Effective Instruction in Culturally Diverse Early Childhood Settings:
Using the CREDE Standards for Effective Pedagogy with Young Children

Lois A. Yamauchi
Seongah Im
Nanette Schonleber

University of Hawai‘i

Effective Instruction in Culturally Diverse Early Childhood Settings:
Using the CREDE Standards for Effective Pedagogy with Young Children

As the US population becomes increasingly more culturally and linguistically diverse (Artiles, 2003; Datnow, Stringfield, & Castellano, 2005), early childhood (EC) educators need to consider how to best meet the needs of children from diverse backgrounds. For example, approximately 30% of children in Head Start programs speak a language other than English as their primary language, and these numbers are growing (U.S. Department of Health and Human Services, 2006). In addition, there are over 140 languages and home cultures represented in Head Start programs nationwide.

When children arrive with diverse cultural and linguistic backgrounds, many EC educators are uncertain about how to simultaneously support the development of the home and school languages and cultures and seek strategies to effectively meet the needs of the diverse children in their care (Head Start, 2008). Similarly, educators in EC programs for Alaska Natives, American Indians and Native Hawaiians often struggle to maintain their communities’ indigenous home languages. EC educators in such programs report that they would like teaching strategies that can help them to revitalize and maintain their indigenous cultures and languages while preparing their children to succeed in majority culture settings (Schonleber, 2006).

While many EC educators use national and state standards to help guide their planning and practice, these guidelines do not necessarily help teachers improve their instruction, nor do they specifically address the needs of diverse students (Boutte, 2008; Kidd, Sanchez & Thorp, 2007; Souto-Manning & Dice, 2007). In teacher preparation programs, EC educators may learn about methods and pedagogical strategies that are effective with multi-ethnic children. However, in these cases, educators often report
that they did not have opportunities to practice these strategies as students (Potter, 2007).

This is not a trivial problem. A substantial body of research indicates that knowledge of content and knowledge of pedagogy are both important factors in the achievement of high academic outcomes for children (Angelo, 1993; Darling-Hammond 1998; Tharp & Entz, 2003). How educators teach is as important as what they teach (Dowsett, Husten, Imes, & Gennetian, 2008; Tharp, Estrada, Dalton, & Yamauchi, 2000).

The Center for Research on Education, Diversity, and Education (CREDE) Standards for Effective Pedagogy are strategies for best practices developed from over 30 years of research on K-12 students from diverse cultural and linguistic groups (Tharp et al., 2000). This study examined how the CREDE Standards could be adapted for use in an EC setting, including changes to an instrument used to measure the Standards in EC education.

**Theoretical Framework**

The CREDE Standards derive from Vygotsky’s (1978) contention that all higher psychological functioning has its roots in social interaction. They are:

- **Standard 1 – Joint Productive Activity.** Teachers and students collaborating to create joint products.
- **Standard 2 -- Language and Literacy Development.** Developing competence in the language(s) of instruction throughout the day.
- **Standard 3 -- Contextualization.** Embedding instruction in the interests, experiences and skills of students’ families and communities.
- **Standard 4 -- Complex Thinking.** Challenging students toward cognitive complexity.
- **Standard 5 -- Instructional Conversation.** Engaging students through dialogue.
- **Standard 6 – Modeling.** Promoting students’ observational learning.
• Standard 7 – Student Directed Activity. Encouraging student decision-making.¹

According to Vygotsky (1978), children’s interactions with others form the basis of higher-level thought. For example, children and caregivers may talk about what they are doing, as they prepare food together. Over time, children construct understandings of those ideas and actions to develop their own ways of thinking about food, food preparation, health, and other concepts discussed in these conversations. In this way, caregivers assist children in performing the routines and traditions of their households and cultures that may later be completed by the children on their own. In an educational context, early childhood educators help children to think about and perform activities in that setting that over time, the children may later enact independently. The CREDE Standards provide guidelines to maximize teachers’ assistance of children in this process.

Evidence for Effectiveness of the CREDE Standards

Recent investigations have examined the effects of teachers’ use of the CREDE Standards through varied research methods including case studies of multiple classrooms, short-term random designs, within-classrooms quasi-experimental designs, and longitudinal studies of entire schools (Doherty et al., 2003; Estrada, 2004, 2005; Hilberg, Tharp, & DeGeest, 2000; Saunders & Goldenberg 1999, 2001). Some studies focused on one or two particular Standards, whereas, others looked at a combination of the CREDE Standards compared to more traditional pedagogical strategies. A number of investigations focused on the relationship between student outcomes and teachers who were either high or low implementers of the Standards. Results of all of these studies indicated positive relationships between use of the CREDE Standards and

¹ Standards 1-5 are considered “generic” principles that were developed for all groups of students. Standards 6 and 7 were developed specifically for indigenous students (Tharp, 2006).
student academic achievement.

For example, Estrada’s (2004, 2005) three years of research showed a consistent relationship between teachers’ use of the CREDE Standards and academic achievement in Grades 1 and 4. She compared student achievement in classrooms of teachers who were strong and weak implementers of the CREDE Standards. The majority of students in strong implementers’ classrooms achieved grade level standards in reading, outperforming those in classrooms of weak implementers. Hilberg, Tharp, and DeGeest (2000) studied the effects of the CREDE Standards on mathematics achievement among American Indian students. The Grade 8 students were randomly assigned to either a CREDE Standards or control condition for mathematics instruction. Results indicated that students who received the CREDE Standards instruction outperformed those in the control group on tests of conceptual learning both at the end of the unit and two weeks later.

A series of studies investigated student outcomes related to teachers’ use of Standard 5, Instructional Conversation. The reading comprehension of students whose teachers used Instructional Conversation was compared to that of those whose teachers used more traditional instruction to teach the same lesson. Both groups demonstrated equal comprehension of more factual understanding of the story (e.g., characters and plot). However, the Instructional Conversation group showed a greater understanding of the story’s theme. These studies were considered by the US Department of Education What Works Clearinghouse in their review of 73 studies focused on language development for English language learners (Institute of Educational Sciences, 2007). The reviewers ranked this use of Instructional Conversation as the most effective method for reading achievement and the second most effective for English language development among programs focused on improving English language literacy among elementary
students who were English language learners (Institute of Educational Sciences, 2007).

Several studies conducted at the CREDE demonstration schools in California showed consistent positive effects of the CREDE Standards. The demonstration schools served predominantly low-income Latino English language learners. The researchers used the Standards Performance Continuum (Doherty, Hilberg, Epaloose, & Tharp, 2002) to measure teachers’ use of the CREDE Standards (See Appendix A). Results indicated that across two years, teachers’ overall use of the Standards predicted standardized test (SAT-9) gains in comprehension, reading, spelling, and vocabulary (Doherty et al., 2003). Students whose teachers were high implementers of the CREDE Standards and who organized their classrooms as multiple, simultaneous, diversified activity settings as specified by the pedagogy model (Tharp et al., 2000), showed greater SAT-9 gains than those in classrooms that were not similarly transformed. Doherty, Hilberg, and Lee (2004) replicated these findings when they compared student achievement in a CREDE demonstration school, where there was high implementation of the Standards, to that of students in another school in a comparable neighboring area.

The CREDE Standards and Early Childhood Education

Although research on the CREDE Standards has been conducted in numerous K-12 classrooms, little research has focused on their use in EC education. In this study, we were interested in how the CREDE Standards could be used with children ages 2- to 5-years-old. For two years, EC educators participated in professional development that focused on using the CREDE Standards in ways that met the developmental needs of young learners. This collaboration resulted in a greater understanding of how EC educators could use the Standards with young children and changes to an instrument used to measure the Standards in EC settings.
Methods

Setting

Our research was conducted at a preschool serving children of students, faculty, and staff at a large, state-funded university. The background of these families reflects the diversity of the University community. Approximately 42% of the children are Asian or Asian American, 40% are European or European American, 15% are Native Hawaiian or Pacific Islander, and the remainder represents a combination of African American and Latino backgrounds. School administrators estimate that approximately 32% of all children at the Center speak English as a second language.

Participants

Participants included the 13 preschool teachers, 2 administrators, and approximately 100 students, ages 2-5. All of the adult participants were female, except for the Center director. The educators’ ages ranged from 20- to 46-years-old. At the start of the study, there was also a wide range of years, between 1 and 17 that the educators had been teaching. The ethnic backgrounds represented by the teachers included Asian American, Native Hawaiian, and European American. All educators held at least a bachelor’s degree in EC education or a related field. One teacher held a master’s degree, and two were enrolled in master’s programs.

Instrumentation

The CREDE ECE-7 is an instrument used to measure use of the CREDE Standards in EC settings (see Appendix B). The CREDE ECE-7 is based on the Standards Performance Continuum, an instrument developed to measure K-12 teachers’ use of the CREDE Standards. The first author and the participating EC educators developed the CREDE ECE-7 during professional development sessions throughout the two years. The first author facilitated discussions about the CREDE Standards and their
use with young children. The group watched selected videotaped excerpts of the teachers’ classrooms and used the Standards Performance Continuum to rate the extent to which the videorecordings reflected use of the CREDE Standards. The group also discussed ways in which criteria from the Standards Performance Continuum needed to be changed in order for it to fit the needs of young learners. Examples of how these criteria changed is reported in the Results section of this paper.

When using the Standards Performance Continuum, coders assign a score of 0 (Not Observed) to 3 (Enacted) to represent the extent to which a lesson reflects use of each CREDE Standard. We originally used the same 0-3 scale for the CREDE ECE-7. However, data analysis revealed that there was little variability of scores. In addition, the coders and teachers felt that the 0-3 scoring did not capture the variation that existed in instructional settings. In particular, they felt that when they assigned a score of 3, this score could be considered either a “high 3” or a “low 3.” In response to these comments, we created an additional level of coding such that each Standard would be coded on a scale of 0-4.

Data Sources

Ratings of videorecorded instruction. Each teacher was videorecorded six times annually for two years. The videorecording took place during a morning instructional period and lasted approximately 60-90 minutes. During the first semester, the Center’s assistant director and the teachers decided when each person would be videorecorded. Thereafter, the order of the videorecordings was randomly determined.

Trained coders rated the 156 videorecordings using the CREDE ECE-7 twice, once using the first version of the instrument with a scale of 0-3, and a second time, using the final version with the scale of 0 to 4. Two coders rated each videorecording
independently and then met to discuss discrepancies. During these meetings, they came to consensus on the final ratings.

*Participant observations.* The first author took field notes during and after the professional development sessions and other meetings about the CREDE Standards and use in EC.

*Data Analysis*

We evaluated the extent of agreement between the two raters’ coding. To determine whether the reliability of scores varied among the seven CREDE Standards, we calculated the inter-rater reliability coefficient of each Standard, as well as the overall coefficient across the seven Standards. Three different methods were employed to give us an idea about different sources of disagreement. First, we measured the proportion of the absolute agreement of the two raters’ coding on the Standards Performance Continuum categories. This index, measuring the proportion of the agreement on the ratings, is denoted as $p$ and ranged from 0 to 1.

Next, we calculated Cohen’s (1960) Kappa ($k$), a chance-corrected measure of agreement that discounts the observed proportion of agreement by the expected level of agreement, given the marginal distributions of the raters’ agreement and the assumption that the raters’ coding is statistically independent. This $k$ coefficient ranges from -1 to 1, with negative or zero coefficients meaning no agreement. The coefficient of 0.6 or higher considered high.

Finally, the generalizability coefficient ($g$), ranging 0 to 1, of the generalizability theory (Shavelson & Webb, 1991) was calculated. The higher $g$ coefficient means a smaller amount of variability between the raters’ coding, i.e., greater consistence. The coefficient of 0.8 or higher is considered high.
We analyzed the qualitative data by creating a narrative about how the CREDE Standards could be used in EC settings and regarding the ways in which the CREDE ECE-7 developed from the Standards Performance Continuum.

Results

How the CREDE Standards Can be Used in Early Childhood

Language development issues. When the UHMCC educators first learned to use the Standards Performance Continuum, they were concerned that the language-oriented Standards (Language and Literacy Development and Instructional Conversation [IC]) were not appropriate for their work with young children. The Standards Performance Continuum criteria for both of these Standards assumed that students used oral language frequently. For example, the criteria for the highest level of IC included students talking more than the teacher, and the educators were not sure if this criterion was appropriate for young children. When the group watched videorecordings of discussions in UHMCC classrooms, they decided that ICs could occur often in EC settings if the focus included both verbal and non-verbal forms of communication. The criteria for IC on the CREDE ECE-7 were written to reflect these understandings.

Moving away from “academic” goals. Another way that educators felt that the rubric to measure use of the CREDE Standards needed to be revised to better fit the development of young children was to delete mention of academic goals. Although the Standards Performance Continuum focused on instruction toward academic goals, the UHMCC educators felt that many developmentally appropriate objectives in EC education encompass goals and skills that might not be considered academic.

Inclusion of the two indigenous Standards. Originally, the project focused on the first five CREDE Standards. However, when UHMCC educators heard about the two additional Standards developed for indigenous educators: Modeling and Self-Directed
Learning, the teachers suggested that they be included in the CREDE model for all young learners. The educators viewed these ideas as universally important in EC education.

*Ratings of the Videorecorded Lessons and Further Instrument Development*

The overall levels of inter-rater reliability using the 0-3 scale were intermediate. Figure 1 presents the sizes of inter-rater reliability using the three methods. The $p$ coefficients of the seven Standards, measuring the absolute agreement between the two raters’ coding, ranged from 0.54 (Standard 2) to 0.88 (Standard 8). The overall coefficient was 0.71. The $k$ coefficients of the seven Standards, measuring the degree of the agreement between the two raters' coding after taking into account the agreement occurring by chance, were somewhat low. The lowest association was 0.22 for Standard 5. The highest one was for Standard 1, which was still intermediate, 0.58. The overall agreement was 0.49. Finally, the $g$ coefficient, reflecting the relative variability between the raters’ coding to the variability within each rater, was calculated. The $g$ coefficients of the seven Standards ranged from 0.50 (Standard 5) to 0.73 (Standard 3). The overall coefficient across the Standards was 0.67.

The $g$ coefficients did not reach 0.8, which is a common standard required for reliable measures. This implies that both raters coded the different teachers’
performance almost equally, indicating that there was little variability within each rater’s coding. For example, the proportion of the agreement of the two raters’ coding was high for Standard 6 and Standard 7. However, the g coefficients were low, suggesting that many teachers received similar or equal scores. The variability among the different teachers may have been too low, compared to the relatively high variability between the two raters’ coding.

Results from this further analysis using the new rubric will be compared to the current one and reported in our presentation.

Discussion
References

Angelo, 1993


Boutte, 2008


Darling-Hammond 1998


Dowsett, Husten, Imes, & Gennetian, 2008


Head Start, 2008

Observational research in U.S. classrooms: New approaches for understanding cultural and linguistic diversity (pp. 48-71). Cambridge, UK: Cambridge University Press.


Kidd, Sanchez & Thorp, 2007

Potter, 2007

Schonleber, 2006


Souto-Manning & Dice, 2007


U.S. Department of Health and Human Services, 2006