences of the reliability and the validity of a self-reported survey inventory designed to assess the degree to which teachers implement recommended instructional practices in the classroom.

Scale Development

A large-scale survey called the High Performance Learning Community (HiPlaces) Assessment was developed as an evaluation tool for one of the biggest evaluation project on recent educational reform (Felner, Kasak, Mulhall, & Flowers, 1997; Felner, Shim, Brand,

Favazza, & Seitsinger, 2000). It was developed to examine the degree to which a broad range of recommendations for effective school reform are implemented in a school as well as to examine more fully their impact on students and staff. Among various dimensions of whole school reform, the HiPlaces Assessment has a section that asks questions on the frequency each teacher uses the recommended instructional practices based on the recommendations of nationwide reform initiatives on middle grades (Turning Points: Carnegie Task Force on Education of Young Adolescents, 1989) as well as national curriculum standards (NCTM, 1989; NCTE, 1996). Teachers reported the frequency with which they used various instructional practices using a 7-point scale with the following response categories: never, several times a year, monthly, several times a month, weekly, several times a week, and daily.

Classroom Instructional Practice Scale (CIPS) was originally developed from the classroom routine section of the Classroom Information Sheet (Wiesz & Cowen, 1976) as well as further items that were written by the authors to assess specific middle school practices. While the original CIS has three subscales, CIS items and the items written by the authors were combined on rational grounds to assess the following constructs: Cognitive development, Interdisciplinary teaching, Parent involvement, Heterogeneous grouping, Health promotion, Student recognition, Rule emphasis, and Group projects. In the 1990-91 sample, acceptable levels of reliability were obtained for the first four scales, while modest levels were found for other scales. An initial revision of the CIPS was conducted for the 1991-92 survey. Based on the initial reliability analyses, Rule emphasis and Group projects scales were dropped. Health promotion and Student recognition were retained for further use as these constructs were intrinsically important for Turning Points goals. On the remaining scales, items were dropped or moved based on item-total correlations and conceptual consistency. The revised scales exhibited moderate to high levels of internal consistency in the 1990-91 and 1991-92 samples.

The CIPS underwent a complete revision for the 1992-93 survey, resulting in the instrument that is currently in use. These revisions were driven by the need to increase the coverage of Turning Points goals and refine the construct definition underlying these

scales. New scales were constructed to assess Small group active instruction, Citizenship and social competence, Health promotion, and Mastery based assessment and student recognition. Further, new items were added to the existing measures of Integration and interdisciplinary practices, Community-based learning opportunities, Critical thinking enhancement, and Heterogeneous grouping practices.

In 1992-93, extensive factor analyses were done with 37 middle schools in Illinois. It is known that the quality of measurement improves if similar items measuring an underlying construct are combined together as one indicator for the construct. Mayer (1999) suggested to use composite scales that combine items measuring the same latent construct rather than individual item for more accurate picture of instructional practices. Eight sub-scales consist of 56 items were emerged as distinct empirical factors and they were also validated by conceptual judgment of a panel of experts¹. By 1996-97, the number of schools participated in the Survey was increased from 39 schools in Illinois to 401 schools in 16 states².

Data and Method

Data for this study was drawn from The HiPlaces Assessment administered to a large number of teachers and students in middle grades across 5 years (1992-93 to 1996-97). In this study, only the teachers who teach middle grades (grades 6 to 8) in typical middle grade schools (6-8, 7-8, 5-8, 7-9, etc.) were selected³. In addition, only the classroom teachers who teach "core" subject areas (mathematics, language art, science, social study, and reading) were selected as we found that instructional practices in non-core subject areas were quite different from those in core subject areas.

Research for the study was conducted with three different statistical techniques: Factor analytic study (both exploratory and confirmatory), reliability study and correlational study. First, the exploratory factor analyses were conducted to identify conceptually meaningful dimensions of CIPS. Factor structures can vary due to sampling fluctuation and differences in factor analytic procedures. Therefore, considerable attention was given to the stability of CIPS factor structures over time. A series of confirmatory factor analyses were also conducted to see whether the proposed measurement model adequately fit the sample data (Jöreskog & Sörbom, 1979). Second, the reliability analyses were conducted to examine the internal consistency in teacher responses using coefficient alpha statistics (Cronbach, 1951). Coefficient alpha was selected because the items on the survey were scored polytomously. Last, correlations

between teacher report and student report of instructional practices, and correlations between teacher report of instructional practices and student achievement were examined to provide evidences of criterion-related validity of CIPS.

Results

Factor analyses

Oblique rotation of 7, 8, and 9 factor analysis was undertaken for 1992-93 data. The 8-factor solution afforded the most meaningful interpretation of the empirical dimensions of the instructional practice construct. Eight-factor solution was applied to the data for later years to see whether the factor structures were stable across years. The factor loadings in 1992-93, 1994-95 and 1996-97 are presented in Table 1⁴. Most of the factors were clean and readily identifiable. Although some items were loaded on multiple factors in later years, extracted factors were, in general, congruent across years. When the items were loaded on multiple factors, they were classified on a conceptual basis judged by the panel of experts. The items "Students read and discuss newspaper articles", and "Additional instruction is provided for students" are examples of those cases.

Maximum likelihood confirmatory factor analyses were employed to examine the goodness of fit of the 8-factor model. EQS for Windows Release 5.7b (Bentler & Wu, 1995) was utilized to estimate the parameters of models consisting of the eight factors. Table 2 presents the goodness of fit indexes across 5 years: χ^2 statistics (Jöreskog, 1969), Nonnormed Fit Index (NNFI: Bentler & Bonnet, 1980) and Comparative Fit Index (CFI: Bentler, 1990)⁵. A baseline model was employed in which each item was allowed to load on only one of the 8 hypothesized latent constructs. These latent variables were allowed to covary, and residual covariances were fixed to zero. While the hypothesized 8-factor model (Model 1) did not fit the data adequately (CFI for the model ranged from .83 to .86), the fit indexes were sufficiently high to suggest that modification would yield models with acceptable fit. As some items are closely related to each other, and some items are loaded with multiple factors, we decided to allow several items to be inter-correlated. Based on the modification indexes provided by the stepwise multivariate LaGranger Multiplier test, the final model (Model 2) with 23 correlated residuals and 6 cross-loaded items was tested. The model with correlated residuals attained a level of fit that is generally considered to be acceptable (Bentler & Bonnet, 1980): NNFI was about .90 and CFI was about .91 across years.

Table 1. Factor Loadings of Classroom Instructional Practice Items Across Years

	1992-93	1994-95	1996-97	
	Small group active instruction			
	-.75-.03-.01 .03-.13 .05-.01-.06	.73-.03 .05 .05-.17-.07 .02-.04	.82-.02-.01 .00-.08-.04-.05-.06	Students participate in cooperative learning.
	-.59 .06 .11 .04-.04 .06-.03-.04	.66 .10 .11 .01-.00-.03 .05 .01	.65 .05-.11-.04 .04 .03 .04 .01	Students engage in group problem solving.
	-.59 .14-.04 .05-.13 .05-.07-.08	.65 .08-.05 .06-.06 .22-.14 .05	.61 .12 .07-.02-.14 .05 .08 .13	Students work on group projects.
	-.45 .10 .08 .05-.09-.04 .07 .01	.41 .08 .14 .04 .02-.07 .19-.12	.47 .05-.08-.08 .10 .15 .02-.09	Students participate in peer tutoring.
	-.44 .09 .05 .05 .26 .08 .14 .01	.61 .05-.08 .11 .17 .17 .03 .09	.56 .07 .09-.02 .19 .11-.05 .17	Students engage in "hands on" learning activities.
	-.41 .00 .03-.02-.33 .08 .10-.15	.55 .01 .08-.01-.19-.06 .05 .11	.52-.00-.05-.05-.12-.01-.04 .08	Assignments are given to allow students to get to know others.
	-.40-.07 .04-.05-.40 .05 .11-.15	.61-.05 .03 .09-.26-.15 .11-.02	.71-.04-.03-.05-.09-.09-.09-.03	Students participate in small group discussions.
	-.35 .07 .02-.01-.05 .13 .28-.04	.53 .03-.03 .02-.02 .06 .20 .04	.47 .09 .02 .00-.13 .08-.14 .07	Students work in heterogeneous ability groups to create projects.
	Community based learning opportunities			
	.01 .77 .01-.01-.11-.04 .05-.06	-.01 .70 .07-.04-.01 .00 .06 .13	.04 .73-.09 .04-.01-.07-.07 .04	Assignments to help students learn about community resources.
	.01 .76-.00 .01-.09-.01 .03 .06	-.04 .72 .01-.04-.08 .05 .02 .09	-.01 .81-.08-.00-.02-.08-.04-.06	Opportunities provided for community experiences to expose students to different cultures and conditions.
	-.02 .76-.03 .04 .05 .02 .01-.03	.02 .75 .01 .09-.00-.07-.05-.06	-.01 .74 .04-.04 .05-.02 .02-.02	Community/service learning opportunities are provided.
	.06 .66 .03 .03-.15-.02 .12-.06	.08 .64 .10-.01-.04-.03-.02 .09	.02 .73-.04-.01-.02-.08-.11 .04	Projects to help students learn about community issues.
	.02 .64 .01 .04 .10 .03-.04-.02	.01 .76-.04-.01 .03-.02 .01-.07	-.02 .64 .02 .01 .03 .07 .03 .01	Supervised youth service experiences in the community are a part of a student's academic program.
	-.18 .57 .09-.01 .02-.05 .06-.07	.18 .46 .05 .00 .03 .13-.02 .03	.22 .25-.07-.12 .06 .12 .06 .06	Students engage in real world learning activities.
	-.01 .45-.00 .19-.03 .06-.13 .07	.01 .54-.02 .12-.06-.06-.01 .03	.01 .54 .03-.05-.05 .00 .11 .09	People from the community are brought in to speak to the class.
	-.10 .34 .04-.04 .07 .04 .06-.02	-.07 .37 .05 .11 .04 .01 .15-.05	.02 .47-.01 .01 .03 .05-.03-.02	Students are linked with outside adult mentors/programs.
	Citizenship and social competence			
	.02-.05 .87 .02 .06 .02 .01 .00	-.02-.02 .84 .04-.10 .01 .01-.09	-.03-.01-.85-.04-.01 .00-.03-.06	Personal growth and development is emphasized.
	-.03-.10 .80-.01 .04 .04 .03-.03	-.02-.03 .84 .05-.08-.03 .01-.10	.01-.00-.75-.01-.07-.02-.08-.06	Social skill development is emphasized.
	.11 .04 .80-.02-.06 .02-.01-.12	-.03 .03 .77 .06-.05 .08-.10-.02	-.00 .02-.73-.05-.02 .06 .03 .01	Emphasis on peer resistance / assertiveness is stressed.
	.06 .11 .73-.02-.10 .00 .00-.10	.06 .10 .60 .01 .05 .06-.02 .15	.03 .05-.58-.02-.02 .11 .03 .09	Coping skills development is emphasized.
	.03 .17 .63-.04-.04 .05-.08 .02	-.07 .17 .59 .02 .02 .01-.02 .13	-.03 .13-.61 .02 .04 .01 .09 .11	Citizenship development is emphasized.
	-.05-.00 .63 .01 .01-.02 .11-.03	.10 .00 .54 .01 .12 .04 .12 .07	.12 .01-.55-.03 .11 .06-.05 .04	Practical applications of course materials are emphasized.
	-.17-.06 .60 .06 .01 .04 .05 .06	.18-.02 .50-.04 .17 .06 .17 .06	.22-.02-.43-.01 .16 .18-.02 .02	Students are taught problem solving / decision-making skills.
	.03 .05 .46 .06-.27 .09-.06-.06	-.02 .09 .49-.05-.16 .03-.02 .25	-.10 .05-.59-.02-.21-.06-.04 .13	Issues related to cultural diversity are emphasized.
	Integration and coverage of health topics/activities			
	-.07 .07 .10 .07 .12 .03-.05-.75	.02 .09 .05 .19 .04-.09 .11 .54	.04 .00-.06-.17 .05 .01-.11 .56	Health topics integrated as part of the broader classroom routines.
	.02 .16 .23-.07 .03 .03-.14-.47	.09 .15 .20 .00 .03 .08-.10 .46	.17 .13-.21 .07-.01 .10 .11 .46	Instructional units address health issues.
	-.15 .01 .03 .11 .00 .02-.07-.30	-.03 .13-.01-.17-.05-.02 .08 .21	.03 .07-.02-.10-.02 .03-.00 .27	Students participate in intramural activities.

1992-93

1994-95

1996-97

Integration and interdisciplinary practices

.03 .07 .05 **.88** .05-.01-.09 .07
-.04-.01 .01 **.78**-.04 .06-.11 .10
.06 .06 .02 **.61**-.03 .01 .19-.19
.01 .08-.06 **.53**-.05-.04 .13-.14
-.10 .10 .05 **.42**-.08-.02 .12-.08
-.02 .01-.04 **.38**-.13 .13 .07-.19
-.04-.05 .18 **.29**-.09-.08 .15-.**40**

-.01 .12-.02 **.82** .04 .06-.10-.10
.01 .06 .02 **.79** .06 .03-.06-.10
.00 .03 .03 **.72** .02 .06 .03 .04
.04 .00-.01 **.48**-.06-.04 .01 .06
.07-.05 .05 **.60**-.04-.04 .09 .12
-.04 .03-.01 **.34**-.16 .07 .07 .19
.05-.08 .16 **.38**-.09-.06 .06 .25

-.03 .05-.01-**.88** .04 .01 .15-.05
-.01-.01-.03-**.81** .04-.03 .07-.02
.04-.02-.00-**.77**-.02 .03-.02 .02
.03 .05-.01-**.38** .00 .01-.07 .10
.09-.04-.06-**.56**-.06-.08-.18 .07
.01 .07-.01-**.35**-.20-.02-.04 .19
.10-.02-.04-**.33**-.03 .05-.23 .08

Teachers from other subjects help to plan/carry out class projects.
Teachers from other subjects help to plan/carry out instructional units.
Classroom curricula are integrated with topics in other subject areas.
The class schedule is changed for instructional purposes.
Interdisciplinary teaching materials are used.
Use of the media center is integrated into lesson plans.
Instructional materials focus on topics relating to early adolescent concerns and interests.

Critical thinking enhancement practices

-.16 .06-.00 .11-**.64** .03-.04 .01
.02 .00 .08 .13-**.61** .04 .21 .03
-.02 .03 .06 .05-**.55**-.00 .07 .08
-.08 .03 .08 .12-**.51** .10-.20 .06
.04 .14 .14 .09-**.47** .01-.02-.17
-.29 .13 .13 .05-**.39**-.02 .03 .05
-.22 .10-.02 .07-**.38** .06-.08 .01
.06 .04 .04-.05-**.37** .29 .13-.06

.10 .09-.05-.00-**.59** .15 .01 .07
-.04-.03 .04 .13-**.57**-.03 .27 .02
.11 .01 .11 .07-**.45** .02-.04-.03
.04 .05-.01-.02-**.29** **.36**-.08 .22
.01 .10 .08 .10-**.20** .02-.01 **.27**
.27 .04 .09 .02-**.44**-.01 .14-.06
.25 .09 .05 .03-**.42** .08-.06 .09
-.04 .13 .02 .01-**.24** .17 .23 .04

.18 .04-.02-.02-**.60** .04 .02 .06
.08-.05-.07-.13-**.54** .10-.07-.02
.13 .03-.08-.01-**.38** .02-.10 .02
-.00 .02-.05 .02-**.43** .15 .13 .22
-.01 .04-.10-.10-**.19**-.09-.10 **.37**
.45 .01-.13-.06-**.26** .07-.02-.15
.35 .08-.05-.06-**.26**-.00 .09 .15
.05 .15 .01-.09-**.31** .27-.01-.05

Students write essay reports and papers.
Students revise their reports and papers.
Students write and keep journals.
Students take essay tests.
Students read and discuss newspaper articles.
Students provide feedback and comments on each other's work.
Students make speeches and/or presentations.
Portfolios of a student's work are used as an indicator of success.

Mastery based assessment and student recognition

-.14-.10 .12-.01 .07 **.60** .09 .08
-.04 .05 .04 .05 .03 **.59**-.10-.04
-.03-.05 .06 .05 .01 **.57** .06-.08
-.03-.01 .14 .03-.04 **.48** .14-.03
.09 .19-.03-.07-.06 **.38**-.06-.05
-.11 .00 .10 .07-.05 **.38** .30-.03
-.07 .03 .03 .15-.10 **.34**-.13 .10
.01 .03 .02-.03-.09 **.32** .22-.08

.07-.10 .18 .05 .02 **.39** .27-.09
.12 .09 .26 .00-.04 **.32** .05 .00
.03-.06 .19 .13-.04 **.45** .07 .03
.06-.05 .16 .09-.01 **.42** .19 .03
-.06 .13 .04 .01-.13 **.28** .14 .04
.11 .02 .06 .05-.08 **.22** **.29** .11
.08 .13-.01 .16-.19 **.34**-.06 .05
-.01-.00 .06-.02-.00 **.25** **.33** .03

.67-.09-.11 .03 .02 **.52**-.11-.03
.07 .05-.23-.04-.09 **.34** .02 .01
.05 .00-.11-.02-.02 **.47**-.03 .09
.08-.05-.09 .01 .00 **.53**-.01 .11
.01 .09 .02-.05-.12 **.33** .08 .02
.10 .10-.07-.05-.16 **.31**-.23 .00
.04 **.27**-.03-.12-.14 **.22** **.25**-.02
-.05 .07-.03-.03-.06 **.39**-.05 .01

Students are given multiple opportunities to improve their grade.
I seek student feedback on how to improve class.
Recognition is given for good behavior.
Students receive special recognition for exemplary work.
Students receive an incomplete until they meet the criteria.
Alternative/authentic assessments are employed to evaluate student.
Student project is shared with the rest of the school.
Students are allowed as much time as needed to demonstrate their knowledge on non-timed tests.

Instructional practices for heterogeneous grouping

-.00 .14 .08 .02-.11 .07 **.55** .06
-.27 .06 .21 .08 .02-.01 **.52** .11
-.20 .02 .25 .06 .10 .03 **.52** .13

-.00 .12 .00 .01-.09 .01 **.55** .00
.28-.04 .07 .06 .02-.09 **.53** .04
.18-.06 .14 .10 .06 .02 **.48** .09

-.03 .08-.04-.08-.05 .20-**.44** .04
.17 .06-.07-.01-.01 .14-**.50** .03
.16 .05-.12 .04 .04 .17-**.44** .07

Students are provided materials at different grade levels.
Strategies to facilitate learning at heterogeneous ability levels are used.
Class activities are designed to present information in a way that matches the students' preferred learning modality.
Additional instruction is provided to students ready to move on.
Additional instruction is provided to students who fall behind.
Self-paced learning materials are utilized.

Table 2. Model Goodness-of-Fit for CIPS

Model	Yr	χ^2	DF	NNFI	CFI	Δ CFI
Model 1	'93	2819	1456	.82	.83	
	'94	5616	1456	.84	.85	
	'95	12203	1456	.85	.86	
	'96	9051	1456	.84	.85	
	'97	10797	1456	.85	.86	
Model 2	'93	2270	1427	.89	.90	.07
	'94	4017	1427	.90	.91	.06
	'95	7902	1427	.91	.91	.05
	'96	6009	1427	.90	.91	.06
	'97	7511	1427	.90	.91	.05

Reliability

Having identified robust and distinctive dimension of instructional practices, we examined the internal consistency of the factorially derived CIPS. Table 3 shows the Cronbach's coefficient alpha statistics for 8 CIPS across 5 years. All scales showed moderate to high level of internal consistency across years. All scales except Integration of health and Mastery based assessment had coefficient alpha ranged from .8 to .91. Mastery based assessment had slightly smaller alpha than .8 (ranged from .76 to .79), whereas Integration of health had alpha ranged from .58 to .62. This scale has only three items combined whereas other scales have 7 to 8 items. This, in part, explains the relatively low reliability of this scale. When the total CIPS based on all 56 items was used, the reliability coefficient was very high across all years (approximately .95). Moreover, levels of internal consistency did not differ substantially between boys and girls, among grade levels, and among students from diverse racial and ethnic and socio-economic backgrounds.

Table 3. Reliability of CIPS

Cronbach's Alpha	N	'93	'94	'95	'96	'97
Small group	8	.88	.85	.87	.87	.86
Community-based learning	8	.84	.85	.85	.84	.84
Citizenship and Social Competence	8	.81	.80	.79	.79	.79
Integration of health	3	.91	.90	.89	.89	.89
Interdisciplinary integration	7	.86	.84	.82	.82	.82
Critical thinking	8	.62	.61	.60	.58	.60
Mastery-based assessment	8	.77	.78	.76	.77	.78
Heterogeneous groups	6	.84	.82	.81	.82	.83
Total CIPS	56	.95	.95	.95	.95	.96

Correlational analyses

We examined the extent to which teacher responses on their instructional practices were congruent with

student responses. Similar items were asked of both teachers and students on two CIPS subscales: Small group instruction, and Interdisciplinary integration. Teachers and students in grades 6 to 8 in middle schools were selected and their responses were aggregated at the school-level. Results are shown in Table 4. It shows a significant relationship between teacher and student report of the instructional practices ($p < .01$).

Table 4. Correlations Between Teacher and Student Responses on Instructional Practices

Staff		Student		
		small group	integration	combined
'93	Small group	.517**	.262	.387*
	Integration	.398*	.758**	.702**
	Combined	.517**	.596**	.629**
'94	Small group	.658**	.462**	.607**
	Integration	.470**	.717**	.673**
	Combined	.622**	.648**	.704**
'95	Small group	.570**	.310**	.472**
	Integration	.464**	.672**	.651**
	Combined	.573**	.544**	.623**
'96	Small group	.665**	.352	.573**
	Integration	.376**	.668**	.622**
	Combined	.591**	.573**	.675**
'97	Small group	.611**	.328**	.531**
	Integration	.328**	.606**	.558**
	Combined	.540**	.533**	.625**

* $p < .05$, ** $p < .01$

Correlations between teacher and student reports of Small group instruction ranged from .52 to .66, whereas correlations between teacher and student reports of Integration ranged from .61 to .76 across 5 years. When teachers more frequently utilized the instructional practices of integration and small group activities, students also reported they engaged in more activities, indicating the validity of teachers' self-reported instructional practices.

Table 5. Correlations between CIPS and Student Achievement in Mathematics

	'94		'95		'96	
	Gr6	Gr8	Gr6	Gr8	Gr6	Gr8
Small group	.55*	.60**	.41*	.43**	.25*	.34*
Citizenship and social competence	.40	.42*	.23	.30*	.53**	.38*
Integration	.70**	.50*	.47**	.38**	.87**	.47**
Critical thinking	.48*	.45*	.19	.22	.57**	.35*
Heterogeneous groups	.24	.14	.46**	.43**	.61**	.23
Total CIP	.52*	.39	.43*	.43**	.48*	.52**
N (Schools)	16	28	30	43	26	36

* $p < .05$, ** $p < .01$

We also examined the correlation between teacher report of their instructional practices and student achievement. In order to make the relationship more comparable, we examined the correlation between teachers' report of instructional practices and students' achievement in mathematics. Achievement data were available for Illinois schools from 1993-94 to 1995-96. Table 5 shows significant and positive correlations between instructional practices of mathematics teachers and student mathematics achievement, especially for Small group instruction and Integration practices. They ranged from .25 to .60 for Small group instruction and .38 to .87 for Integration.

Summaries and discussions

This study examines the reliability and the validity of self-reported survey data on instructional practices. It is based on nation-wide survey with more than 25,000 teachers in over 1,000 schools across 5 years. Our study shows that survey data provide a fairly accurate description of how often teachers use various instructional practices that are consistent with the recommendations of several reform initiatives. There was consistent and solid agreement between what teachers reported and what students perceived in terms of their classroom activities. The CIPS was positively related to student achievement in mathematics. Instead of using individual indicators, we found that grouped items, measuring the same underlying characteristics, provide more reliable measures of instructional practices. We proposed 8 dimensions of quality instruction. They measure distinctive constructs of instructional practices empirically and conceptually. Their factor structures were stable over 5 years and the hypothesized model fit the data well. As policymakers focus more on assessing instructional trends, it is not plausible to rely on in-depth studies of a small number of classrooms. Survey data will provide the most cost-effective way of measuring national trends in instruction.

References

- Bentler, P.M., & Wu, E.J.C. (1995). EQS for windows user's guide. Encino, CA: Multivariate Software.
- Bentler, P.M., & Bonnet, D.G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bentler, P.M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, 107, 238-246.
- Burstein, L, McDonnell, L.M., Van Winkle, J., Ormseth, T., Mirocha, J., & Guitton, G. (1995). *Validating national curricular indicators*. Santa Monica, CA: RAND.
- Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Felner, R.D., Shim, M., Brand, S., Favazza, A., & Seitsinger, A. (2000). *Improving achievement in the middle grades in mathematics and related areas: Lessons from the Project on High Performance Learning Communities: Proceedings of the National Convocation on Mathematics Education*. National Academy of Sciences: Washington, D.C.
- Felner, R.D., Kasak, D., Mulhall, P., & Flowers, N. (1997). The Project on High Performance Learning Communities. *Phi Delta Kappan*, 78, 520-527.
- Jöreskog, K.G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34, 183-202.
- Jöreskog, K.G., & Sörbom, D. (1979). *Advances in factor analysis and structural equation models*. Cambridge, MA: Abt.
- Mayer, D.P. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis*, 21, 29-45.
- National Council of Teachers in Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.
- National Council of Teachers in English & International Reading Association. (1997). *Standards for the English language art*. Reston, VA: Author.
- Task Force on Education of Young Adolescents. (1989). *Turning Points: Preparing American Youth for the 21st Century*. Washington D.C.: Carnegie Council on Adolescent Development.
- Weisz, P.V. & Cowen, E.L. (1976). Relationships between teachers' perceptions of classroom environments and school adjustment problems. *American Journal of Community Psychology*, 4, 181-187.

¹ Consisted of university researchers, principals and teachers.

² No. of schools: 37('93), 143('94), 361('95), 258('96), 401('97)
No. of teachers: 545('93), 2453('94), 6181('95), 4992('96), 8021 ('97)

³ In preliminary analyses, we found that teachers who taught other than middle grades had somewhat different instructional practices than those who taught middle grades. We also found that having younger students at school would affect the instructional practices of teachers who taught middle grades in K-8 schools.

⁴ The factor loadings for other years were not reported in Table 1 as they were quite similar to those reported in the table.

⁵ χ^2 statistics has been criticized as being dependent on sample size. Hu & Bentler (1995) reported that ML-based NNFI and CFI