

Classroom Quality and Time Allocation in
Tulsa's Early Childhood Programs

Deborah Phillips
Psychology Department
Georgetown University

William T. Gormley, Jr.
Public Policy Institute
Georgetown University

Amy Lowenstein
Psychology Department
Georgetown University

Paper presented at the biennial meetings of the Society for Research in Child
Development, Boston, MA, March 30, 2007

ACKNOWLEDGMENTS

We would like to thank the research assistants from the University of Tulsa who worked tirelessly to collect the classroom data (Lisa Adams, Sara Campbell, Pam Denton, Sara House, Nicole Larson, Katie Livingood, Jaclyn Niebergall, and Melissa Young), our research assistants at Georgetown University who coded and analyzed the data (especially Dan Cullinan and Mireya Almazan), and the leadership, principals, and teachers in the Tulsa Pre-K and Head Start programs without whom this project would not have been possible. We would also like to thank Diane Early for sharing the NCEDL data with us. Finally, we would like to thank the Spencer Foundation, the Foundation for Child Development, the David and Lucile Packard Foundation, and the A.L. Mailman Family Foundation for their generous financial support. All conclusions drawn in this paper are exclusively those of the authors.

INTRODUCTION

In recent years, several states have established universal pre-Kindergarten (UPK) programs in an effort to enhance the school readiness of young children. Oklahoma took this step in 1998. The Oklahoma UPK program is especially interesting because it reaches more four-year-olds than any other program in the U.S. (U.S. GAO, 2004) and because it employs an unusually rigorous set of classroom quality requirements. Under state law, every lead teacher in a state-funded pre-K classroom must have a bachelor's degree with a teaching certificate in early childhood education. Classrooms must maintain 10/1 child/staff ratios, keep total class size at 20 or fewer children, and lead teachers are paid public school wages and benefits. The majority of pre-K teachers have undergone mandated training in early childhood reading and math instruction. Instructional content is guided by a set of state learning standards, but there are no specific curriculum requirements.

On two separate occasions (2001 and 2003 waves of data collection), we have documented impressive learning gains for four-year-olds who participated in Oklahoma's UPK program (Gormley & Phillips, 2005; Gormley & Gayer, 2005; Gormley, Gayer, Phillips, & Dawson, 2005). The effect sizes of 0.79 for pre-reading skills, 0.64 for pre-writing skills, and 0.38 for pre-math skills exceed those reported for other state-funded pre-K (Gilliam & Zigler, 2001, Xiang & Schweinhart, 2002), for pre-K programs generally (Magnuson et al., 2007), and for high-quality child care programs (NICHD and Duncan, 2003). In fact, they fall just short of the substantial effect sizes associated with

the Abecedarian project (Campbell, et al., 2002; Ramey et al., 2000) and the Perry Preschool program (Schweinhart et al., 2005). These gains, based on tests administered to kindergarten and pre-K children in the fall of 2003 in Tulsa, Oklahoma, extend to diverse racial and ethnic groups and to children from diverse socioeconomic backgrounds. Gains for Hispanic children are particularly impressive, as are gains for disadvantaged children.

Although these findings are extremely encouraging, we do not know what accounts for them. While Oklahoma, along with about half of all state pre-K programs (Early et al., 2006; Barnett, Hustedt, Robin, & Schulman, 2005), requires the BA degree for all pre-K teachers, as well as an early childhood certificate, the teachers vary widely in their years of experience, in the college attended, and in the coursework they took. With this variation in mind, we do not know which teachers are more successful than others, nor do we know if some teacher characteristics, attitudes, and approaches to instruction matter more than others as children prepare for kindergarten. We have also not previously observed the Tulsa pre-K classrooms, leaving open questions regarding variation in classroom quality and children's exposure to academic instruction, which are central to understanding the processes that underlie the documented success of this program.

This paper represents the first report from the 2006 wave of data collection on the Tulsa Pre-K program, which included observational data on virtually all pre-K classrooms, including the Head Start collaboratives, and extensive data from teachers on

their educational and occupational histories and attitudes about children. Specifically, we will provide a descriptive picture of pre-K (and Head Start) classrooms in Tulsa, OK and situate this portrait within the broader context of pre-K (and Head Start) quality in the U.S. We will also examine the relation between selected attributes of pre-K teachers (teacher education and training, years of experience, curricular choice, and Spanish language skills), the distribution of classroom time spent on academic content, and the emotional and instructional quality of the classroom. In subsequent analyses, we will examine the relation between these teacher and classroom features, and the children's Fall 2006 achievement test scores, as well as teacher ratings of their social-emotional behavior.

Parents and policymakers want to know which early childhood programs offer the highest-quality experiences for young children and, as such, provide the greatest support for their early development. Prior research suggests that the typical school-based preschool program is of higher quality than the typical child care program, particularly among programs serving low-income children (Goodson & Moss, 1992; Phillips, Voran, Kisker, Howes, & Whitebook, 1994). Head Start programs tend to fall in the middle of this spectrum, thus protecting children from the very poor-quality environments that can be found in child care, but perhaps failing to provide – on a consistent basis – the exemplary experiences that are found in some pre-K programs (Barnett et al. 2005; Goodson & Moss, 1992).

Nevertheless, emerging descriptive data indicate that pre-K programs, like child care and Head Start, are characterized by extensive variation. This extends beyond variation in the teacher and classroom structural requirements discussed above, to variation in what children actually experience in these different settings (Early et al., 2006; Clifford et al., 2005; Gilliam & Marchesseault, 2005). For example, in a recent report of pre-K practices in 11 states (Early et al., 2006), 12% of classrooms received total (ECERS-R) quality scores in the minimal range and 8% received scores in the good to excellent range. The proportion of time spent on reading and pre-reading activities ranged from 0% to 36% and time on math ranged from 0% to 44%. Interestingly, these indicators of children's actual experiences, sometimes called process quality, were typically lower than were concurrent assessments of the structural features (e.g., ratios, teacher qualifications) of the classrooms, which easily met the high standards established by the National Association for the Education of Young Children. The question we address is whether the Oklahoma pre-K program, with its relatively stringent structural and teacher requirements, and strong support for focused instructional in-service training, exhibits a significantly different distribution of classroom quality and instructional time than other pre-k programs. We also examine whether the Head Start collaboratives in Tulsa differ in quality and instructional time from the public school programs and from other Head Start programs.

There is also a pressing need to understand which teacher attributes and behaviors predict variation in classroom quality and child outcomes. While 38 states (serving more than 800,000 children) now provide some publicly funded pre-K (Barnett et al., 2005),

many of these programs are in their early stages of implementation, with decisions regarding their design, organization, and staffing still in flux. However, there is a dearth of research focused on pre-K environments that looks inside the pre-K “black box” to identify the specific practices and characteristics of programs, classrooms and teachers that contribute to the pattern of promising results that the Oklahoma program and others have been reporting (see also Gilliam & Zigler, 2001, 2004; Henry et al., 2001, 2003; Magnuson et al., 2004).

This has left the field largely reliant on the center-based child care literature, which may have dubious generalizability to state pre-K programs as a result of child care’s typically more poorly educated and compensated staff, lesser focus on early childhood instruction, and lower likelihood of being co-located with an elementary school (see LaParo, Pianta, & Stuhlman, 2004). What this literature does tell us is that the ingredients of quality child care that best predict positive outcomes for enrolled children are better trained and educated teachers, smaller child-teacher ratios and group sizes, and the sensitivity and responsiveness of the teacher towards the child (Howes, Hamilton, & Matheson, 1994; NICHD ECCRN, 1999, 2000, 2001b, 2002; Phillips, McCartney, & Sussman, 2006; Phillips, Mekos, Scarr, McCartney, & Abbott-Shim, 2001). Efforts to look inside child care and pre-K classrooms to identify effective elements of early childhood instruction are also informative. Key elements appear to be child-centered and flexible instructional practices, time spent on explicit subject-matter learning (as distinct from time spent on transitions and personal care), clear and efficient time management and classroom organization, ample opportunities for child-teacher

conversation and discussion, and a classroom climate characterized by warm, contingent interactions (Bowman, Donovan, & Burns, 2001; Clements, Sarama & DiBaise, 2003; Connor, Morrison, & Slominski, 2006; Howes & Ritchie, 2002; National Council of Teachers of English and International Reading Association, 1996; Snow, Burns, & Griffin, 1998; Stipek & Byler, 2003).

Emerging efforts to examine questions about predictors of quality pre-K programs, as distinct from child care centers, have both informed and complicated this literature. Notably, results from a recent, large-scale study of pre-K programs in six states are indicating that teacher attitudes and beliefs about children are more important factors in predicting classroom quality than are teachers' educational levels, degrees, or credentials (LaParo, Pianta, & Stuhlman, 2004). Moreover, the ratios of teachers to children did not predict variation in pre-K quality. These findings are similar to those emerging from the Head Start FACES data showing that, while higher-quality classrooms are associated with teachers who have higher levels of education (especially a BA or AA degree or some graduate education) and a teacher certificate, these background characteristics appear to affect classroom quality through their positive influence on teacher attitudes and knowledge (Resnick & Zill, 2003). Closer examinations of teacher background variables in the pre-K study revealed that teachers with a BA plus some form of early childhood training provided higher-quality early education as compared to teachers with no formal training in early childhood (Pianta et al., 2005), and that teachers with *more than* a BA degree provided higher-quality educational experiences than teachers with an AA degree (Early et al., 2006). Neither the B.A. degree taken alone nor

state certification in 4-year-old education was associated with higher-quality classrooms. These findings contrast starkly with the child care literature and thus lend support to questions now being raised about the applicability of the child care literature to pre-K settings, as well as to efforts to examine the ingredients of quality that may be specific to pre-K classrooms. The findings we report today will contribute to the small, but important, body of research on this topic.

In sum, we address two questions: (1) What is classroom quality and time on instruction in Tulsa's publicly funded four-year-old classrooms and how does it compare to multi-state assessments of quality in early childhood classrooms? (2) What predicts higher-quality early education in Tulsa's publicly funded four-year-old classrooms and how do these findings compare to the prior literature on the essential ingredients of quality care and early education? We will address the first question with data from pre-K and Head Start. We will address the second question with data from pre-K only, because the Head Start data are still being analyzed.

METHODS

The focus of our inquiry is pre-K classrooms run by the Tulsa Public Schools and four-year-old Head Start classrooms run by the Community Action Project of Tulsa County. Both sets of classrooms are funded by the state of Oklahoma, under a universal pre-K system established in 1998. The Head Start classrooms are eligible for funding because they have established a "collaborative" relationship with Tulsa Public Schools,

and agreed to meet the pre-K quality requirements of the TPS school system. As of the fall of 2006, all four-year-old Head Start programs run by the Community Action Project of Tulsa County were collaboratives.

Classroom Sample

At the outset of our investigation, we identified 100 pre-K classrooms and 29 Head Start classrooms that might be suitable for classroom visits. Because pre-K classrooms might be expected to differ from morning to afternoon, we decided to focus exclusively on morning experiences. For this reason, we excluded afternoon-only classrooms and afternoon sessions of full-day programs from consideration, thus leaving a total of 80 pre-K classrooms and 29 Head Start classrooms eligible for observation.

In early 2006, we trained a team of eight University of Tulsa students so that they would be able to observe the Tulsa Public Schools pre-K and CAP Head Start classrooms. During the spring, these students were trained on two observational instruments, discussed below. Following certification as reliable observers, the students visited a total of 78 TPS pre-K classrooms and 28 Head Start classrooms. Students did not visit two schools at all (Celia Clinton, Patrick Henry) because of problems with scheduling a classroom visit. Of 59 full-day TPS programs, students visited 58, always in the morning. Of 21 morning TPS programs, students visited 20. Of 29 Head Start classrooms, students visited 28, always in the morning. In short, our sample constitutes

virtually the entire universe of state-funded morning classrooms for four-year-olds in Tulsa, Oklahoma.

Observation Instruments

The students worked in pairs and spent an entire morning, from student arrival until dismissal for lunch, observing each classroom. They used two instruments to capture the children's interactions and activities in the classrooms: the Classroom Assessment Scoring System or CLASS (developed by Robert Pianta and his colleagues at the University of Virginia); and the Emerging Academics Snapshot (developed by Carollee Howes and her colleagues at UCLA).¹

The CLASS provides an assessment of classroom quality that encompasses instructional quality, emotional support, and classroom organization. The observer assesses 11 separate dimensions of classroom climate, using a scale of 1-7, with 1-2 indicating low levels, 3-5 medium levels, and 6-7 high levels. For example, the Quality of Feedback – one of the 11 dimensions – measures the extent to which the teacher's feedback focuses on expanding learning and understanding, as opposed to correctness of the end product. Each cycle of observation consists of a 20-minute period of watching

¹ In one school (ECDC), we were unable to administer CLASS, due to an illness. In another school (Phillips), we were unable to administer the Snapshot, also due to an illness. Thus our working sample for CLASS is 77 public school classrooms (and 28 Head Start classrooms), and our working sample for the Snapshot is also 77 public school classrooms (and 28 Head Start classrooms).

classroom interaction and taking notes, followed by a 10-minute period for recording codes. The observer completes as many observation cycles as possible, and codes during all activities except recess. In the majority of cases, our CLASS observers completed five cycles per classroom. In addition to reporting the separate dimensions, we group each classroom's scores into two broad categories that emerged from a factor analysis conducted by scholars using data from 11 states (La Paro et al. 2002). These two factors are Instructional Support and Emotional Support. Although subsequent research suggests the presence of three factors, the use of two factors helps us to compare our findings with earlier work by other scholars. In future work, we will report the results of our own factor analysis.²

The Child Engagement section of the Emerging Academics Snapshot provides detailed information on the time children spend engaged in specific academic activities, such as reading and pre-reading activities, expressive language development, math, science, and social studies. The observer focuses on four children, selected at random, for a period of 20 minutes. She devotes one minute to each child, observing for 20 seconds and coding for 40 seconds, then turns to the next child. Once the 20-minute cycle is completed, the observer shifts to another group of four children, selected at random, and the process begins again. The observer codes throughout all activities, including recess, unless the child being observed leaves the classroom. According to the

² CLASS observers were tested for reliability just prior to data collection. On average, 88% of observer responses were either exactly the same as or within 1 scale-point of the expert's responses. All observers achieved at least 81% reliability.

protocol, as many as 16 children may be observed.³ In the majority of cases, Snapshot observers completed six segments. The Child Engagement section of the Emerging Academics Snapshot consists of 15 items that are coded as present or absent within the 20-second observation period (Ritchie et al. 2002). For example, letter and sound learning is defined as practicing vowel sounds, alliterations, or rhymes; talking about sound-letter relationships; identifying letters; sounding out words; syllabification of words through clapping games; taking an oral spelling test; or reading spelling words to classmates.⁴

Other Data

In addition to the classroom observations, we obtained data on teacher characteristics through a teacher survey. In this survey, we asked pre-K and Head Start teachers about their educational background and training, work experience, and use of particular curricula. We also asked them about their fluency in English and Spanish. Most of the questions for pre-K and Head Start teachers were identical, but we did modify our question wording somewhat, where it seemed appropriate. For example, we asked Head Start teachers when they were first hired by CAP Head Start and whether they ever taught in another Head Start program. As noted earlier, we only utilize teacher survey data for the TPS sub-sample, because the Head Start data are still being analyzed.

³ In some cases, as many as 20 children were observed. In this respect, our observers departed somewhat from the protocol.

⁴ Snapshot observers were tested for reliability just prior to data collection. On average, 92% of observer responses were exactly the same as the expert's responses. All observers achieved at least 90% reliability.

Statistical Comparisons

Because this is the first paper to report on our 2006 classroom findings, we rely primarily on descriptive statistics. We compare TPS pre-K and CAP Head Start classrooms, by conducting a difference of means test, with t-tests to determine statistical significance. We compare TPS pre-K with a national sample of public school based pre-K classrooms and CAP Head Start with a national sample of Head Start classrooms, using the same techniques.⁵ When we use our own data (Figures 1-7), we calculate statistical significance for a difference of means by first determining whether the variances are equal; when we compare our data to national data (Figures 8-21), we calculate statistical significance for a difference of means by assuming that the variances are unequal (lacking the full data set, we cannot make any assumptions about this).

FINDINGS

Intercorrelations among the Quality Measures

Almost across the board, the CLASS and Snapshot scores are positively correlated with each other, such that higher CLASS scores on “positive” dimensions or composites (e.g., Productivity, Emotional Support) are related to larger proportions of

⁵ The national sample comes from the NCEDL pre-K project, which included pre-K programs from 11 states (California, Georgia, Illinois, Kentucky, New York, Ohio, Massachusetts, New Jersey, Texas, Washington, and Wisconsin). The sample is not a random sample of classrooms, but the 11 states account for approximately three-fourths of all children participating in state-funded pre-K programs (Early et al., 2005, p. 3). School-based pre-K programs refer to the physical site of the pre-K program, which overlaps considerably with auspices but is not identical to it.

time spent engaged in Snapshot activities (e.g., Reading, Science). In keeping with this pattern, for example, a high score on the Negative Climate CLASS dimension is related to smaller proportions of time spent engaged in Snapshot activities. The one exception is that neither the Instructional Support CLASS Composite nor any of its three dimensions (Concept Development, Quality of Feedback, or Language Modeling) is significantly correlated with any of the Snapshot scores.

The magnitude of these associations ranges from .167 to .355, suggesting that the two instruments (CLASS and Snapshot), while interrelated in meaningful ways, capture distinct aspects of classroom quality.

Comparisons within Tulsa

In this section, we compare CLASS scores and Emerging Academic Snapshot scores for the Tulsa Public Schools (TPS) pre-K and CAP Head Start four-year-old programs in Tulsa. T-tests indicate whether each difference is statistically significant.

CLASS comparisons (Figures 1-3) suggest that public school-based and Head Start-based classrooms are very similar, in terms of how teachers interact with students. Overall instructional support and emotional support scores are virtually identical, however there are minor differences in classroom management. If we take a closer look at each dimension, we see that TPS pre-K classrooms score higher on behavior

management ($p < .05$), which captures teachers' methods for preventing and redirecting misbehavior.

Clearer differences between the two types of programs emerge if we shift our focus from CLASS scores to Emerging Academics Snapshot scores (Figures 4-7). Within the critical category of language and literacy, TPS teachers and their students devote substantially more time to practicing letters and sounds ($p < .01$). Within the math and science category, TPS teachers and their students devote substantially more time to math engagement ($p < .001$). In concrete terms, TPS classrooms allocate 12.5% of their time to practicing letters and sounds and 17.3% of their time to math engagement. In contrast, Head Start classrooms allocate 7.2% of their time to practicing letters and sounds and 10.8% of their time to math engagement.

Differences in time devoted to other educational activities are also apparent. Specifically, CAP Head Start teachers and their students devote much more time to social studies ($p < .001$) and fantasy play ($p < .001$) than TPS teachers and their students (these two categories overlap quite a bit, as the social studies category encompasses fantasy play).

Interestingly enough, significant differences emerge even in motor activities. Whereas CAP Head Start students are more likely to be engaged in gross motor activities ($p < .05$), TPS students are more likely to be engaged in fine motor activities ($p < .05$).

In short, differences in classroom time allocation practices are fairly striking, while differences in instructional support, emotional support, and classroom management are much more subtle. Whereas TPS pre-K and CAP Head Start students receive similar levels of support from their teachers, they are spending their mornings in rather different ways.

Comparisons, Tulsa v. National Sample

In this section, we compare classroom quality and time allocation scores for Tulsa pre-K programs with those from a national sample of school-based pre-K programs. We also compare classroom quality and time allocation scores for Tulsa Head Start programs with those from a national sample of Head Start programs. T-tests indicate whether a difference is statistically significant.

One difficulty with comparing CLASS scores from Tulsa with those from a national sample of school-based pre-K programs is that one of the emotional support dimensions used in the national sample (over-control) was later discarded in favor of a new dimension (regard for student perspectives), which we used in Tulsa. To avoid an inappropriate comparison, we have simply deleted this dimension from consideration when comparing emotional support scores. For the same reason, we do not compare the Tulsa emotional support composite score with that from the national sample (because their underlying dimensions differ somewhat).

A comparison of instructional support scores for Tulsa pre-K programs and the national sample (Figure 8) indicates that the Tulsa scores are higher for four of four dimensions (all differences are statistically significant).

A comparison of emotional support scores for Tulsa pre-K programs and the national sample (Figure 9) indicates that the Tulsa scores are significantly better for one dimension (a lower negative climate score), significantly worse for another dimension (a lower positive climate score), and no different for the other two dimensions.

Overall, the instructional support composite score for Tulsa's pre-K programs (3.06) exceeds that for the national pre-K sample (2.01) (see Figure 10). We cannot directly compare emotional support composite scores because the underlying dimensions differ somewhat.

When we shift our attention from CLASS scores to Emerging Academics Snapshot observations, we see much clearer differences between TPS pre-K and national school-based pre-K classrooms.

If we focus on language and literacy activities, we see that TPS pre-K teachers devote much more time to key language and literacy activities than their national pre-K counterparts. Most notably, TPS pre-K teachers spend twice as much time (10.5%) as their national counterparts (5.6%) reading to students, and they spend three times as

much time (12.5%) as their national counterparts (4.7%) practicing letters and sounds (see Figure 11).

Differences in math and science are also striking (see Figure 12). TPS pre-K students and their teachers are engaged in math 17.3 percent of the time, as opposed to a national norm of 8.6 percent. TPS pre-K students and their teachers are engaged in science 17.2 percent of the time, as opposed to a national norm of 10.8 percent. Comparisons for computer use were not possible, because these data were not available from the national sample.

Differences in time devoted to other educational activities are more subtle. TPS pre-K students are somewhat more likely to be engaged in aesthetics than their national peers (18.6% vs. 14.4%), but other differences are not statistically significant (see Figure 13). Differences in the amount of time devoted to motor skills are negligible (see Figure 14).

When we turn our attention to Head Start, we also see some differences between the Tulsa experience and national data. For three of four instructional support dimensions, the Tulsa Head Start program scores significantly higher than the national Head Start sample (see Figure 15). Not surprisingly, the instructional support composite scores also differ. The overall score for CAP Head Start of Tulsa County is 3.03, while the overall score for the national sample is 2.06 (see Figure 17).

In contrast, there are no statistically significant differences in emotional support between the Tulsa Head Start program and the national Head Start sample (see Figure 16).

Differences in classroom time allocation are even more striking. The CAP Head Start program of Tulsa County devotes significantly more time to language and literacy than the national Head Start sample in four of five areas (see Figure 18). The Tulsa Head Start program also devotes significantly more time to math, science, social studies, and aesthetics than the national Head Start sample (see Figures 19 and 20). Finally, the Tulsa Head Start program also devotes significantly more time to gross motor activities, but not to fine motor activities, than the national Head Start sample (see Figure 21).

CLASS and Snapshot Score Predictors, TPS

In this section, we report the results of OLS models that predict Tulsa Public Schools pre-K classroom CLASS and Emerging Academics Snapshot scores based on teacher characteristics. The independent variables in each model are: whether the teacher speaks Spanish; a three-category variable for the field in which the teacher received the BA (early childhood education, education, and other, where BA in other field is the omitted category); whether the teacher is in her/his first or second year of teaching; total years as a classroom teacher; whether the teacher uses the Direct Instruction curriculum in the classroom; whether the teacher uses the Waterford Early Learning Program

curriculum in the classroom; whether the teacher uses the TPS Elementary Standard District Framework in the classroom; and whether the teacher participated in the Tulsa Reads training program during the 2001-02 school year. (See Appendix A for descriptive statistics for the variables included in these models.)

As Table 1 indicates, teacher characteristics and curricular choices have some impacts on CLASS scores. Teachers who speak some Spanish score higher on emotional support. Teachers who are in their first or second year of teaching score lower on student engagement. Teachers with more years of teaching experience score higher on classroom management. Teachers who use the Direct Instruction curriculum score higher on classroom management. No other differences are statistically significant.

As Table 2 indicates, teacher characteristics and curricular choices also have some impacts on Snapshot scores. In classrooms taught by newer teachers, more time is devoted to student reading but less time is devoted to expressive language skills. In classrooms taught by teachers with more years of experience, more time is devoted to student reading. In classrooms where the Waterford Curriculum is used, more time is devoted to practicing letters and sounds. In classrooms where the TPS Elementary Standard District Framework is used, less time is devoted to student reading. In classrooms where the teacher participated in the Tulsa Reads training program (2001-02), teachers spend less time reading to students but more time is spent practicing letters and sounds.

Discussion

States have high aspirations for the investments they are making in pre-K education. The expectation is that high-quality, educationally focused programs for four-year-olds will reap benefits in improved school performance for all participating children. The pre-K program in Oklahoma, in particular, has received substantial attention from researchers and decision-makers alike. Not only is this program unusual in its universality, near exclusive reliance on school-based programs, and strict requirements for teacher qualifications and classroom size, but it has generated strong and replicated short-term results for children's achievement (Gormley, Gayer, Phillips, & Dawson, 2005; Gormley and Gayer, 2005).

The Tulsa Public Schools' pre-K classrooms devote more time to pre-literacy and pre-math activities than the Tulsa Head Start program's classrooms. This probably reflects both the more explicit educational focus of a public school system and the more comprehensive goals of Head Start (education, health care, family support). While Head Start children are somewhat less exposed to pre-literacy and pre-math activities than children in school-based classrooms, they are also somewhat more exposed to social studies and fantasy play. At the same time, the quality of instructional and emotional support provided to children in the pre-K and Head Start classrooms is virtually identical. Thus, while the distribution of time spent on various activities differs somewhat, the

quality of children's experiences and interactions in classrooms is roughly comparable across these two settings.

This is encouraging in light of prior and much older evidence that the quality of school-based programs exceeded the quality of Head Start programs (Goodson & Moss, 1992). Perhaps the national effort to upgrade the qualifications of Head Start teachers, and, in Tulsa, the specific requirement that all Head Start teachers participating in the pre-K program have a B.A. with an early childhood certification, has paid off in the quality of children's experiences. It remains to be seen whether this classroom feature, or the differences in actual instructional time allocation, play the larger role in predicting children's academic outcomes at the end of their pre-K year.

When compared to other school-based pre-K programs nationally, the Tulsa pre-K classrooms provide significantly higher instructional support and devote much more time to pre-literacy, pre-math, and science activities. The Head Start classrooms in Tulsa also provide significantly higher instructional support and devote more time to academic skills compared to a nationally representative Head Start sample. While comparisons across research programs must be interpreted with caution, it is noteworthy that our research team was trained by the same experts who were involved with the NCEDL study. This pattern of results parallels the larger learning gains that have also been found for the Tulsa pre-K program as compared to other state pre-K programs (Gilliam & Zigler, 2004; Magnuson et al., 2007), and suggests that higher classroom quality begets greater learning. This remains, however, to be tested directly, which is the next step in our program of research.

The critical question for policy purposes concerns how to produce the kinds of pre-K classrooms that generate learning gains. In our initial effort to address this question, we examined associations between classroom quality and both teacher characteristics and teacher curricular choices. Of the teacher characteristics, experience seemed the most important factor. More experienced teachers possess stronger classroom management skills and are able to devote more time to student reading, without forfeiting time devoted to other academic skills. Freshly minted teachers (less than two years of experience) also devote more time to student reading, but they devote less time to expressive language skills. Also, newer teachers are less successful at engaging their students. At this point, we cannot judge which pre-reading strategies are more effective. However, we can say that teacher experience affects the choices that teachers make.

We also found that teachers who used highly scripted curricula that were developed for low-performing students (Direct Instruction and Waterford Curriculum) offer somewhat different classroom experiences to their students. Teachers who use Direct Instruction were more successful in managing their classrooms. Teachers who use Waterford spent more time practicing letters and sounds, a strategy that follows directly from the foci of this particular curriculum on mastering letters, basic writing skills, and language stories (see www.pearsondigital.com/waterford). In contrast, reliance on the TPS Elementary Standard District Framework discouraged time for student reading. Tulsa Reads, a district-wide training program for pre-K teachers (in 2001-02), appears to

have led to less emphasis on reading to students but more emphasis on practicing letters and sounds.

It is important, in this context, to reiterate that this study was not designed as an assessment of different curricula. While we report some interesting associations, they may be the product of both measured and unmeasured teacher, classroom, and student characteristics that are associated with both the teachers' choice of curricula and classroom quality and practices. Future analyses will examine interactions between teacher characteristics and use of various curricula to ensure that the associations we have reported here are not mediated, rather than direct, relationships. It may be the case, for example, that newer teachers are less likely to use certain curricula. We know that newer teachers are less likely to have participated in the Tulsa Reads training program, which took place before they were hired.

In light of prior evidence that Hispanic children showed the strongest benefits from the Tulsa pre-K program (Gormley et al., 2005), it is of interest that Spanish-speaking teachers scored higher than their non-Spanish speaking colleagues on emotional support in the classroom. This scale captures an atmosphere of enjoyment and respect, the teacher's provision of encouragement, and his/her lack of regimentation and over-control. While none of the Tulsa pre-K classrooms utilizes bilingual instruction, it may be the case that bilingual teachers are more cross-culturally conversant and thus more sensitive to the backgrounds of the Hispanic children in their classrooms. Future analyses

will examine our findings by the racial-ethnic composition of the classrooms and by the race-ethnicity of the individual students (outcome data).

These findings have implications for state pre-K and national Head Start policies, both of which have as explicit goals the provision of high-quality, educationally rich environments for preschool-age children. This appears to have been achieved in Oklahoma with its relatively stringent requirements for teacher qualifications and classroom size. There is, however, substantial variation within the Tulsa pre-K classrooms and some features appear to support higher-quality experiences for young children than others. Our findings suggest that teacher experience matters, though it is not yet clear whether newer or more experienced teachers have made wiser choices. The findings regarding curricula suggest that highly structured learning sequences can support effective classroom management and, in the case of Waterford and Direct Instruction, ensure that students spend relatively more time on pre-literacy activities. We have yet to examine whether student academic and behavioral outcomes vary with the pre-K curriculum and the classroom activities to which students were exposed. The same conclusion applies to the findings regarding teachers who speak some Spanish. While they appear to provide more supportive, respectful, and flexible classroom environments, it remains to be seen if these qualities also facilitate higher scores on tests of learning and behavior.

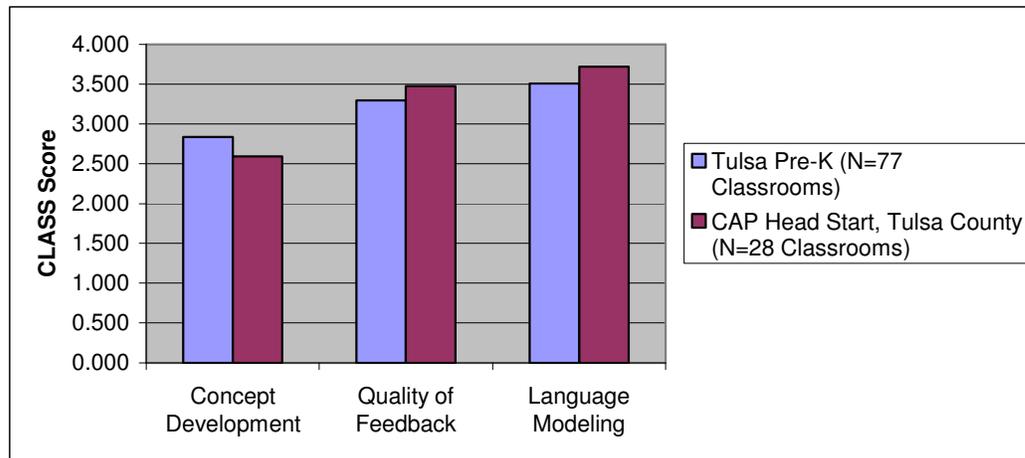
Perhaps most importantly, we have documented important differences between pre-K classroom experiences in Tulsa and comparable settings in other jurisdictions. We

can now offer a partial answer to the question: why is the Tulsa pre-K program producing such substantial learning gains for young children? The quality of Tulsa's instructional support is higher than elsewhere, and Tulsa's pre-K teachers are devoting much more time to pre-reading and pre-math skills than their counterparts in other jurisdictions. These strategic choices appear to have paid off. We have not yet examined the effectiveness of Tulsa's Head Start program, but it also seems to be superior to other Head Start programs, in terms of both instructional support and time devoted to pre-reading and pre-math activities. In the future, we hope to examine the effectiveness of this program and to link test outcomes to teaching practices.

For the time being, we can note that the two Tulsa programs have several things in common: teachers with a B.A. degree, teachers who are early childhood certified, and teachers who are relatively well paid. These factors may help to explain variations between Tulsa and the rest of the nation, while other factors (teacher experience, curricular choices, Spanish language ability) help to explain variations within Tulsa itself.

Figure 1

Classroom Assessment Scoring System (CLASS), Instructional Support Dimensions, TPS Pre-K v CAP Head Start of Tulsa County				
CLASS Score	Tulsa Pre-K (N=77 Classrooms)	CAP Head Start, Tulsa County (N=28 Classrooms)	T-Test, P-value	
Concept Development	2.835	2.591	0.360	
Quality of Feedback	3.293	3.477	0.406	
Language Modeling	3.507	3.719	0.341	

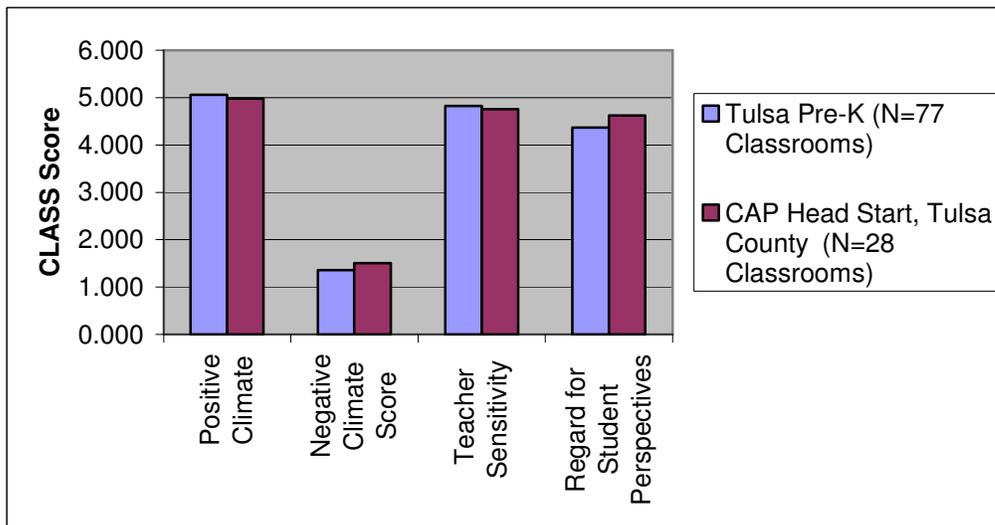


Source: CROCUS, 2007

Figure 2

**Classroom Assessment Scoring System (CLASS),
Emotional Support Dimensions,
TPS Pre-K v CAP Head Start of Tulsa County**

CLASS Score	Tulsa Pre-K (N=77 Classrooms)	CAP Head Start, Tulsa County (N=28 Classrooms)	T-Test, P-value
Positive Climate	5.059	4.983	0.679
Negative Climate Score	1.357	1.501	0.356
Teacher Sensitivity	4.827	4.757	0.717
Regard for Student Perspectives	4.372	4.624	0.185



Source: CROCUS, 2007

Figure 3

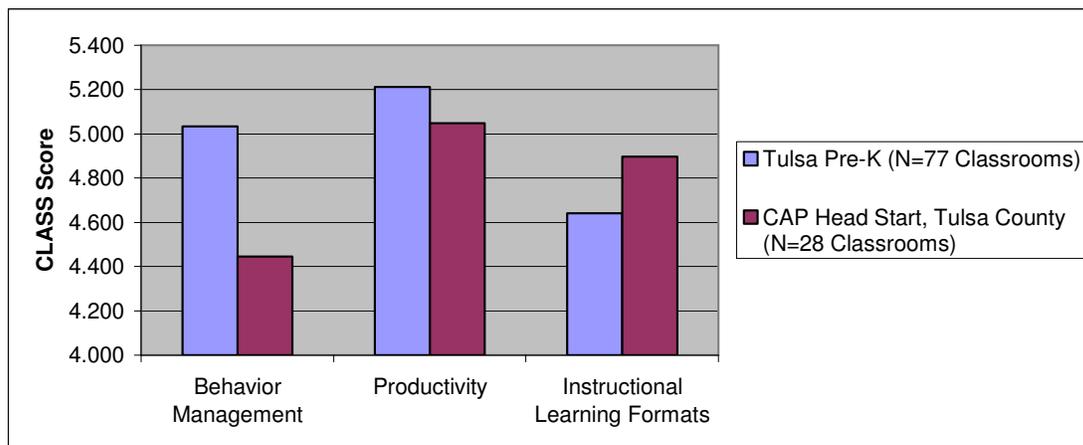
**Classroom Assessment Scoring System (CLASS),
Classroom Management Dimensions,
TPS Pre-K v CAP Head Start of Tulsa County**

CLASS Score	Tulsa Pre-K (N=77 Classrooms)	CAP Head Start, Tulsa County (N=28 Classrooms)	T-Test, P-value
Behavior Management	5.034	4.445	0.030 **
Productivity	5.212	5.049	0.402
Instructional Learning Formats	4.642	4.898	0.169

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Source: CROCUS, 2007

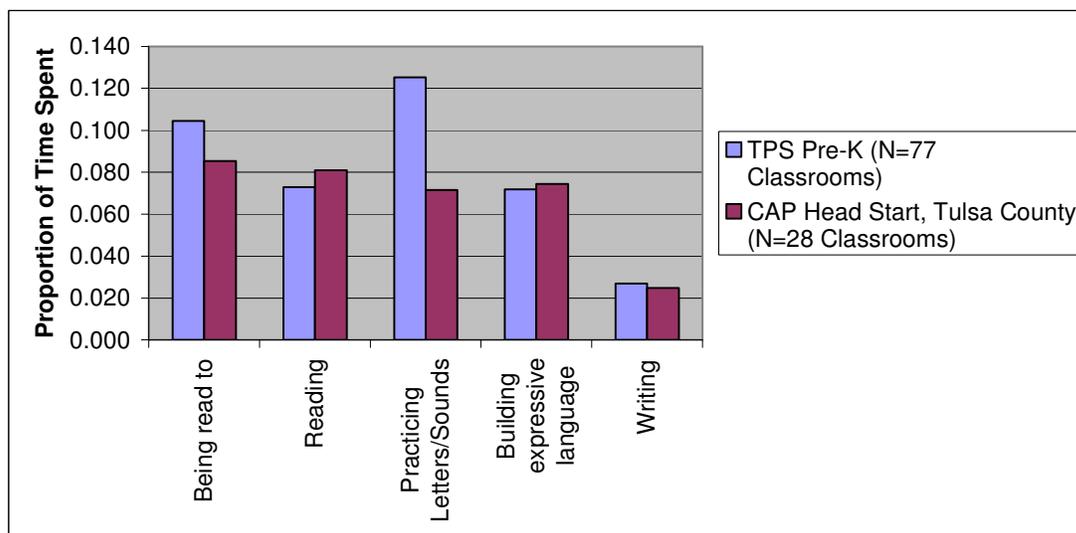
Figure 4

Classroom Time Allocation, Language & Literacy (Snapshot): TPS Pre-K v CAP Head Start of Tulsa County			
Proportion of Time Spent	TPS Pre-K (N=77 Classrooms)	CAP Head Start, Tulsa County (N=28 Classrooms)	T-Test, P-value
Being read to	0.105	0.085	0.194
Reading	0.073	0.081	0.603
Practicing Letters/Sounds	0.125	0.072	0.001 ***
Building expressive language	0.072	0.074	0.856
Writing	0.027	0.025	0.740

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Source: CROCUS, 2007

Figure 5

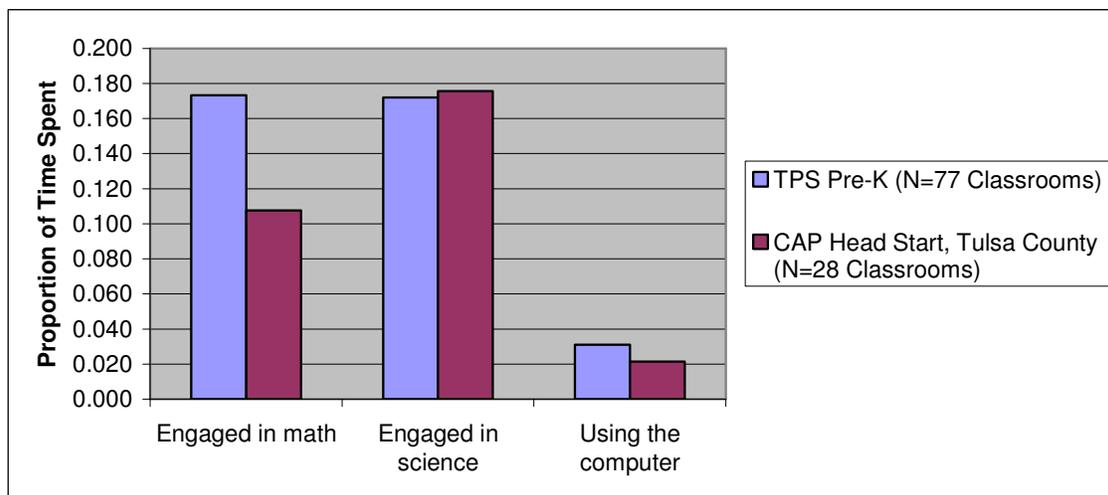
**Classroom Time Allocation, Math & Science (Snapshot):
TPS Pre-K v CAP Head Start of Tulsa County**

Proportion of Time Spent	TPS Pre-K (N=77 Classrooms)	CAP Head Start, Tulsa County (N=28 Classrooms)	T-Test, P-value
Engaged in math	0.173	0.108	0.000 ***
Engaged in science	0.172	0.176	0.874
Using the computer	0.031	0.021	0.255

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Source: CROCUS, 2007

Figure 6

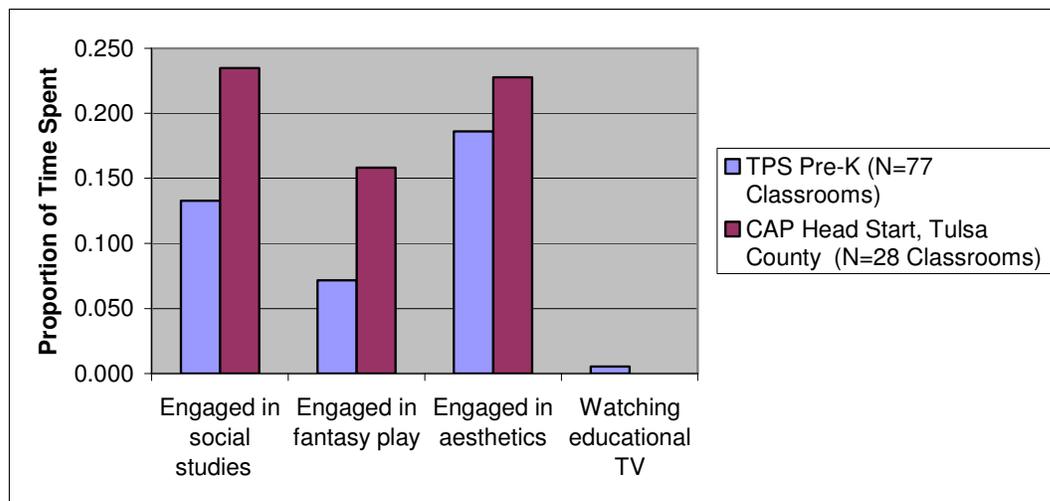
**Classroom Time Allocation, Other Educational Activities (Snapshot):
TPS Pre-K v CAP Head Start of Tulsa County**

Proportion of Time Spent	TPS Pre-K (N=77 Classrooms)	CAP Head Start, Tulsa County (N=28 Classrooms)	T-Test, P-value
Engaged in social studies	0.133	0.235	0.000 ***
Engaged in fantasy play	0.072	0.158	0.000 ***
Engaged in aesthetics	0.186	0.228	0.152
Watching educational TV	0.005	0.000	0.065 *

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Source: CROCUS, 2007

Figure 7

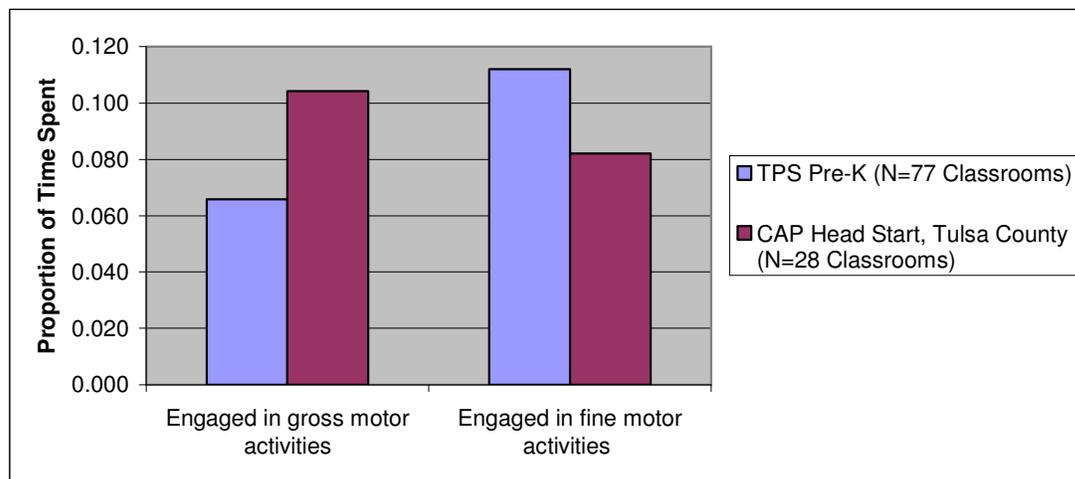
**Classroom Time Allocation, Motor Activities (Snapshot):
TPS Pre-K v CAP Head Start of Tulsa County**

Proportion of Time Spent	TPS Pre-K (N=77 Classrooms)	CAP Head Start, Tulsa County (N=28 Classrooms)	T-Test, P-value
Engaged in gross motor activities	0.066	0.104	0.023 **
Engaged in fine motor activities	0.112	0.082	0.032 **

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Source: CROCUS, 2007

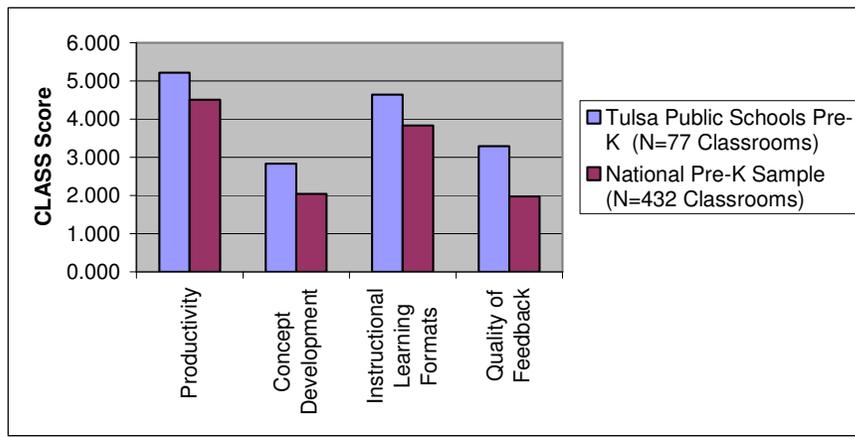
Figure 8

Classroom Assessment Scoring System (CLASS) Scores, Instructional Support Dimensions: Tulsa Public Schools Pre-K v National Pre-K Sample				
CLASS Score	Tulsa Public Schools Pre-K (N=77 Classrooms)	National Pre-K Sample (N=432 Classrooms)	T-Test, P-value	
Productivity	5.212	4.508	0.000	***
Concept Development	2.835	2.041	0.000	***
Instructional Learning Formats	4.642	3.837	0.000	***
Quality of Feedback	3.293	1.976	0.000	***

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP)).

Figure 9

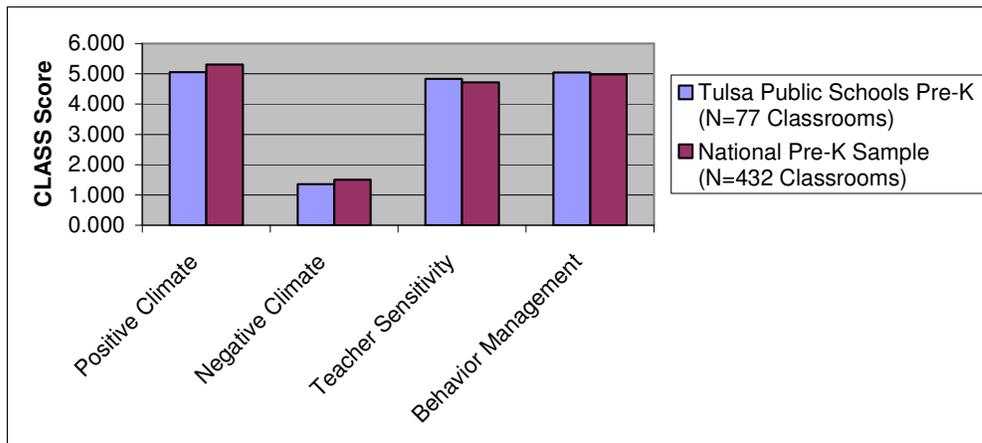
**Classroom Assessment Scoring System (CLASS) Scores, Emotional Support Dimensions:
Tulsa Public Schools Pre-K v National Pre-K Sample**

CLASS Score	Tulsa Public Schools Pre-K (N=77 Classrooms)	National Pre-K Sample (N=432 Classrooms)	T-Test, P-value
Positive Climate	5.059	5.304	0.016 **
Negative Climate	1.357	1.505	0.041 **
Teacher Sensitivity	4.827	4.715	0.209
Behavior Management	5.034	4.982	0.632

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP)).

Figure 10

Classroom Assessment Scoring System (CLASS) Scores, Composite Measures: Tulsa Public Schools Pre-K v National Pre-K Sample			
CLASS Score	Tulsa Public Schools Pre-K (N=77 Classrooms)	National Pre-K Sample (N=432 Classrooms)	T-Test, P-value
Instructional Climate Composite	3.064	2.008	0.000 ***
Emotional Climate Composite	5.225	Not Comparable	NA

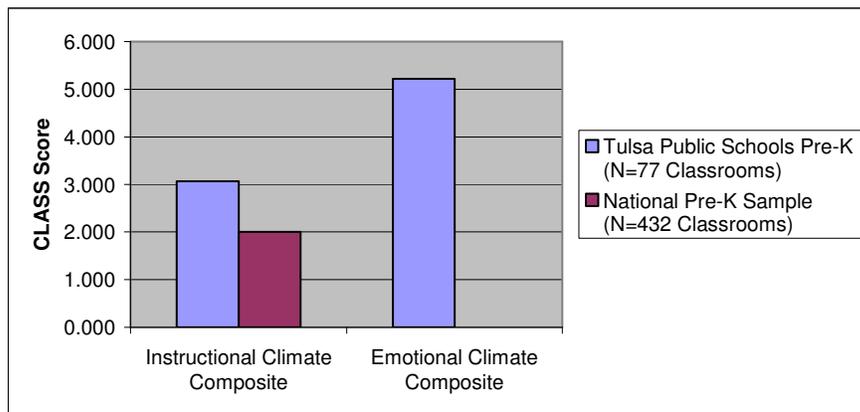
Instructional Climate is a composite of Concept Development and Quality of Feedback.

Emotional Climate is a composite of Positive Climate, Negative Climate (reversed), Teacher Sensitivity, and Regard for Student Perspectives

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP)).

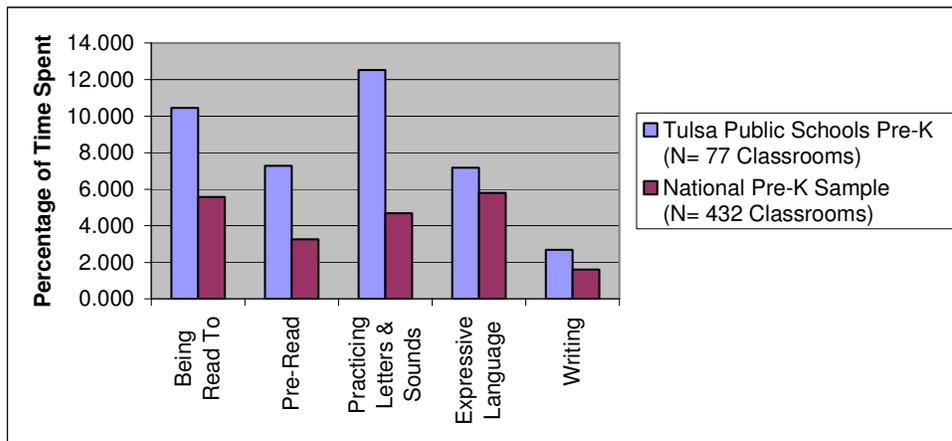
Figure 11

Classroom Time Allocation, Language & Literacy (Snapshot): Tulsa Public Schools Pre-K v National Pre-K Sample				
Percentage of Time Spent	Tulsa Public Schools Pre-K (N= 77 Classrooms)	National Pre-K Sample (N= 432 Classrooms)	T-Test, P-value	
Being Read To	10.450	5.570	0.000	***
Pre-Read	7.290	3.250	0.000	***
Practicing Letters & Sounds	12.520	4.690	0.000	***
Expressive Language	7.180	5.800	0.042	**
Writing	2.690	1.610	0.008	***

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEET)).

Figure 12

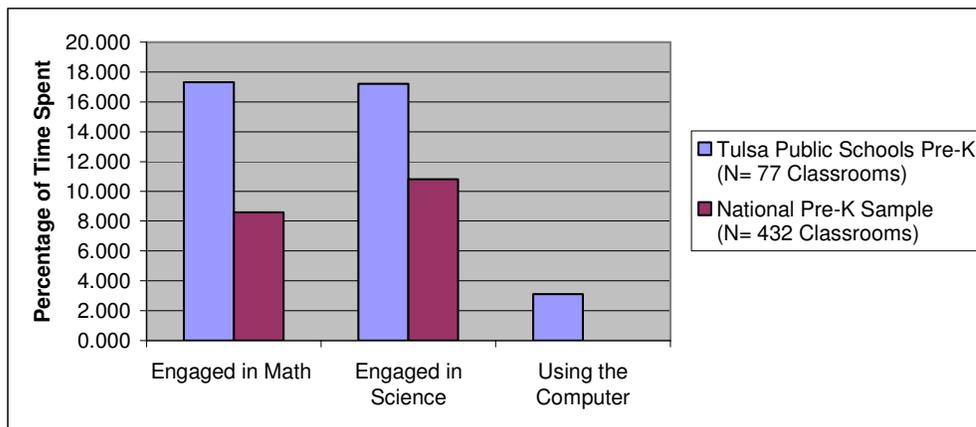
**Classroom Time Allocation, Math & Science (Snapshot):
Tulsa Public Schools Pre-K v National Pre-K Sample**

Percentage of Time Spent	Tulsa Public Schools Pre-K (N= 77 Classrooms)	National Pre-K Sample (N= 432 Classrooms)	T-Test, P-value
Engaged in Math	17.320	8.600	0.000 ***
Engaged in Science	17.210	10.810	0.000 ***
Using the Computer	3.100	.	NA

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEET)).

Figure 13

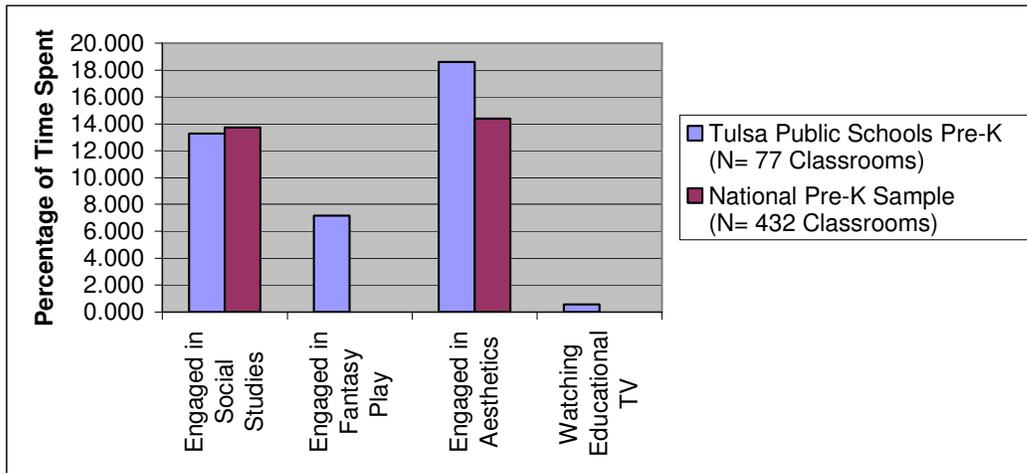
**Classroom Time Allocation, Other Educational Activities (Snapshot):
Tulsa Public Schools Pre-K v National Pre-K Sample**

Percentage of Time Spent	Tulsa Public Schools Pre-K (N= 77 Classrooms)	National Pre-K Sample (N= 432 Classrooms)	T-Test, P-value
Engaged in Social Studies	13.270	13.720	0.653
Engaged in Fantasy Play	7.170	.	NA
Engaged in Aesthetics	18.600	14.390	0.001 ***
Watching Educational TV	0.540	.	NA

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP)).

Figure 14

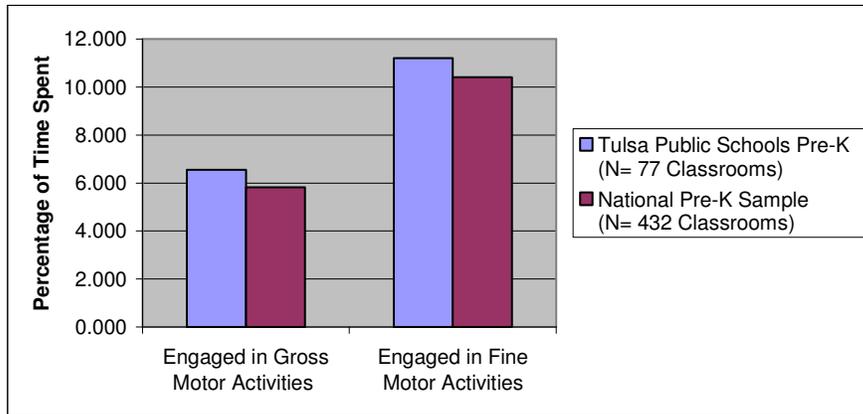
**Classroom Time Allocation, Motor Activities (Snapshot):
Tulsa Public Schools Pre-K v National Pre-K Sample**

Percentage of Time Spent	Tulsa Public Schools Pre-K (N= 77 Classrooms)	National Pre-K Sample (N= 432 Classrooms)	T-Test, P-value
Engaged in Gross Motor Activities	6.550	5.820	0.336
Engaged in Fine Motor Activities	11.210	10.410	0.451

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEET)).

Figure 15

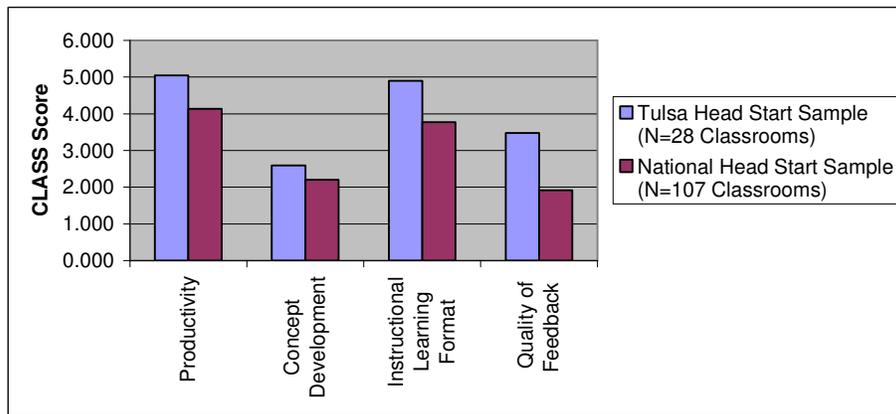
**Classroom Assessment Scoring System (CLASS) Scores, Instructional Support Dimension:
Tulsa Head Start v National Head Start Sample**

CLASS Score	Tulsa Head Start Sample (N=28 Classrooms)	National Head Start Sample (N=107 Classrooms)	T-Test, P-value
Productivity	5.045	4.134	0.000 ***
Concept Development	2.591	2.203	0.128
Instructional Learning Format	4.898	3.770	0.000 ***
Quality of Feedback	3.477	1.916	0.000 ***

* (10% significance level)

** (5% significance level)

*** (1% significance level)



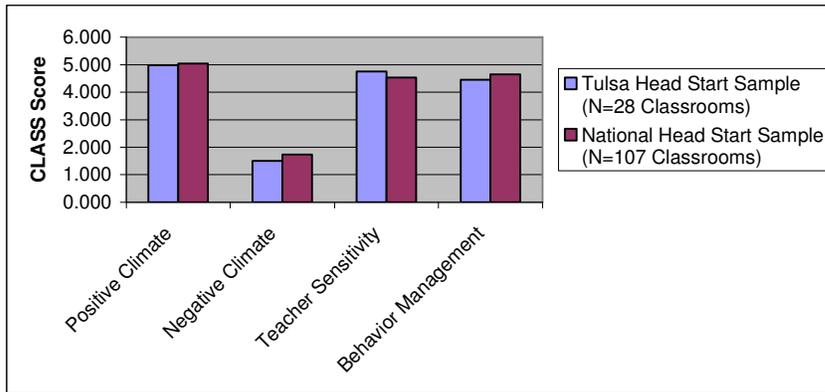
Sources: CROCUS, 2007; Personal communication with Diane Early, March 19, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP)).

Figure 16

**Classroom Assessment Scoring System (CLASS) Scores, Emotional Support Dimension:
Tulsa Head Start v National Head Start Sample**

CLASS Score	Tulsa Head Start Sample (N=28 Classrooms)	National Head Start Sample (N=107 Classrooms)	T-Test, P-value
Positive Climate	4.983	5.051	0.731
Negative Climate	1.501	1.727	0.163
Teacher Sensitivity	4.757	4.535	0.280
Behavior Management	4.445	4.642	0.465

* (10% significance level)
 ** (5% significance level)
 *** (1% significance level)



Sources: CROCUS, 2007; Personal communication with Diane Early, March 19, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP)).

Figure 17

**Classroom Assessment Scoring System (CLASS) Scores, Composite Measures:
Tulsa Head Start v National Head Start Sample**

CLASS Score	Tulsa Head Start Sample (N=28 Classrooms)	National Head Start Sample (N=107 Classrooms)	T-Test, P-value
Instructional Climate Composite	3.034	2.060	0.000 ***
Emotional Climate Composite	5.225	Not Comparable	NA

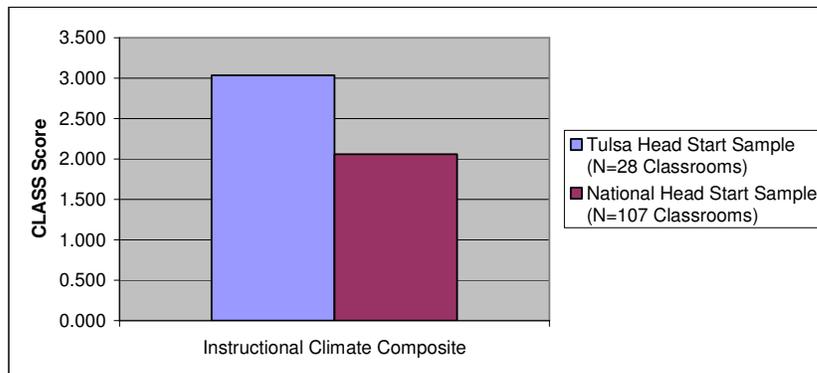
Instructional Climate is a composite of Concept Development and Quality of Feedback.

Emotional Climate is a composite of Positive Climate, Negative Climate (reversed), Teacher Sensitivity, and Regard for Student Perspectives

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS, 2007; Personal communication with Diane Early, March 19, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP)).

Figure 18

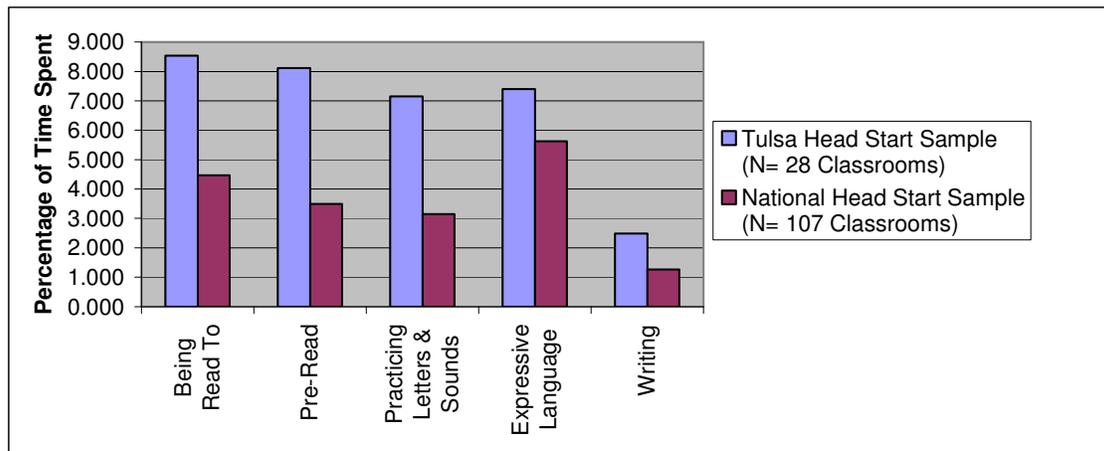
**Classroom Time Allocation, Language & Literacy (Snapshot):
Tulsa Head Start v National Head Start Sample**

Percentage of Time Spent	Tulsa Head Start Sample (N= 28 Classrooms)	National Head Start Sample (N= 107 Classrooms)	T-Test, P-value
Being Read To	8.530	4.463	0.002 ***
Pre-Read	8.110	3.497	0.003 ***
Practicing Letters & Sounds	7.150	3.151	0.002 ***
Expressive Language	7.400	5.623	0.123
Writing	2.480	1.261	0.030 **

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP)).

Figure 19

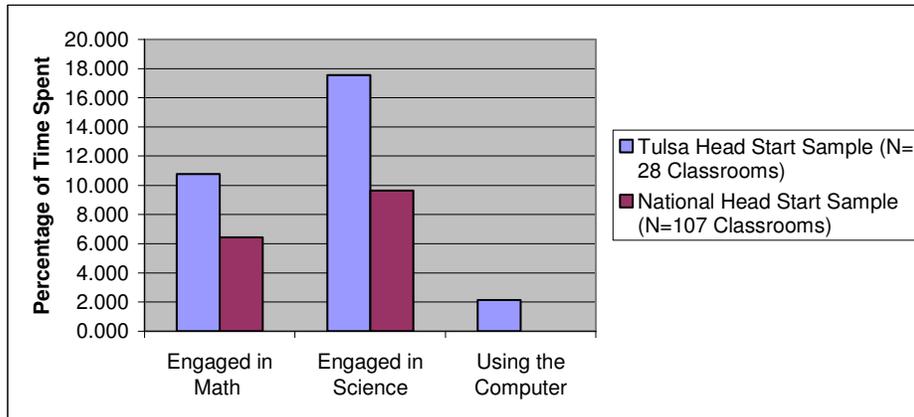
**Classroom Time Allocation, Math & Science (Snapshot):
Tulsa Head Start v National Head Start Sample**

Percentage of Time Spent	Tulsa Head Start Sample (N= 28 Classrooms)	National Head Start Sample (N=107 Classrooms)	T-Test, P-value
Engaged in Math	10.770	6.427	0.000 ***
Engaged in Science	17.570	9.636	0.001 ***
Using the Computer	2.130	.	NA

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEET)).

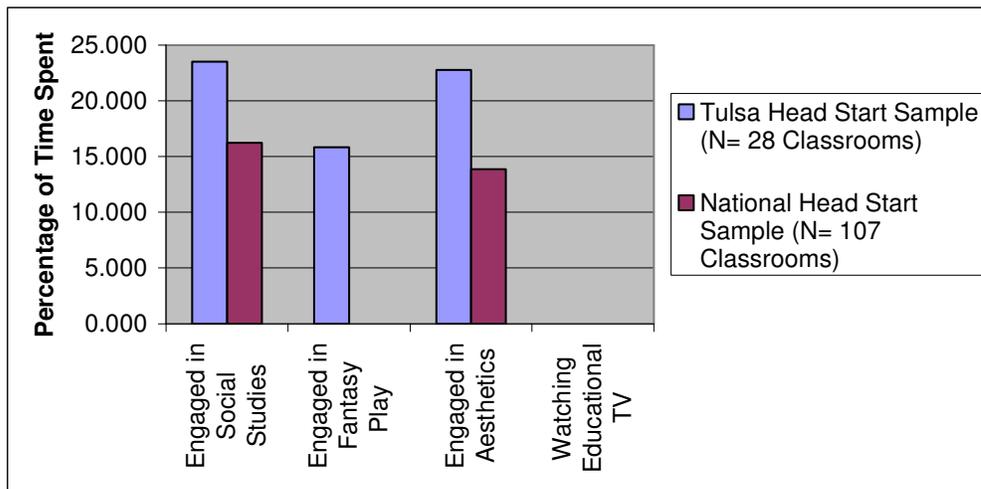
Figure 20

Classroom Time Allocation, Other Educational Activities (Snapshot): Tulsa Head Start v National Head Start Sample			
Percentage of Time Spent	Tulsa Head Start Sample (N= 28 Classrooms)	National Head Start Sample (N= 107 Classrooms)	T-Test, P-value
Engaged in Social Studies	23.500	16.241	0.000 ***
Engaged in Fantasy Play	15.840	.	NA
Engaged in Aesthetics	22.770	13.843	0.003 ***
Watching Educational TV	0.000	.	NA

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEET)).

Figure 21

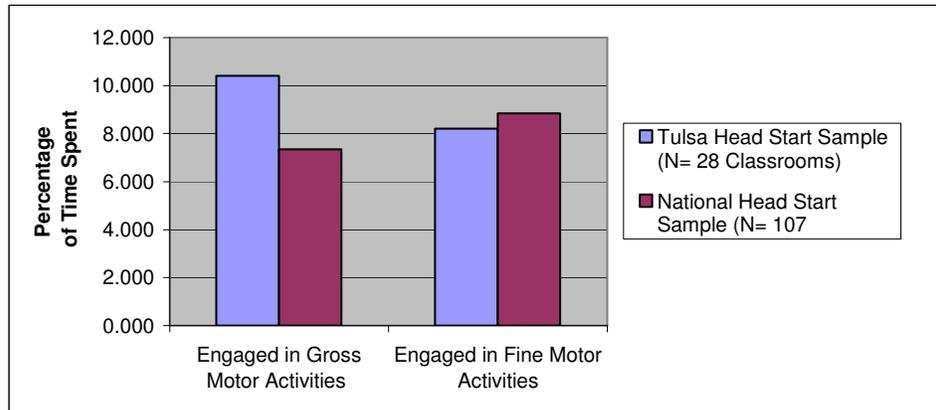
**Classroom Time Allocation, Motor Activities (Snapshot):
Tulsa Head Start v National Head Start Sample**

Percentage of Time Spent	Tulsa Head Start Sample (N= 28 Classrooms)	National Head Start Sample (N= 107 Classrooms)	T-Test, P-value
Engaged in Gross Motor Activities	10.420	7.351	0.054 *
Engaged in Fine Motor Activities	8.210	8.839	0.561

* (10% significance level)

** (5% significance level)

*** (1% significance level)



Sources: CROCUS 2007; Personal communication with Diane Early, March 9, 2007 (NCEDL Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP)).

Table 1: Predictors of CLASS Scores

	(1)	(2)	(3)	(4)
	Emotional Support composite	Classroom Management composite	Instructional Support composite	Average Students' Engagement score
Teacher speaks Spanish	0.292	0.245	-0.147	0.214
	(0.061) *	(0.150)	(0.572)	(0.165)
Teacher has BA in Early Childhood Education	-0.102	-0.195	-0.406	-0.214
	(0.688)	(0.376)	(0.207)	(0.302)
Teacher has BA in Education	-0.004	-0.015	-0.051	-0.209
	(0.988)	(0.954)	(0.894)	(0.414)
Teacher's first or second year of teaching	-0.039	-0.138	-0.360	-0.483
	(0.844)	(0.590)	(0.322)	(0.054) *
Total years as classroom teacher	0.011	0.020	-0.012	0.016
	(0.316)	(0.036) **	(0.471)	(0.174)
Direct Instruction used	0.040	0.259	0.249	0.138
	(0.748)	(0.061) *	(0.343)	(0.437)
Waterford Early Learning Program used	0.170	0.231	-0.002	0.245
	(0.229)	(0.183)	(0.995)	(0.196)
TPS Elementary Standard District Framework used	-0.062	-0.215	0.010	-0.270
	(0.643)	(0.198)	(0.967)	(0.151)
Attended Tulsa Reads training program	-0.145	-0.124	-0.258	-0.193
	(0.452)	(0.511)	(0.425)	(0.412)
Constant	5.135	4.881	3.716	5.419
	(0.000) ***	(0.000) ***	(0.000) ***	(0.000) ***
Observations	73	73	73	73
R-squared	0.106	0.163	0.069	0.204
Robust p values in parentheses				
* significant at 10%; ** significant at 5%; *** significant at 1%				

Note: Teacher has BA in Other Field is the omitted category for BA Field.

Table 2: Predictors of Snapshot Scores

	(1)	(2)	(3)	(4)	(5)	(6)
	Proportion of time spent being read to	Proportion of time spent reading	Proportion of time spent practicing letters/sounds	Proportion of time spent building expressive language	Proportion of time spent writing	Proportion of time spent engaged in math
Teacher speaks Spanish	-0.019 (0.259)	0.014 (0.407)	0.001 (0.986)	0.008 (0.526)	0.001 (0.905)	0.030 (0.314)
Teacher has BA in Early Childhood Education	-0.023 (0.357)	0.001 (0.945)	-0.017 (0.638)	-0.004 (0.775)	0.010 (0.283)	0.004 (0.902)
Teacher has BA in Education	-0.030 (0.111)	-0.035 (0.118)	-0.040 (0.265)	-0.020 (0.208)	0.002 (0.829)	0.008 (0.816)
Teacher's first or second year of teaching	-0.045 (0.106)	0.070 (0.010)***	0.028 (0.410)	-0.028 (0.086)*	0.015 (0.255)	-0.010 (0.836)
Total years as classroom teacher	-0.000 (0.874)	0.003 (0.020)**	0.001 (0.541)	0.001 (0.502)	-0.001 (0.347)	0.002 (0.298)
Direct Instruction used	0.018 (0.400)	0.026 (0.182)	0.044 (0.170)	0.004 (0.760)	-0.001 (0.920)	0.007 (0.810)
Waterford Early Learning Program used	0.029 (0.154)	0.008 (0.633)	0.046 (0.077)*	0.014 (0.368)	-0.012 (0.228)	0.021 (0.504)
TPS Elementary Standard District Framework used	0.023 (0.219)	-0.043 (0.013)**	0.026 (0.343)	0.017 (0.217)	0.001 (0.898)	0.022 (0.434)
Attended Tulsa Reads training program	-0.046 (0.054)*	0.016 (0.327)	0.038 (0.098)*	-0.018 (0.226)	0.017 (0.101)	-0.020 (0.522)
Constant	0.140 (0.000)***	0.035 (0.190)	0.072 (0.070)*	0.071 (0.002)***	0.018 (0.142)	0.134 (0.001)***
Observations	73	73	73	73	73	73
R-squared	0.124	0.281	0.170	0.100	0.080	0.059
Robust p values in parentheses						
* significant at 10%; ** significant at 5%; *** significant at 1%						

Note: Teacher has BA in Other Field is the omitted category for BA Field.

Appendix A: Descriptive Statistics, OLS Variables, TPS Pre-K Programs

Teacher-Questionnaire Independent Variable	Obs	Mean	Std. Dev.	Min	Max
Teacher speaks Spanish	74	0.324324	0.471317	0	1
Teacher has BA in Early Childhood Education	74	0.527027	0.502677	0	1
Teacher has BA in Education	74	0.27027	0.447131	0	1
Teacher has BA in other field (included in intercept)	74	0.202703	0.404757	0	1
Teacher's first or second year of teaching	74	0.216216	0.414473	0	1
Total years as classroom teacher	74	8.824324	7.768232	0	30
Direct Instruction used	74	0.297297	0.460189	0	1
Waterford Early Learning Program used	74	0.324324	0.471317	0	1
TPS Elementary Standard District Framework used	74	0.472973	0.502677	0	1
Attended Tulsa Reads training program	74	0.527027	0.502677	0	1

CLASS Dependent Variable	Obs	Mean	Std. Dev.	Min	Max
Emotional Support composite	73	5.2265	0.567206	3.25	6.8
Classroom Management composite	73	4.983355	0.661968	3.166667	6.5
Instructional Support composite	73	3.201837	0.943578	1.4	5.933333
Average Students' Engagement score	73	5.247097	0.690422	3.25	6.6

Snapshot Dependent Variable	Obs	Mean	Std. Dev.	Min	Max
Proportion of time spent being read to	73	0.103239	0.076314	0	0.38
Proportion of time spent reading	73	0.075025	0.07238	0	0.306569
Proportion of time spent practicing letters/sounds	73	0.127313	0.098322	0	0.474138
Proportion of time spent building expressive language	73	0.070377	0.051403	0	0.188889
Proportion of time spent writing	73	0.028187	0.033429	0	0.151261
Proportion of time spent engaged in math	73	0.171011	0.100496	0.009009	0.466019

REFERENCES

- Barnett, W.S., Hustedt, J., Robin, K., & Schulman, K. (2005). *The State of Preschool: 2005 State Preschool Yearbook*. New Brunswick, N.J.: National Institute for Early Education Research, 2005.
- Bowman, B., Donovan, M.S., & Burns, M.S., eds. (2000). *Eager to learn: Educating our preschoolers*. Washington, DC: National Academy Press.
- Campbell, F.A., Ramey, C.T., Pungello, E., Sparlin, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project: *Applied Developmental Science*, 6, 42-57.
- Clements, D., Sarama, J., & DiBiase, A. (2003). *Engaging young children in mathematics: Standards for early childhood mathematics education*. Mahwah, N.J.: Erlbaum.
- Connor, Carol McDonald, Morison, Frederick J., & Slominski, Lisa (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, 98(4), 665-689.
- Clifford, R.M., Barbarin, O, Change, F., Early, D, Bryant, D., Howes, C., Burchinal, M., & Pianta, R. (2005). What is pre-kindergarten? Characteristics of public pre-kindergarten programs. *Applied Developmental Science*, 9(3), 126-143.
- Early, D.M., Bryant, D.M., Pianta, R.C., Clifford, R.M., Burchinal, M.R., Ritchie, S., Howes, C, & Barbarin, O. (2006). Are teachers' education, major, and credential related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly*, 21, 174-195.
- Early, D.M., Barbarin, O., Bryant, D., Burchinal, M., Chang, F., Clifford, R., Crawford, G., Weaver, W., Howes, C., Ritchie, S., Kraft-Sayre, M., Pianta, R., & Barnett, W.S. (2005). *Pre-Kindergarten in Eleven States: NCEDL's Multi-State Study of Pre-Kindergarten & Study of State-Wide Early Education Programs (SWEEP)*. Chapel Hill, N.C.: NCEDL Working Paper, May 4.
- Gilliam, W. & Marchesseault, C. 2005. *From Capitols to Classrooms, Policies to Practice: State-Funded Prekindergarten at the Classroom Level*. New Haven, Conn.: Child Study Center, Yale University, March 30.
- Gilliam, W., & Zigler, E. (2001). A critical meta-analysis of all evaluations of state-funded preschool from 1977 to 1998: Implications for policy, service delivery and program evaluation, *Early Childhood Research Quarterly*, 15, 441-473.

- Gilliam, W., & Zigler, E. (2004). *State efforts to evaluate the effects of prekindergarten: 1977 to 2003*. Unpublished manuscript: Yale University Child Study Center.
- Goodson, B. & Moss, M. (1992). Analysis report: Predicting quality of early childhood classrooms. *Observational Study of Early Childhood Programs*. Boston, MA: Abt Associates, Inc.
- Gormley, W., Gayer, T., Phillips, D. & Dawson, B. (2005). The effects of universal pre-K on cognitive development, *Developmental Psychology*, 41(6), 872-884.
- Gormley, W. & Gayer, T. (2005). Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-K program, *Journal of Human Resources* 40 (3), 533-558.
- Gormley, W. & Phillips, D. (2005). The effects of universal pre-K in Oklahoma: Research highlights and policy implications, *Policy Studies Journal* 33 (February): 65-82.
- Henry, G., Gordon, C., Mashburn, A. & Ponder, B. (2001). *Pre-K Longitudinal Study: Findings from the 1999-2000 School Year*. Atlanta, Georgia: Georgia State University, Applied Research Center.
- Henry, G.T., Henderson, L.W., Ponder, B.D., Gordon, C.S., Mashburn, A.J., & Rickman, D.K. (2003). *Report of the Findings from the Early Childhood Study: 2001-02*. Atlanta, GA: Georgia State University, Andrew Young School of Policy Studies.
- Howes, C., Matheson, C., & Hamilton, C. (1994). Maternal, teacher, and child care correlates of children's relationships with peers. *Child Development*, 65, 264-273.
- Howes, C., & Ritchie, S. (2002). *A Matter of Trust: Connecting Teachers and Learners in the Early Childhood Classroom*. New York, NY: Teachers College Press.
- La Paro, Karen, Robert Pianta, and Megan Stuhlman. 2004. "The Classroom Assessment Scoring System: Findings from the Prekindergarten Year," *The Elementary School Journal* 104 (May), pp. 409-26.
- Maguson, K., Ruhm, C., and Waldfogel, J. (2007). Does Prekindergarten Improve School Preparation and Performance? *Economics of Education Review* 26, 33-51.
- Magnuson, K., Meyers, M., Ruhm, C., and Waldfogel, J. (2004). Inequality in Preschool Education and School Readiness, *American Educational Research Journal* 41 (Spring), 115-157.
- National Council of Teachers of English and International Reading Association (1996). *Standards for the English language arts*. Urbana, IL: Author.

- NICHD Early Child Care Research Network. (1999). Child outcomes when child care center classes meet recommended standards for quality. *American Journal of Public Health, 89*, 1072-1077.
- NICHD Early Child Care Research Network. (2000). The relation of child care to cognitive and language development. *Child Development, 71*(4), 960-980.
- NICHD Early Child Care Research Network. (2001). Child care and the children's peer interactions at 24 and 36 months. *Child Development, 72*, 1478-1500.
- NICHD Early Child Care Research Network. (2002). Structure>Process>Outcome: Direct and indirect effects of caregiving quality on young children's development. *Psychological Science, 13*, 199-206.
- NICHD Early Child Care Research Network & Duncan, G. Y. (2003). Modeling the impacts of child care quality on children's preschool cognitive development. *Child Development, 74*, 1454-1475.
- Phillips, D., McCartney, K., and Sussman, A. (2006). Child care and early development. In McCartney, K., & Phillips, D. *The Handbook of Early Child Development*. (pp. 471-489). Blackwell Publishers.
- Phillips, D., Mekos, D., Scarr, S., McCartney, K., & Abbott-Shim, M. (2001). Within and beyond the classroom door: Assessing quality in child care centers. *Early Childhood Research Quarterly, 15*, 475-496.
- Phillips, D.A., Voran, M., Kisker, E., Howes, C., & Whitebook, M. (1994). Child care for children in poverty: Opportunity or inequity? *Child Development, 65*, 472-492.
- Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science, 9*(3), 144-159.
- Ramey, C.T., Bryant, D.M. & Suarez, T.M. (1985). Preschool compensatory education and the modifiability of intelligence: A critical review. In D.K. Detterman (Ed.), *Current topics in human intelligence. Vol. 1. Research methodology*. Norwood, N.J.: Ablex Publishing Corp., 247-296.
- Ramey, C.T., Campbell, F.A., Burchinal, M., Skinner, M.L., Gardner, D.M. & Ramey, S.L. (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science 4* (1): 2-14.

- Resnick, G., & Zill, N. (April 2003). *Understanding Quality in Head Start Classrooms: The Role of Teacher and Program-Level Factors*. Paper presented at the Biennial Meetings of the Society for Research in Child Development, Minneapolis, MN, April 2003.
- Ritchie, Sarah, Carolee Howes, M. Kraft-Sayre, and Billie Weiser. 2002. *Emerging Academics Snapshot*. Los Angeles: University of California at Los Angeles.
- Schweinhart, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R., & Nores, M. (2005). *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40*. Monographs of the High/Scope Educational Research Foundation, 14). Ypsilanti, MI: High/School Press.
- Snow, C., Burns, M.S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Stipek, D., & Byler, P. (2003). *The Early Childhood Classroom Observation Measure*. Unpublished Manuscript: Stanford, CA: Stanford University.
- U.S. Government Accountability Office (2004). *Prekindergarten: Four selected states expanded access by relying on schools and existing providers of early education and care to provide services*. Washington, D.C.: September 9.
- Xiang, Z., & Schweinhart, L. (2002). *Effects Five Years Later: The Michigan School Readiness Program Evaluation Through Age 10*. Paper prepared for the Michigan State Board of Education. Ypsilanti, Michigan: High/Scope Educational Research Foundation, January 23.