

---

## Five Standards and Student Achievement

---

R. William Doherty, R. Soleste Hilberg, America Pinal, and Roland G. Tharp  
University of California, Santa Cruz

---

### Abstract

Two studies examine the influence of the Standards for Effective Pedagogy on student achievement gains. Participants were 15 teachers and 266 students (grades 3 to 5) in a public elementary school serving predominantly low-income Latino English Language Learners (ELLs). Study 1 found that higher use of the standards by teachers reliably predicted student achievement gains on SAT-9 tests of comprehension, reading, spelling, and vocabulary. Further analysis found teachers' use of the standards reliably predicted gains in *English* language achievement when English was the language of instruction. Study 2 found that achievement gains in comprehension, reading, spelling, and vocabulary were greatest for students whose teachers had transformed both their pedagogy and the organization of instructional activities as specified by the Standards for Effective Pedagogy model. These teachers used the standards extensively, both directly at the teacher center and indirectly through multiple, simultaneous, diversified learning activities. Implications for teaching practice and research are discussed.

### Five Standards and Student Achievements

Tharp and his colleagues (Tharp, Estrada, Dalton, & Yamauchi, 2000) have identified five teaching standards, the Standards for Effective Pedagogy, that are critical for improving learning outcomes for all students, and especially those at risk of academic failure due to cultural, linguistic, or economic factors. The first standard is to facilitate learning through joint productive activity in which teachers and students work together on a common product or goal and have opportunities to converse about their work. The second standard is to develop language and literacy across the curriculum; that is, develop competence in the language and literacy of instruction *and* in the academic disciplines through extended reading, writing, and speaking activities. The third standard is to contextualize instruction in the experiences and skills of students' homes and communities. The fourth standard is teaching complex thinking through challenging activities requiring the application of content knowledge to achieve an academic goal, with clear standards and systematic feedback on performance. The fifth standard is to teach dialogically using planned, goal-directed instructional conversations between a

teacher and a small group of students. This study reports on the relationship between teachers' use of these standards and students' language arts achievement gains in a school serving predominantly low-income Latino English Language Learners (ELLs).

Grounded in a sociocultural perspective of teaching and learning, the Standards for Effective Pedagogy (Five Standards) are the essential elements of the theory of teaching and learning first proposed by Tharp and Gallimore (1988) and later more fully elaborated by Tharp et al. (2000) in *Teaching Transformed: Achieving Excellence, Fairness, Inclusion and Harmony* (see also Tharp, 1994, 1997; Tharp, Dalton, & Yamauchi, 1994), but some clarifications are necessary here. First, these pedagogy standards are not intended to reflect the full spectrum of complex tasks that comprise teaching; rather, they represent guiding principles for instructional activities that promote active, effective student learning and that must be adapted to varying contexts and diverse student needs. Second, these standards do not stand in opposition to small-group direct instruction. Tharp and Gallimore used the terms direct and effective interchangeably in their report on reading comprehension in the Kamehameha Early Education Program (Tharp, 1982). Third, we do not propose these standards should be used to the exclusion of other strategies: In fact, our data suggests that teachers who use the standards at higher rates are more likely, not less, to use a variety of other effective teaching strategies (Doherty & Pinal, 2002). Finally, the terms transformed and untransformed are used here for the dual purposes of (a) denoting the pedagogy and classroom organization proposed by Tharp et al., and (b) differentiating teachers' use of these theoretically essential elements of effective teaching and learning. Transformed pedagogy denotes instruction based on the Five Standards; transformed organization denotes the use of multiple, simultaneous, diversified activity settings. No connotations outside this theoretical framework are intended.

The Five Standards are the result of three decades of research across cultural, linguistic, and economic contexts. This line of research began with the Kamehameha Elementary Education Program (KEEP), a program for at-risk K-3 Native Hawaiian students that operated from 1970 through 1988 with fidelity to its original self-description. Numerous publications have described that program (e.g., Au et al., 1986; Au & Jordan, 1981; Calkins et al., 1989; Tharp, 1982; Tharp et al., 1984). After reorganizing KEEP classrooms into peer-oriented small-group activity settings, students in the program evidenced significant improvement in reading achievement (Tharp, 1982) and higher rates of industriousness (Antill & Tharp, 1974), on-task behavior, and peer-directed cooperative behavior toward school-related goals (Tharp & Gallimore, 1988). Through many years of upscaling into fifteen multicultural public schools in Hawaii, evaluation results remained above non-KEEP programs' academic achievement and these effects continued until the program was formally terminated in 1997 after expansion pressures and reduction of resources eroded fidelity to the initial model (Calkins et al., 1989; Gallimore, Tharp, Sloat, Klein, & Troy, 1982; Klein, 1988; Klein & Calkins, 1988; Tharp, 1982; Yap, Estes, & Nickel, 1988).

The KEEP model was extended into Rough Rock Elementary School (Navajo) in Arizona in 1984 (Jordan, 1995; Vogt, Jordan, & Tharp, 1992), and a program "naturalized" in the Navajo locale took root (Begay et al., 1995; Dick, Estell, & McCarty, 1994; Sells, 1994). Its operations are fully congruent with the pedagogy and organization

proposed by Tharp et al. (2000), and literacy in both English and Navajo are significantly higher than in comparison groups.

Key elements of the KEEP model and the Five Standards have also been extended to the schooling of American Indian groups from Alaska (Barnhardt, 1982; Blum, 1998, n.d.; Demmert, 1994, 2001; Lipka, 1986, 1990, 1994; Preston, 1991; Scollon, 1981; Swisher & Deyhle, 1987) to the American Southwest (Hilberg, Tharp, & DeGeest, 2000; Hilberg, Doherty, Epaloose, & Tharp, 2001; Jordan, 1995; Jordan, Tharp, & Vogt, 1985; Yamauchi & Tharp, 1995), and low income populations such as Latinos (Padron & Waxman, 1999; Waxman, Huang, Anderson, & Weinstein, 1997; Waxman & Huang, 1997; Waxman, Huang, & Padron, 1995) and Appalachians (McIntyre, Kyle, Hovda, & Stone, 1999; McIntyre, Rosebery, & Gonzalez, 2001; McIntyre & Stone, 1998).

### **The Standards for Effective Pedagogy and Student Outcomes**

There is growing support for the theory of teaching and learning proposed by Tharp and his colleagues (2000). Consistent findings from correlational, quasi-experimental, and true experimental designs have documented a systematic relationship between use of the Five Standards and a broad range of affective, behavioral, and cognitive indicators of improved student performance.

In classrooms of largely Latino ELL students in which the Five Standards were used moderately or only slightly, students spent more time on-task, perceived greater cohesion in the classroom, and perceived themselves as better readers having less difficulty with their work than students in classrooms where the standards were not used at all (Padron & Waxman, 1999). Estrada (2000) found that teachers' use of the standards in literacy instruction was related to higher reading and language scores on the SABE for first graders, and SAT-9 language scores for fourth graders. Doherty and Pinal (2002) found that teachers' use of joint productive activity (JPA) during language arts instruction reliably predicted students' self-reported use of effective cognitive reading strategies, which in turn predicted achievement gains on standardized tests of comprehension. This study also found a direct relationship between teachers' use of JPA and student gains.

In a quasi-experimental design randomly assigning groups of eighth-grade American Indian students to either Transformed (Five Standards) or Traditional (Whole Class) mathematics instruction, Hilberg, Tharp and DeGeest (2000) found that students in the Transformed classes reported improved attitudes toward mathematics. They also evidenced more conceptual learning on tests at the end of the math unit and higher retention of unit content two weeks later.

In a series of true experimental designs, Saunders and Goldenberg (in press) found that students with varying levels of English proficiency taught using instructional conversation (IC) demonstrated greater understanding of story theme than students taught using direct instruction, although both groups demonstrated equivalent levels of literal comprehension. Students taught using both IC and contextualization showed significantly better reading comprehension and thematic understanding than students taught using either IC or contextualization separately (Saunders & Goldenberg, 1999).

## Theoretical Bases for Effectiveness

The Five Standards are based on sociocultural tenets (Tharp & Gallimore, 1988) that learning occurs best when (a) teachers and students work together on a common task or goal and have opportunities to converse during collaboration, (b) instructional activities are meaningfully connected to students' prior experience and knowledge; and (c) instruction occurs within the learner's zone of proximal development (ZPD), defined by Vygotsky (1978) as ". . . the difference between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers." From this perspective, teaching is assisting student performance with the goal of increasing that which students can do unassisted by the teacher; learning represents improved performance, or movement through the ZPD toward increased competence and autonomy (Tharp et al., 2000; Tharp & Gallimore, 1988).

The effectiveness of the Five Standards in teaching and learning can also be explained by decades of research in cognitive science. Wittrock (1978) favored a focus on how teaching style influences learners' attention, motivation, and understanding, which in turn influence student performance, rather than on how teaching style directly influences student performance. Central to this discussion is how teachers' use of the Five Standards promotes student elaboration -- the process of forming associations between new information and prior knowledge (Dansereau, 1988; Wittrock, 1978, 1986).

One line of research in cognitive psychology has established that context influences both the encoding and retrieval of information from long-term memory (Baddeley, 1990). Context facilitates encoding by organizing new information in a meaningful manner (Dodd & White, 1980; Schvaneveldt & McDonald, 1981), and features of the context encoded with the new information serve as cues that facilitate recall by priming associated memory structures (Baddeley, 1990; Dansereau, 1988; Howes, 1990). Memory performance will be greatest when there is a good match between the context in which information is encoded and the context in which it is retrieved (Tulving, 1983; Tulving & Thompson, 1971). A second line of research has shown that new information encoded using elaborative strategies is better retained and retrieved than when encoding involves rehearsal (Howes, 1990). Although learning can result simply by maintaining new material in short-term memory (Kolors & Brison, 1984), elaboration is more effective than rehearsal because it directly transfers new information from short-term to long-term memory. A single elaboration may result in learning; a single rehearsal is unlikely to do so. A third line of research has shown that the extent and type of processing used to encode new information influences memory performance. Processing that focuses on conceptual features such as its meaning, personal and social relevance, or relationship to prior knowledge and experience facilitates encoding, storage, and recall of new information and produces more durable memory ( Craik & Lockhart, 1972; Lockhart & Craik, 1990).

Integrating these three lines of research, learning may be defined as associating new information with prior knowledge (Baddeley, 1990; Lockhart & Craik, 1990). For learning to occur, relevant prior knowledge in long-term memory must be activated and the new information must undergo some form of processing. Processing that involves elaborative strategies focusing on conceptual characteristics of the new material improves learning. Also, the richer the associations made between the new and the known, the

more likely the new material will be retained and recalled (Baddeley, 1990; Lockhart & Craik, 1990). As Howes (1990) states, an extended body of elaborated, meaningful material is more easily encoded and retrieved than an extended body of unrelated elements.

A cognitive elaboration perspective of teaching and learning may serve to bridge the divisions between cognitive and sociocultural theorists. For example, Wittrock (1978) viewed learning from instruction as a generative process in which learners play an active, constructive role in forming associations between the new information and prior knowledge. From this point of view, effective instruction facilitates the learners' ability to construct meaning from experience, and the teachers' role is "to design different treatments for different students in different situations to actively induce mental elaborations that relate previous learning and schemata to stimuli" (p. 1). Tharp and Gallimore (1988) define teaching as assisting the performance of learners. Learning, then, is a collaborative process in which both teacher and learner play an active, constructive role. The teacher's role is to (a) design challenging learning activities that generate associations between the new information and students' prior knowledge from home, school, and community, and (b) to use the Five Standards to promote, guide, and sustain students' cognitive elaborations. For example, joint productive activities require sufficient elaboration of the new material to complete the task or achieve the goal of the learning activity. Contextualization makes new information more meaningful and, by activating students' prior knowledge, facilitates the forming of associations between the new and the known. Instructional conversations give the teacher a great deal of control over the extent and type of processing students use.

A highly abstracted Five Standards instructional model consists of a teacher and a small group of students having an instructional conversation while collaborating on a cognitively challenging activity contextualized in students' personal, social, or cultural knowledge and experience. Other students engage in multiple, diverse activities occurring simultaneously. The overarching goals of instruction are to foster complex thinking by all students, and language and literacy development in the language of instruction as well as in the content domains. The latter is especially important for diverse students and English Language Learners.

A more concrete example of this model can be drawn using instructional units developed in an American Indian school using the Five Standards to guide their reform efforts. An eighth-grade teacher team developed a thematic unit on current issues affecting the community. Tribal leaders were invited to speak to students at an assembly, after which teachers created subject area units related to the issues presented. In mathematics, students collaborated in small groups to generate and administer surveys on students' attitudes about school. The data from the surveys served as the basis for a unit on fractions, decimals, and percents. The survey results were later presented in multiple representations, such as pie charts, graphs, and frequency distributions, and used in student presentations and letters to the tribal council. This unit encouraged students to investigate the world around them, organize their thoughts and actions, and use mathematical concepts to communicate their ideas and knowledge. Mathematical principles that may have seemed of little practical value became tools for communicating personal concerns and influencing their world. In science, students focused on the issue of local water quality. Collaborating in small groups with the teacher, they sampled local

water sources, including water from fountains in the school and bottled water. By examining the water samples for their chemical content and pollutants, the periodic table, once simply a poster on their science classroom wall, became a lens for looking at their school, homes, community, and the world in an academically engaging and challenging way.

Contextualizing instruction by situating new information in meaningful contexts activates students' prior knowledge, making it more available for association with new information. This is, of course, in stark contrast to presenting new material in an atomistic, decontextualized, drill-like manner in which facts are presented in isolation. Contextualizing new information in students' everyday lives not only makes it more relevant and meaningful during encoding, linking schooled and everyday concepts makes the schooled concepts more readily retrieved in and pertinent to students' daily lives (Tharp & Gallimore, 1988).

Effective joint productive activities generate associations between the new and the known and provide a rich context for conceptual processing. For example, new information must be sufficiently elaborated to accomplish the task or achieve the goal of the learning activity. When both goals and feedback are present in an activity, self-evaluative mechanisms may generate elaborations such as planning (goal-setting and the allocation of resources prior to learning), information management strategies (organizing or summarizing), comprehension monitoring (assessment of one's learning or use of strategies), and evaluation (analyzing learning and strategies after the learning episode) to modify performance to greater accordance with the goals (Bandura & Cervone, 1983; Bandura & Schunk, 1981; Schraw & Sperling-Dennison, 1994). Joint productive activities also provide opportunities for the teacher to model his/her language, thinking, and problem-solving strategies. Modeling is widely acknowledged as one of the most powerful forms of teaching (Bandura, 1977, 1986), and its role in assisting the performance of learners has been discussed from a cognitive-behavioral perspective by Tharp and Gallimore (Gallimore & Tharp, 1995; Tharp & Gallimore, 1988). The articulation of problem-solving strategies by a more experienced person increases learners' performance on subsequent attempts compared to non-modeled trial-and-error learning (Gonci & Rogoff, 1998). By observing others, one forms rules of behavior that subsequently serve as a guide for action and, because people can learn through modeling before they perform any behavior themselves, they are spared the costs of faulty effort (Bandura, 1986). Modeling requisite skills increases students' understanding of verbal and conceptual explanations, thus increasing the potential for learning (Yamauchi & Tharp, 1995). This may be especially useful for assisting English Language Learners who may not fully understand what is being said in the classroom.

Instructional conversations during joint productive activity can prime students' prior knowledge, making it more accessible for associations with new information, and thereby assisting the integration of new material into long-term memory. Instructional conversations are academic, goal-directed discussions between a teacher and a small group of students. The teacher's role is to facilitate the active construction of students' mental and verbal elaborations, listen attentively, and assess and assist student understanding (Wittrock, 1978). By questioning students on their views, judgments, and rationales, as well as their experience, attitudes, values, and beliefs in relation to academic concepts, the teacher is able to activate a broad range of cognitive and affective

content in long-term memory. The IC allows the teacher to extend students' cognitive and verbal elaborations, assist them in their association-making efforts, and regulate the type of processing (i.e., conceptual) students apply to the new information.

If teachers' use of the Five Standards influences student cognition, which in turn influences student learning, then a direct relationship between use of the standards and student performance would be expected. There is some research evidence of this relationship. Doherty and Pinal (2002) found that teachers' use of the Five Standards during language arts instruction for primarily Latino ELL students reliably predicted students' self-reported use of effective comprehension strategies, and students' self-reported use of effective comprehension strategies reliably predicted their achievement gains on standardized comprehension tests (SAT-9). Teachers' use of the standards also predicted gains in students' comprehension achievement directly. This study examined the relationship between teachers' use of the standards and student achievement across a broader range of learning measures. Our specific hypothesis was that higher implementation of the standards during language arts instruction, as indicated by ratings of teachers' use of the Five Standards, would predict greater achievement gains on end-of-year standardized tests of comprehension, language, reading, spelling, and vocabulary.

## Method

### Participants

Participants were 15 teachers (2 men, 13 women) and 266 students (137 boys, 129 girls) in a public elementary school in Central California. Situated in a rural area, the school serves a community of predominantly low-income Hispanic families: 90% of students in the school are Hispanic, 78% receive free or reduced-price lunch, 68% are limited English proficient, and 38% are from migrant families. The school ranked in the second decile statewide on standardized test scores the previous year. Seven teachers in the sample taught third grade, and eight taught combined fourth- and fifth-grade classes. Teachers' years of experience ranged from 1 to 26 ( $M = 6.01$ ,  $SD = 5.42$ ). Of the students, 82 were in third grade (31%), 101 were in fourth grade (38%), and 83 were in fifth grade (31%).

### Measures

#### The standards performance continuum (SPC)

The SPC (Doherty, Hilberg, Epaloose, & Tharp, 2002) is a five-point rubric measuring teachers' performance of the Five Standards. Levels of standards performance are: *Not Observed* – the standard is not present; *Emerging* – elements of the standard are implemented; *Developing* – the standard is partially implemented; *Enacting* – the standard is fully implemented; and *Integrating* – at least three standards are implemented simultaneously in a single instructional activity (all Enacting level ratings in the same activity then become Integrating). In addition to individual subscale scores (range = 0 – 4), each teacher receives a SPC Total score, found by summing across subscales. Inter-coder agreement for SPC Total scores in this study was .91 (Kendall's W).

## English language proficiency (ELP)

All students in the school district are rated as either English-speaking, Fully English Proficient, or Limited English Proficient. Due to a highly transient student population, 43% of students' ELP data were unavailable. Of the 151 students with ELP data 19% were English-speaking, 11% Fully English Proficient, and 70% Limited English Proficient. To serve the needs of students with varying levels of English proficiency, the school provides three modes of instruction: English Only (EO), Structured English Instruction (SEI), and Bilingual (BIL; 28%, 18%, and 54%, respectively). With 43% of ELP data missing, and a Spearman's rank-order correlation coefficient of .80 (150;  $p < .001$ ) between ELP and mode of instruction, mode of instruction was used as the most reasonable indicator of students' English proficiency.

## Procedures

SPC data were gathered through live observations over the course of one semester by a pair of trained observers. Two 45-minute observations of language arts instruction were made of each teacher. Observations were separated by approximately seven weeks.

## Results

### Study 1

The outcome measures in this study were six indicators of student achievement estimated by year-end SAT-9 subtest scores: Comprehension, Language, Reading, Spelling, Vocabulary, and Overall NCE, found by averaging across all subtests. To control for the tendency of scores to regress to the mean on repeated measures of parallel tests (Soar, 1978), estimated gain scores (EGS) were computed for each SAT-9 subtest. The EGS were found by subtracting students' predicted scores, based on their prior year's test performance, from observed scores on each subtest. The independent variables were teachers' years of K-12 teaching experience (Teacher Experience), students' grade level (Grade) and language proficiency (Mode), and teachers' SPC Total scores averaged across the two observations. Preliminary analyses found teachers' SPC Total scores ranged from 4.00 to 16.50 with a mean of 10.11 and standard deviation of 2.55. Nonparametric tests (Kruskal-Wallis) found differences in SPC Total scores in third-grade classes ( $M = 9.09$ ,  $SD = 2.42$ ) and combined fourth- and fifth-grade classes ( $M = 10.59$ ,  $SD = 2.46$ ) did not differ significantly,  $\chi^2(2) = 1.71$  ( $p = .43$ ); SPC Total scores did differ significantly between modes of instruction,  $\chi^2(2) = 57.14$  ( $p < .001$ ). Teachers in SEI classrooms had the highest SPC Total scores, followed by those in EO and Bilingual classes ( $M_s = 10.91$ ,  $10.08$ , and  $9.81$ ;  $SD_s = .37$ ,  $1.43$ , and  $3.33$ , respectively). SPC Total scores were negatively correlated with teachers' years of experience,  $r(14) = -.66$  ( $p = .02$ ). SPC subscale inter-correlations ranged from .15 to .82 (see Table 1).

Table 1  
*Inter-correlations of SAT-9 Subtests*

Variable	Comprehension	Language	Reading	Spelling	Vocabulary	Overall NCE
Comprehension	1.00					
Language	.25*	1.00				
Reading	.69*	.33*	1.00			
Spelling	.29*	.29*	.28*	1.00		
Vocabulary	.15*	.32*	.63*	.25*	1.00	
Overall NCE	.65*	.66*	.82*	.63*	.69*	1.00

*Note.* \* =  $p < .05$ .

SPC Total scores are ordinal data and not likely to be normally distributed (Harwell & Gatti, 2001). Inspection of the scatterplot found that indeed SPC Total scores were positively skewed. A square root transformation was used to make the distribution more normal (Tabachnick & Fidell, 1989).

To estimate the contribution of SPC Total scores to the prediction of student achievement, --above that afforded by factors such as grade level, language ability, or teacher experience -- hierarchical regression analyses were run on each of the dependent variables. In the first step, Teacher Experience was entered in the model. In the second step, Grade and Mode were entered. Transformed SPC Total scores were entered in the final step. Table 2 presents the standardized regression coefficient ( $\beta$ ) and t-test of its significance for each variable at the step entered, and the multiple correlation coefficient ( $R$ ),  $R^2$ , degrees of freedom, and F statistic for each model tested.

Overall NCE

After step 1, with Teacher Experience in the equation,  $\underline{R}^2 = .00$ ,  $F_{inc}(1, 264) = .03$ ,  $p = .87$ . After step 2, with Grade and Mode added to the equation,  $\underline{R}^2 = .02$ ,  $F_{inc}(3, 262) = 1.87$ ,  $p = .14$ . After step 3, with transformed SPC Total scores added to the equation,  $\underline{R}^2 = .06$ ,  $F_{inc}(4, 261) = 3.97$ ,  $p = .004$ . The addition of transformed SPC Total scores resulted in a significant increase in  $\underline{R}^2$  of .04 ( $p = .002$ ) above the variance accounted for by Grade and Mode.

Table 2  
*Hierarchical Analysis of SAT-9 Subtests*

Variable	B	t	R	R <sup>2</sup>	df	F
<b>Comprehension</b>						
Teacher Experience	-.006	-.10	.01	.00	1, 264	.01
Grade	-.030	-.44	--	--	--	--
Mode	-.134	-2.17*	.14	.02	3, 262	1.66
SPC Total	.223	2.45*	.20	.04	4, 261	2.77*
<b>Language</b>						
Teacher Experience	-.046	-.75	.05	.00	1, 264	.57
Grade	-.214	-3.12*	--	--	--	--
Mode	-.070	-1.14	.21	.04	3, 262	3.93*
SPC Total	.050	.56	.21	.04	4, 261	3.02*
<b>Reading</b>						
Teacher Experience	-.059	-.95	.06	.00	1, 264	.91
Grade	.054	.77	--	--	--	--
Mode	-.099	-1.60	.12	.02	3, 262	1.34
SPC Total	.258	2.85*	.21	.05	4, 261	3.06*
<b>Spelling</b>						
Teacher Experience	.069	1.13	.07	.01	1, 264	1.28
Grade	-.008	-.11	--	--	--	--
Mode	-.151	-2.47*	.17	.03	3, 262	2.47
SPC Total	.290	3.24*	.26	.07	4, 261	4.54*
<b>Vocabulary</b>						
Teacher Experience	-.001	-.01	.00	.00	1, 264	.00
Grade	.126	1.81	--	--	--	--
Mode	-.045	-.73	.12	.01	3, 262	1.25
SPC Total	.177	1.94*	.17	.03	4, 261	1.89
<b>Overall NCE</b>						
Teacher Experience	-.010	-.16	.01	.00	1, 264	.03
Grade	-.025	-.36	--	--	--	--
Mode	-.143	-2.32*	.15	.02	3, 262	1.87
SPC Total	.286	3.17*	.24	.06	4, 261	3.97*
<b>Instruction in English</b>						
<b>Language</b>						
Teacher Experience	.121	1.34	.12	.02	1, 120	1.80
Grade	-.249	-2.53*	.26	.07	2, 119	4.15*
SPC Total	.208	2.29*	.32	.11	3, 118	4.62*
<b>Vocabulary</b>						
Teacher Experience	.114	1.25	.11	.01	1, 120	1.57
Grade	-.012	-.12	.11	.01	2, 119	.79
SPC Total	.243	2.61*	.26	.07	3, 118	2.82*
<b>Instruction in Spanish</b>						
<b>Language</b>						
Teacher Experience	-.069	-.82	.07	.01	1, 141	.68
Grade	-.171	-1.75	.16	.03	2, 140	1.87
SPC Total	-.136	-.87	.18	.03	3, 139	1.50
<b>Vocabulary</b>						
Teacher Experience	-.019	-.22	.02	.00	1, 141	.05
Grade	.285	2.97*	.24	.06	2, 140	4.43*
SPC Total	.045	.29	.24	.06	3, 139	2.96*

Note. \* =  $p < .05$ .

The effects of our independent variables on students' Overall NCE gain scores typified the pattern of effects on the Comprehension, Reading, Spelling, and Vocabulary gain scores. As shown in Table 2, Teacher Experience was not a reliable predictor of achievement gains on any test. Grade and Mode combined only accounted for a significant portion of variance on the Language gain scores. Even on Vocabulary gain scores, on which the final model was not significant, the t-test for transformed SPC Total scores was significant ( $p = .05$ ).

We then repeated the preceding analyses using the untransformed SPC Total scores. The only result that differed significantly was for the Vocabulary subtest: although the standardized coefficient did not change, the p-value increased from .05 to .10.

The null and marginal effects of transformed SPC Total scores on the Language and Vocabulary subtests, respectively, were other than predicted and attenuate the contention that teachers' use of the standards fosters English language development by English Language Learners. It is possible that (a) the theory proposed by Tharp et al. (2000) is not valid, or that (b) the SPC does not validly predict student achievement. Another possibility is that the language of instruction is an intervening variable: the language of instruction differentially affects student performance on tests of English language and vocabulary achievement. To test the latter possibility, we removed Mode of Instruction from the model and collapsed this three-level categorical variable (English Only, Structured English Instruction, and Bilingual) into a dichotomous variable: Instruction in English vs. Instruction in Spanish. Although some instruction occurred in English in the latter category, both teacher and student talk were primarily in Spanish. We then ran hierarchical regression models containing Teacher Experience, Grade, and the transformed SPC Total scores on the Language and Vocabulary subtests separately for Instruction in English and Instruction in Spanish. The results of these analyses (presented in the lower portion of Table 2) show that, for students instructed in English, SPC Total scores reliably predicted their achievement gains in both Language and Vocabulary. The proportion of variance accounted for by SPC Total scores, above that accounted for by Teacher Experience and Grade, in both tests was significant: an additional 4% for Language and 5% for Vocabulary. Conversely, for students taught primarily in Spanish, teachers' use of the standards was unrelated to their performance on tests that measure English language and vocabulary achievement.

## Study 2

This study examined the influence on student learning of both the pedagogy *and* organization proposed by Tharp et al. (2000). The transformed organizational model consists of multiple, simultaneous, diversified activity settings organized around inter-related learning tasks. Tharp et al. define activity settings as the organizational structures “in which children engage, and the language and problem solving that accompany them” (p. 46). A constant feature in this model is the teacher center, “Center One,” to which small, homogeneous groups of students regularly rotate, enabling the teacher to more responsively assess and assist students’ language and literacy development, contextualize new information, and establish and maintain appropriately challenging instruction. Students not with the teacher are engaged in meaningful learning activities in heterogeneous groupings. This organization is held as a necessary condition for supporting the full implementation of the proposed pedagogy. For example, neither joint

productive activity nor instructional conversation --the features fundamental to this model -- can be fully implemented in a whole-class setting.

Multiple, diversified learning activities based on the Five Standards that generate connections between new information and students' prior knowledge should provide multiple, differentiated modes of elaborating new information and, consequently, improve learning (Wittrock, 1978). Conceptualizing pedagogy and organization as two dimensions ranging from "untransformed" to "transformed" might be useful for capturing differential effects on student learning associated with variations in pedagogy and organization. This approach produced the following 2 x 2 taxonomy of transformed teaching:

*Untransformed Organization/Untransformed Pedagogy (UO/UP)*

At this level, teachers employ whole-class organization and use of the standards is limited.

*Untransformed Organization/Transformed Pedagogy (UO/TP)*

Teachers employ whole-class organization and use the standards as extensively as this organization allows.

*Transformed Organization/Untransformed Pedagogy (TO/UP)*

Teachers use simultaneous, multiple, diversified activity settings, but use of the standards is limited.

*Transformed Organization/Transformed Pedagogy (TO/TP)*

Teachers use simultaneous, multiple, diversified activity settings and use of the standards is extensive.

The hypotheses tested in this study were: (1) the proposed taxonomy could be mathematically derived from quantitative data on pedagogy and organization, and (2) student achievement gains would be greater for students whose teachers had transformed both their pedagogy and organization than for students whose teachers had not similarly transformed their teaching.

## **Method**

### **Procedure and Measures**

The data for Studies 1 and 2 were gathered concurrently from the same sample by two trained observers. While one observer used the SPC to assess implementation of the standards following the procedures reported in Study 1, the second observer used the SPC MAP, a modified version of the SPC, to separately assess (a) teachers' *direct* use of the standards at Center One, and (b) teachers' *indirect* use of the standards at each activity setting observed.

Prior to scoring with the SPC MAP, a quick assessment must be made of each activity setting in terms of its generativity (i.e., the degree to which the activity promotes conceptual processing). For example, activities that generate extended reading, writing, or discourse related to academic topics are ranked higher than activity settings relying on rehearsal (e.g., worksheets) or that generate perceptual processing of information (e.g., listening to recorded books), and activities that are recreational or social (e.g., games, puzzles, or free play). Although useful in promoting socioemotional, language, and cognitive development, these activity settings are ranked lowest. Unplanned, non-academic, transitory activity settings spontaneously formed by students are neither

ranked nor scored. When the initial ranking is completed, scoring begins with Center One and proceeds through the activity settings in descending rank order. This scoring procedure is used to produce a data set denoting activity settings of decreasing generativity for subsequent statistical analysis to identify groups of teachers similar in their pedagogy and organization. After initial ratings are made, all activity settings are monitored for the remainder of the observation period, and scores are modified as necessary.

### Results

To test our first hypothesis that taxonomy of transformed teaching could be generated representing the four variations of pedagogy and organization, the scores for each of the five standards at each activity setting were summed to form new variables. With one teacher using nine activity settings in one observation, nine variables were created (AS1 through AS9) to form a rectangular matrix for cluster analysis. If a teacher had only one activity setting during the observation, AS2 through AS9 were all zero; if a teacher had two activity settings, AS3 through AS9 were zero, and so forth. The rank data were then converted to distances, and a four-cluster solution was specified in hierarchical cluster analysis using Ward's method (metric = squared Euclidean distance). Table 3 presents the means for all activity settings for each cluster of teachers. Although the number of activity settings and their corresponding means were less than ideal, the solution provided a reasonable approximation of the proposed taxonomy.

Table 3  
*Activity Setting (AS) Means for Four- and Three-cluster Solutions*

<u>Teacher Group</u>	<u>AS1</u>		<u>AS2</u>		<u>AS3</u>		<u>AS4</u>		<u>AS5</u>	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
<u>Four Clusters</u>										
UO/UP <sup>1</sup>	6.40	0.89	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
UO/TP <sup>2</sup>	10.40	2.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TO/UP <sup>3</sup>	7.00	3.26	5.18	1.78	3.18	1.83	2.54	2.66	1.27	1.85
TO/TP <sup>4</sup>	11.00	3.08	8.80	1.79	7.60	2.30	1.80	1.48	0.80	1.79
<u>Three Clusters</u>										
UO/TP	9.07	2.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
TO/UP	6.22	3.07	4.66	1.50	3.89	1.05	3.11	2.62	1.56	1.94
TO/TP	10.63	2.45	8.38	1.51	6.00	4.17	2.00	2.45	1.63	3.29

*Note.* <sup>1</sup>Untransformed Organization/Untransformed Pedagogy; <sup>2</sup>Untransformed Organization/Transformed Pedagogy; <sup>3</sup>Transformed Organization/Untransformed Pedagogy; <sup>4</sup>Transformed Organization/Transformed Pedagogy.

Table 4  
Teacher Group Means on SAT-9 Subtests and Overall Achievement

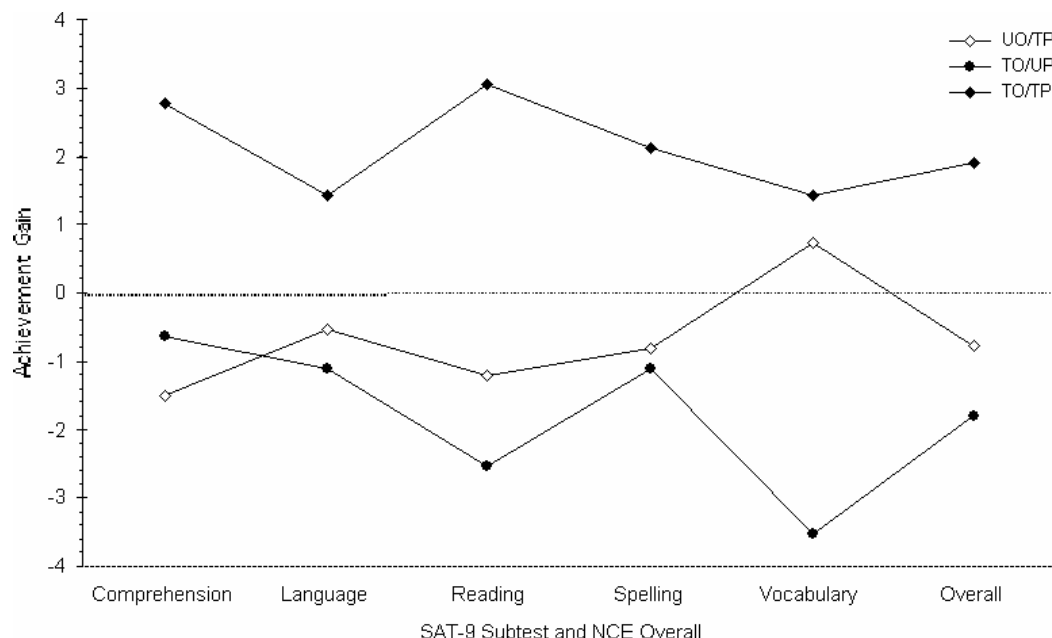
Teacher Group	TN <sup>1</sup>	SN <sup>2</sup>	Comprehension		Language		Reading		Spelling		Vocabulary		Overall NCE	
			M	SE	M	SE	M	SE	M	SE	M	SE	M	SE
UO/UP <sup>3</sup>	4	77	-1.34 <sup>a</sup>	1.25	.41 <sup>b</sup>	1.60	.52 <sup>a</sup>	1.27	-0.03 <sup>a</sup>	1.55	2.43 <sup>b</sup>	1.56	.35 <sup>a</sup>	.98
UO/TP <sup>4</sup>	3	57	-1.43 <sup>a</sup>	.98	-1.01 <sup>a</sup>	1.25	-2.21 <sup>a</sup>	1.00	-1.12 <sup>a</sup>	1.22	-1.51 <sup>a</sup>	1.22	-1.46 <sup>a</sup>	.79
TO/UP <sup>5</sup>	6	93	.54 <sup>a</sup>	1.07	-.96 <sup>a</sup>	1.37	-.07 <sup>a</sup>	1.10	-1.41 <sup>a</sup>	1.34	-1.86 <sup>a</sup>	1.35	-.75 <sup>a</sup>	.85
TO/TP <sup>6</sup>	2	39	4.23 <sup>b</sup>	1.51	3.72 <sup>b</sup>	1.93	4.66 <sup>b</sup>	1.54	5.86 <sup>b</sup>	1.88	3.72 <sup>b</sup>	1.89	4.45 <sup>b</sup>	1.19
<u>Three Clusters</u>														
UO/TP	2	29	-1.51 <sup>a</sup>	.77	-.52 <sup>b</sup>	.99	-1.22 <sup>a</sup>	.79	-.83 <sup>b</sup>	.97	.74 <sup>b</sup>	.98	-.77 <sup>a</sup>	0.62
TO/UP	9	150	-.64 <sup>b</sup>	1.74	-1.12 <sup>b</sup>	2.24	-2.53 <sup>b</sup>	1.79	-1.11 <sup>b</sup>	2.18	-3.52 <sup>b</sup>	2.21	-1.79 <sup>a</sup>	1.39
TO/TP	4	87	2.78 <sup>b</sup>	1.01	1.42 <sup>b</sup>	1.30	3.04 <sup>b</sup>	1.04	2.13 <sup>b</sup>	1.27	1.43 <sup>b</sup>	1.28	1.92 <sup>b</sup>	.80

Note. <sup>1</sup>Number of teachers in each group. <sup>2</sup>Number of students in each group. <sup>3</sup>Untransformed Organization/Untransformed Pedagogy; <sup>4</sup>Untransformed Organization/Transformed Pedagogy; <sup>5</sup>Transformed Organization/Untransformed Pedagogy; <sup>6</sup>Transformed Organization/Transformed Pedagogy. Contrasts between Teacher Groups use IO/TP as the reference group. Different subscripts indicate significant differences from TO/TP.

To test our second hypothesis, cluster membership was converted into a four-level categorical variable (Teacher Group) representing the categories of the taxonomy. Teacher Group was then entered as the independent variable into a multivariate analysis of covariance (MANCOVA) with Teacher Experience as the covariate. The five SAT-9 subtest gain scores constructed in Study 1 were the dependent variables. Because Overall NCE was a linear combination of the five subtest scores, it was not entered into the multivariate model, but was tested in subsequent univariate analyses of covariance (ANCOVA), controlling for Teacher Experience. Fisher's  $F$  ratios for all multivariate tests are based on Wilks' lambda, and effect sizes are reported as eta squared ( $\eta^2$ ). Simple contrasts were specified a priori in all tests, with the TO/TP condition as the reference group in mean comparisons with the other conditions. An alpha level of .05 was used for all statistical tests.

This model produced a significant multivariate effect for Teacher Group,  $F(15, 709) = 2.17, p = .006 (\eta^2 = .04)$ , but Teacher Experience had no reliable effect. As shown in the upper portion of Table 4, students in the TO/TP condition showed the greatest gains on all subtests and Overall NCE. Significant univariate effects for Teacher Group were found for Comprehension,  $F(3, 261) = 4.15, p = .007 (\eta^2 = .05)$ ; Reading,  $F(3, 261) = 5.94, p = .001 (\eta^2 = .06)$ ; Spelling,  $F(3, 261) = 3.63, p = .01 (\eta^2 = .04)$ ; Vocabulary,  $F(3, 261) = 3.31, p = .02 (\eta^2 = .04)$ ; and Overall NCE,  $F(3, 261) = 6.53, p < .000 (\eta^2 = .07)$ . The univariate effect for Language was not significant,  $F(3, 261) = 1.93, p = .12 (\eta^2 = .02)$ . The a priori simple contrasts indicated achievement gains in the TO/TP condition were greater than in all other conditions on all tests except Vocabulary, for which the comparison with UO/UP was not significant ( $p = .58$ ).

**Figure 1**  
*Achievement gains by students of teachers in the Untransformed Organization/Transformed Pedagogy (UO/TP), Transformed Organization/Untransformed Pedagogy (TO/UP), and Transformed Organization/Transformed Pedagogy (TO/TP) groups.*



Although these effects provide evidence that students' achievement gains were greatest when both pedagogy and organization were transformed, a closer inspection of the data revealed that the TO/TP condition contained only students in the English Only mode of instruction. Thus, this solution permitted no interpretation of the influence of the proposed instructional model on students with limited English proficiency. To address this limitation, SPC MAP data were hierarchically clustered with a three-cluster solution specified. As shown in the lower portion of Table 3, a reasonable approximation of the three highest levels of the proposed taxonomy emerged. With the Untransformed Organization/Untransformed Pedagogy group now absorbed into the three new clusters, all activity setting means decreased. Two teachers from the Bilingual classes, however, entered the TO/TP cluster.

To make a direct comparison of gains between students with higher and lower English proficiency, a 2 x 3 MANCOVA model was constructed using the two-level Mode of Instruction variable (Instruction in English vs. Instruction in Spanish) and Teacher Group as independent variables, with Teacher Experience as the covariate. Comprehension, Language, Reading, Spelling, and Vocabulary gains scores were the dependent variables. Overall NCE gain scores were analyzed separately in a 2 x 3 ANCOVA. This model found a significant multivariate effect for Teacher Group,  $F(5, 256) = 2.11, p = .02 (\eta^2 = .04)$ . The effects for Teacher Experience, Mode of Instruction, and the interaction term were not reliable. While scores in the TO/TP condition were highest in all comparisons (see Figure 1 and the lower portion of Table 4), the only significant univariate effects for Teacher Group were for Comprehension,  $F(2, 260) = 5.68, p = .004 (\eta^2 = .04)$ ; Reading,  $F(2, 260) = 5.28, p = .006 (\eta^2 = .04)$ ; and Overall NCE,  $F(2, 260) = 4.44, p = .01 (\eta^2 = .03)$ . The effect for Spelling was marginal,  $F(2, 260) = 2.20, p = .11 (\eta^2 = .02)$ . Although differences in gains were greatest between the TO/TP and TO/UP conditions for Reading and Overall NCE, simple contrasts found only the differences in gains between the TO/TP and UO/TP conditions significant on all three comparisons. This effect was likely due to the smaller standard error resulting from having only two teachers in the UO/TP group. As shown in Table 5, the main effect of Teacher Group and the non-significant interaction term indicated that, regardless of English proficiency, students in the TO/TP condition showed greater gains in comprehension and reading, and less declines overall than students whose teachers had not transformed both their pedagogy and organization.

Table 5  
*Teacher Group Means and Standard Deviations for Achievement Gains in English-Only and Bilingual Classes*

Variable	Comprehension				Reading				Overall NCE			
	High ELP <sup>1</sup>		Low ELP <sup>2</sup>		High ELP		Low ELP		High ELP		Low ELP	
Teacher Group	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE
UO/TP <sup>2</sup>	-0.45	1.03	-2.56	1.15	-0.39	1.32	-1.01	1.47	-.51	0.81	-1.10	0.91
TO/UP <sup>3</sup>	--	--	-.64	1.75	--	--	-1.12	2.24	--	--	-1.79	1.38
TO/TP <sup>4</sup>	4.30	1.50	1.26	1.35	3.72	1.93	.87	1.74	4.45	1.19	-.13	1.07

Note. <sup>1</sup>English Language Proficiency. <sup>2</sup>Untransformed Organization/Transformed Pedagogy. <sup>3</sup>Transformed Organization/Untransformed Pedagogy. <sup>4</sup>Transformed Organization/Transformed Pedagogy. Only main effects for Teacher Group were significant. See Table 4 for mean comparisons.

### Discussion

The two studies reported here examined the relationships between pedagogy, classroom organization, and achievement gains of predominantly low-income Latino students with varying levels of English proficiency. Study 1 examined the influence of teachers' use of the Standards for Effective Pedagogy during language arts instruction on gains in student achievement. Study 2 examined the relationship between the pedagogy *and* classroom organization proposed by Tharp et al. (2000) and student achievement gains. After determining a meaningful four-level taxonomy of transformed teaching, the achievement gains of students whose teachers had transformed both their pedagogy and organization were compared with students whose teachers had not similarly transformed their teaching.

The findings of Study 1 indicate a consistent, positive, and significant relationship between teachers' use of the Five Standards and students' performance on year-end standardized tests (SAT-9). After accounting for the effects of teachers' years of experience and students' grade level and English proficiency, higher SPC Total scores predicted greater achievement gains than would be predicted by students' SAT-9 scores from the previous year. This relationship was found for overall achievement gains as well for the comprehension, reading, spelling, and vocabulary SAT-9 subtests. There was no effect on the language subtest, and the relationship was only marginally significant for vocabulary. Further analysis found that teachers' use of the standards when English was the language of instruction reliably predicted English language and vocabulary achievement, whereas, as might be predicted, in classes where instruction was in Spanish, teachers' use of the standards was unrelated to student gains in *English* language and vocabulary.

The results of Study 1 may have important implications for improving the educational outcomes of all students, and especially for ELL students. Emphasizing the use of the Five Standards can help teachers design instructional activities that are meaningfully connected to students' everyday lives and that foster complex thinking. Instruction that generates extended writing and academic discourse is critical if students are to master the language of instruction and the content areas. Finally, the teacher's use

of instructional conversations during joint productive activity provides rich opportunities to assess and assist student performance at that point where assistance is most needed.

These findings should not be construed as support for providing instruction only in English for all English Language Learners. With no outcome measures of Spanish language and vocabulary achievement for this sample, the relationship between teachers' use of the standards and students' Spanish vocabulary and language development when Spanish is the language of instruction remains a question for further study.

Study 2 found that teachers' use of the Five Standards and their use of multiple, simultaneous, diversified activity settings had a significant effect on their students' achievement gains. Students whose teachers had transformed both their pedagogy and classroom organization had significantly greater overall achievement gains, as well as greater gains in comprehension, reading, spelling, and vocabulary than students whose teachers had not similarly transformed their teaching. No meaningful differences in student achievement gains were found between the three conditions in which either, or neither, pedagogy or organization was transformed. Further analysis using a three-cluster solution for grouping teachers found instruction using the transformed pedagogy and organization more effective for all students, regardless of their English proficiency. For example, students in Bilingual classes whose teachers had transformed their teaching showed greater overall achievement gains, and greater gains in comprehension and reading in particular, than students whose teachers had not transformed their teaching.

Study 2 may also have important implications for practice. Using the Five Standards to transform a classroom produces fundamental changes that spread the benefits of the standards throughout the classroom. Rows of students working in quiet isolation becomes a classroom in which small groups of students are engaged in meaningful activities accompanied by mutual assistance and rich discourse. As important, this organization allows the teacher to engage a small group of students in cognitively complex joint productive activity and to converse about the task as they work. As Gallimore and Tharp (1995) wrote, the language that accompanies joint productive activity is the major vehicle for the development of the higher cognitive processes necessary for reading comprehension. The proximity of the teacher also provides substantially more opportunities for assisting students' language use, inarguably critical for the language development of ELL students.

At first glance, the findings of Studies 1 and 2 may appear contradictory. Study 1 found teachers' use of the Five Standards predicted achievement gains. Study 2 found that, for all practical purposes, transformed pedagogy alone was not more effective than untransformed pedagogy. The simple explanation in this study is that the four teachers in the three-cluster solution who used the standards more effectively, both directly and indirectly, contributed substantially to the predictive validity of SPC scores. A more complex explanation rests in how the SPC and the SPC MAP differ: the SPC assesses pedagogy and organization combined and indicates how intensively the standards are used; the SPC MAP assesses pedagogy and organization uniquely, and indicates how intensively and extensively the standards are used. The findings here suggest that transforming pedagogy is necessary for improving the achievement of ELL students, but it alone is not sufficient. Both pedagogy and organization must be transformed to maximize the effectiveness of the Standards for Effective Pedagogy.

Any implications for practice based on these two studies can only be offered tentatively. These were correlational studies of the relationship between teachers' use of the standards, classroom organization, and student achievement gains. Although student performance from the prior year was used to control for individual differences contributing to gains, this changes neither the design of the study nor the inferences that can be made based on the findings. If cognitivist assertions that new information must be sufficiently processed to be learned are sound, then the findings suggest that teaching transformed as proposed may be an effective instructional model for promoting the elaboration of instructional content which, in turn, may improve learning. These findings do not offer a simple solution to a complex problem. In fact, the complexity of the solution offered by Tharp et al. (2000) matches the complexity of the problem. In order for these standards to be used effectively, teachers will have to re-envision their teaching practices. The classroom environment necessary to reform education is fundamentally different from what most teachers experienced as students, and for teachers to be successful in taking on new roles and changing practices that have withstood decades of reform efforts, they will need a new articulation of the role of the teacher and clear standards for transforming their teaching. Likewise, principals and district administrators must re-envision their roles to that of assisting teachers in their efforts to improve their practice (Tharp et al., 2000; Tharp & Gallimore, 1988).

Such a transformation of American classrooms may not be feasible. Many argue that the creation of a pedagogical system that reliably creates the setting events that in turn produce language and cognitive development is impractical. In this view, the ponderous bureaucracies, institutional inertias, and limitations on the competence of schools make such a vision impossible to realize. The evidence available suggests otherwise. In a long-term demonstration of scalability, the Kamehameha Early Education Program (KEEP) operated a statewide system of such classrooms in public schools throughout the Hawaiian Islands for more than 20 years (Calkins et al., 1989; Jordan et al., 1985; Klein, 1988; Yap et al., 1988).

This paper has offered primarily a cognitive explanation for the influence of the Five Standards on student learning. The focus on the links between teaching, extended cognitive and verbal elaborations, and learning represents a fundamental congruence between cognitive perspectives on learning and sociocultural theories of crucial processes in effective pedagogy. For example, Hart & Risley (1995; 1999) have recently demonstrated conclusively that sheer quantity of elaborated verbal exchanges between caregiver and young children from birth to three years predicts school success and measured IQ on entry into school and extending into the elementary school years. From this perspective, the Standards for Effective Pedagogy constitute an effective instructional design to provide the classroom settings that maximize elaborated verbal exchanges during shared meaningful activities with more knowledgeable participants.

Thus, both cognitive and sociocultural theories provide compatible explanations of the findings of this study, and predictive hypotheses for future work. Our laboratories are currently planning true-experimental designs to replicate the findings in this study and further test the efficacy of the Standards for Effective Pedagogy model, as well as preparing for scaling up to a variety of linguistic and cultural communities.

## References

- Antill, E., & Tharp, R. G. (1974). A comparison of the industriousness levels of KEEP and public school students (Technical Report 55). Honolulu: Kamehameha Schools.
- Au, K. H., Crowell, D. C., Jordan, C., Sloat, K. C. M., Speidel, G. E., Klein, T. W., & Tharp, R. G. (1986). Development and implementation of the KEEP reading program. In J. Orasanu (Ed.), *Reading comprehension: From research to practice* (pp. 235-252). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Au, K. H., & Jordan, C. (1981). Teaching reading to Hawaiian children: Finding a culturally appropriate solution. In H. Trueba, G. P. Guthrie, & K. H. Au (Eds.), *Culture in the bilingual classroom: Studies in classroom ethnography* (pp. 139-152). Rowley, MA: Newbury House.
- Baddeley, A. D. (1990). Human memory: Theory and practice. Boston: Allyn and Bacon.
- Bandura, A. (1977). Social learning theory. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1986). Social foundations of thought and action. New York: Prentice Hall.
- Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology*, 45, 1017-1028.
- Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology*, 41, 586-598.
- Barnhardt, C. (1982). Tuning-in: Athabaskan teachers and Athabaskan students. In R. Barnhardt (Ed.), *Cross-cultural issues in Alaskan education* (Vol. 2). Fairbanks, AL: Center for Cross-Cultural Studies.
- Begay, S., Dick, G. S., Estell, D. W., Estell, J., McCarty, T. L., & Sells, A. (1995). Change from the inside out: A story of transformation in Navajo community school. *Bilingual Research Journal*, 19(1), 121-139.
- Blum, R. (1998, November). Lessons from Alaska and more. Paper presented at the Greenland Colloquium, Bellingham, WA.
- Blum, R. (n.d.). Closing the gap: Alaska. Portland, OR: Northwest Regional Educational Laboratory.
- Calkins, R., Klein, T. W., Guili, C., Au, K., Cunningham, L., & Springer, R. (1989). Kamehameha Elementary Education Program: An evaluative summary (Technical report). Honolulu: The Kamehameha Schools Center for the Development of Early Education.
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Dansereau, D. F. (1988). Cooperative learning strategies. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation* (pp. 103-120). Orlando, FL: Academic Press.
- Demmert, W. (1994). Blueprints for Indian education: Languages and cultures. ERIC Digest.

- Demmert, W. (2001). Improving academic performance among Native American students: A review of the research literature. Charleston, WV: ERIC Clearinghouse on Rural Education and Small Schools.
- Dick, G. S., Estell, D. W., & McCarty, T. L. (1994). Saad Naakih Bee'enootihji Na'alkaa: Restructuring the teaching of language and literacy in Navajo community school. *Journal of American Indian Education*, 33, 31-46.
- Dodd, D. H., & White, R. M. (1980). Cognition, mental structures and processes. Boston: Allyn and Bacon, Inc.
- Doherty, R. W., Hilberg, R. S., Epaloose, G., & Tharp, R. G. (2002). Standards Performance Continuum: Development and validation of a measure of effective pedagogy. *Journal of Educational Research*, 96(2), 78-89.
- Doherty, R. W., & Pinal, A. (2002, November). Joint productive activity, cognitive reading strategies, and achievement. Paper presented at the Annual Conference of the National Council of Teachers of English, Atlanta, GA.
- Estrada, P. (2000, October). Pedagogy, professional development and reading performance in six linguistically diverse classrooms. Paper presented at the UC ACCORD First Annual Conference on Education & Equity: Research, Policy, & Practice, San Jose, CA.
- Gallimore, R., & Tharp, R. G. (1995). Teaching mind in society: Teaching, schooling, and literate discourse. In L. Moll (Ed.), *Vygotsky and education: Instructional implications and applications* (pp. 175-205). New York: Cambridge University Press.
- Gallimore, R., Tharp, R. G., Sloat, K. C. M., Klien, T. W., & Troy, M. E. (1982). Analysis of reading achievement test results for the Kamehameha Early Education Project: 1972-1979 (Technical Report No. 102). Honolulu: Kamehameha Schools, Bishop Estate, Center for the Development of Early Education.
- Goncii, A., & Rogoff, B. (1998). Children's categorization with varying adult support. *American Educational Research Journal*, 35(2), 333-349.
- Hart, B., & Risley, T. R. (1995). Meaningful differences in the everyday experience of young American children. Baltimore, MD: Brookes Publishing Company, Inc.
- Hart, B., & Risley, T. R. (1999). The social world of children learning to talk. Baltimore, MD: Baltimore, MD: Brookes Publishing Company, Inc.
- Harwell, M. R., & Gatti, G. G. (2001). Rescaling ordinal data to interval data in educational research. *Review of Educational Research*, 71(1), 105-131.
- Hilberg, R., S., Tharp, R. G., & DeGeest, L. (2000). The efficacy of CREDE's standards-based instruction in American Indian mathematics classes. *Equity and Excellence in Education*, 33(2), 32-39.
- Hilberg, R. S., Doherty, R. W., Epaloose, G., & Tharp, R. G. (2001). CREDE project 5.6 final report. Santa Cruz: University of California, Center for Research on Education, Diversity & Excellence (CREDE).
- Howes, M. B. (1990). The psychology of human cognition. New York: Pergamon Press.
- Jordan, C. (1995). Creating cultures of schooling: Historical and conceptual background of the KEEP/Rough Rock Project. *The Bilingual Research Journal*, 19(1), 83-100.

- Jordan, C., Tharp, R., & Vogt, L. (1985). Compatibility of classroom and culture: General principles with Navajo and Hawaiian instances (Working paper No. 18). Honolulu: Kamehameha Schools, Bishop Estate, Center for the Development of Early Education.
- Klein, T. W. (1988). Program evaluation of the Kamehameha Elementary Education Program's reading curriculum, in Hawai'i public schools: The cohort analysis 1978-1986. Honolulu: Kamehameha Schools, Bishop Estate, Center for the Development of Early Education.
- Klein, T. W., & Calkins, R. (1988). The typical pattern of student achievement in KEEP from grade one to grade three: A look at different achievement measures across the years (Technical report). Honolulu: Kamehameha Schools, Bishop Estate, Center for the Development of Early Education.
- Kolers, P. A., & Brison, S. J. (1984). Commentary: On pictures, words, and their mental representations. *Journal of Verbal Learning and Verbal Behavior*, 23, 105-113.
- Lipka, J. (1986). School-community partnerships in rural Alaska. *Rural Educator*, 7(3), 11-14.
- Lipka, J. (1990). Integrating cultural form and content in one Yup'ik Eskimo classroom: A case study. *Canadian Journal of Native Education*, 17, 18-32.
- Lipka, J. (1994). Culturally negotiated schooling: Toward a Yup'ik mathematics. *Journal of American Indian Education*, Spring, 14-30.
- Lockhart, R. S., & Craik, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology*, 44, 87-122.
- McIntyre, E., Kyle, D. W., Hovda, R. A., & Stone, N. (1999). Nongraded primary programs: Reform for Kentucky's children. *Journal of Education for Students Placed at Risk*, 4(1), 47-64.
- McIntyre, E., Rosebery, A., & Gonzalez, N. (2001). Classroom diversity: Connecting curriculum to students' lives. Westport, CT: Heinemann.
- McIntyre, E., & Stone, N. J. (1998). Culturally contextualized instruction in Appalachian-Descent and African-American classrooms. *National Reading Conference Yearbook*, 47(1), 209-220.
- Padron, Y. N., & Waxman, H. C. (1999). Classroom observations of the Five Standards of Effective Teaching in urban classrooms with English language learners. *Teaching and Change*, 7(1), 79-100.
- Preston, V. (1991). Mathematics and science curricula in elementary and secondary education for American Indian and Alaska Native students. Washington, DC: U.S. Department of Education, Indian Nations At Risk Task Force.
- Saunders, W., & Goldenberg, C. (1999). The effects of instructional conversations and literature logs on the story comprehension and thematic understanding of English proficient and limited English proficient students. Santa Cruz, CA: Center for Research on Education, Diversity & Excellence, University of California.
- Saunders, W., & Goldenberg, C. (in press). The effects of an instructional conversation on transition students' concepts of friendship and story comprehension. In R.

- Horowitz (Ed.), *The evolution of talk about text: Knowing the world through classroom discourse*. Newark, DE: International Reading Association.
- Schraw, G., & Sperling-Dennison, R. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460-475.
- Schvaneveldt, R. W., & McDonald, J. E. (1981). Semantic context and the encoding of words: Evidence for two modes of stimulus analysis. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 9, 544-555.
- Scollon, R. (1981). *Tempo, density and silence: Rhythms in ordinary talk*. Fairbanks: University of Alaska, Center for Cross-Cultural Studies.
- Sells, A. (1994, April). Initiating and sustaining positive change: Navajo tribal perspective on the KEEP-Rough Rock experience. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- Soar, R. S. (1978). Problems in analyzing process-product relationships in studies of teacher effectiveness. *Journal of Educational Research*, 160(4), 96-116.
- Swisher, K., & Deyhle, D. (1987). Styles of learning and learning of styles: Educational conflicts for American Indian/Alaskan Native youth. *Journal of Multilingual and Multicultural Development*, 8, 345-360.
- Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistics* (2nd ed.). New York: Harper Colins Publishers, Inc.
- Tharp, R. G. (1982). The effective instruction of comprehension: Results and description of the Kamehameha Early Education Program. *Reading Research Quarterly*, 17(4), 503-527.
- Tharp, R. G. (1994). Research knowledge and policy issues in cultural diversity and education. In B. McLeod (Ed.), *Language and learning: Educating linguistically diverse students* (pp. 129-167). Albany, NY: SUNY Press.
- Tharp, R. G. (1997). *From at-risk to excellence: Research, theory, and principles for practice* (Research Report No. 1). Washington, DC: Center for Applied Linguistics and Center for Research on Education, Diversity, and Excellence.
- Tharp, R. G., Dalton, S. S., & Yamauchi, L. A. (1994). Principles for culturally compatible Native American education. *Journal of Navajo Education*, 11(3), 21-17.
- Tharp, R. G., Estrada, P., Dalton, S. S., & Yamauchi, L. (2000). *Teaching transformed: Achieving excellence, fairness, inclusion, and harmony*. Boulder, CO: Westview Press.
- Tharp, R. G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context*. New York: Cambridge University Press.
- Tharp, R. G., Jordan, C., Speidel, G. E., Au, K. H., Klein, T. W., Calkins, R. P., Sloat, K. C. M., & Gallimore, R. (1984). Product and process in applied developmental research: Education and the children of a minority. In M. E. Lamb & A. L. Brown & B. Rogoff (Eds.), *Advances in developmental psychology* (Vol. 3, pp. 91-144). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Tulving, E. (1983). *Elements of episodic memory*. London: Oxford University Press.
- Tulving, E., & Thompson, D. M. (1971). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 353-373.

- Vogt, L. A., Jordan, C., & Tharp, R. G. (1992). Explaining school failure, producing school success: Two cases. In E. Jacob & C. Jordan (Eds.), *Minority education: Anthropological perspectives* (Vol. 18, pp. 53-66). Norwood, NJ: Ablex.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole & V. John-Steiner & S. Scribner & E. Souberman, Trans.). Cambridge, MA: Harvard University Press.
- Waxman, H. C., Huang, S.-L. Y., Anderson, L., & Weinstein, T. (1997). Classroom process differences in inner-city elementary schools. *Journal of Educational Research*, 91(1), 49-59.
- Waxman, H. C., & Huang, S. L. (1997). Classroom instruction and learning environment differences between effective and ineffective urban elementary schools for African American students. *Urban Education*, 32(1), 7-44.
- Waxman, H. C., Huang, S. L., & Padron, Y. N. (1995). Investigating the pedagogy of poverty in inner-city middle level schools. *Research in Middle Level Education*, 18(2), 1-22.
- Wittrock, M. C. (1978). The cognitive movement in instruction. *Educational Psychology*, 13, 15-29.
- Wittrock, M. C. (1986). Students' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 297-314). New York: Macmillan.
- Yamauchi, L. A., & Tharp, R. G. (1995). Culturally compatible conversations in Native American classrooms. *Linguistics and Education*, 7, 349-367.
- Yap, K. O., Estes, G. D., & Nickel, P. R. (1988). A summative evaluation of the Kamehameha Early Education Program (Draft report). Honolulu, HI: Northwest Regional Educational Laboratory.

### **Endnote**

This work was supported under the Education Research and Development Program, PR/Award No. R306A60001, the Center for Research on Education, Diversity & Excellence (CREDE), as administered by the Office of Educational Research and Improvement (OERI), National Institute on the Education of At-Risk Students (NIEARS), US Department of Education (USDOE). The contents, findings and opinions expressed here are those of the author and do not necessarily represent the positions or policies of OERI, NIEARS, or the USDOE.