

New Waves—Educational Research & Development

New Waves—Educational Research and Development is a refereed journal published by the *Chinese American Educational Research and Development Association*, a non-profit professional organization. *New Waves* can be accessed from the *CAERDA* website, the Directory of Open Access Journals and the EBSCO databases. **ISSN: 1526-8659**

Volume 14, Number 1, 2011

Editor-in-Chief: Wen Ma, Le Moyne College

Associate Editor: Binbin Jiang, Kennesaw State University

Editorial Review Board:

Tak C. Chan, Kennesaw State University
Yi Ding, Fordham University
Ran Hu, East Carolina University
Miranda Lin, Illinois State University
Susan Sun, West Virginia University
Ko-Yin Sung, Utah State University
Yea-Ling Tsao, Minnesota State University, Mankato
Lihshing L. Wang, University of Cincinnati
Rong Yuan, Defense Language Institute
Jing Zhang, Defense Language Institute
Aiping Zhao, Shandong University

Table of Contents

Notes from the Editor	1
A Comparison of Progress Monitoring Scores and End-of-Grade Achievement	3
Bob Algozzine, Chuang Wang, Anatoli Boukhtiarov University of North Carolina at Charlotte	
Modeling the Effects of Home and Student Factors on Text Comprehension	22
Qiuying Wang, Oklahoma State University Ze Wang, University of Missouri Steven J. Osterlind, University of Missouri	
A Comparative Study on Pedagogical Content Knowledge of Mathematics Teachers in China and the U.S.	35
Xiaobo She, Texas Tech University William Y. Lan, Texas Tech University Jennifer Wilhlem, University of Kentucky	
Faculty Roles in University Governance at Two Universities in the U.S. and China	50
Hongchi Jiang, China University of Geosciences Michael Wei, University of Missouri	
Chinese as a Foreign Language Learners' Strategy Use and Writing Achievement	67
Ko-Yin Sung, Utah State University	

Notes from the Editor

Welcome to another packed issue of *New Waves-Educational Research and Development (NWERD)*! Before I begin to tell you about the interesting articles in this issue, I want to take a moment to thank the Board members, Webmaster E. Shen, President Ya-yu Lo, and past President Chuang Wang, all of the *Chinese American Educational Research and Development Association (CAERDA)*, for their support of my work. I also like to thank the Editorial Review Board members (see the list of reviewers for this issue) for volunteering their valuable time to provide insightful feedback in the often multiple rounds of review processes. I want to thank the numerous authors, too, for choosing *NWERD* as the venue to disseminate their scholarship within the *CAERDA* community and beyond. Of course I should thank the countless readers who have given enthusiastic suggestions and comments on *NWERD* publications, as well as kind encouragement to me. Colleagues and friends, my sincere thanks to you all!

This issue contains five rich research articles, which reflect very diverse subject areas, conceptual and theoretical underpinnings, research methods, and educational implications. Needless to say, this is in keeping with *CAERDA*'s core values as an all-inclusive, multicultural professional organization. To begin with the issue, Algozzine, Wang, and Boukhtiarov focus on what they call how "performance on widely-used progress monitoring measures relates to end-of-grade achievement on statewide assessments." Through a strenuous evaluation of the STAR Reading and SRI-I performance by more than 1,000 students from diverse backgrounds, Algozzine et al. conclude that the students' scores were statistically significantly correlated to the Florida

Comprehensive Assessment Test results. These findings have broad implications, especially in terms of increasing K-12 students' learning outcomes. The rigorous methodology of the study itself, I think, is exemplary of fine statistical analysis practices.

Next, Wang, Wang, and Osterlind's study includes even more participants (over 8,000 fourth graders) from the United Kingdom and Hong Kong. Drawing on the large-scale secondary data from the Progress in International Reading Literacy Study 2001 (PIRLS 2001), Wang et al. explore how home and parental factors may influence students' text comprehension as measured by their attitude toward reading and reading self-concept, and whether there exists any overarching difference between the Eastern (Chinese) and the Western (UK) cultures. As the article suggests, such research yields useful insight on how to improve reading instructions in a Chinese or a Western sociocultural context.

She, Lan, and Wilhlem's article continues with an international focus, this time involving eight mathematics teachers from Sichuan in China and Texas in the United States. Specifically, She et al. employ an interview protocol consisted of multiple algebraic questions to help them probe any similarities and differences in these educators' pedagogical content knowledge. Their findings are quite revealing: the four U.S. teachers were more prone to use concrete models and practical approaches in promoting their students' problem-solving and mathematical thinking skills, whereas their Chinese counterparts were more inclined to utilize theories and procedures in instruction. While there may be both merits and demerits in both approaches, the four U.S. mathematics teachers appeared to lack more in-depth

understanding of mathematical concepts and cross-conceptual interconnections, the four Chinese teachers seemed to be more at ease in integrating concepts into a more meaningful conceptual network. When supported by larger sample size and more participant diversity, such findings can be very useful for effective mathematics instruction and assessment in both countries.

Jiang and Wei investigate another important area of study in the higher education field: university governance. Drawing on structured interviews of 20 faculty members from a Midwestern university in the United States and a university in Central China, Jiang and Wei discover drastically different governance structures, which in turn have impacted the curriculum, faculty roles, student lives, and institutional organization. The authors' suggestions on effective governance, while meaningful for both settings, seem to be particularly relevant to the ongoing university reforms in China.

The last article by Sung showcases four Chinese as a Foreign Language (CFL) college students' strategy use and their writing achievement. Using a variety of data, including survey, learner interviews, classroom observations, and learners' writing samples, Sung concludes that CFL learners regularly exposed to explicit writing strategy instruction tend to write at a higher proficiency level than those only required to focus on oral language skills. In light of the growing interest in teaching and learning CFL in the United States and around the world, the development of CFL learners' balanced listening, speaking, reading, and writing skills through learning strategies will be quite meaningful.

Finally, this is the last issue under my editorship as I am now completing my two-year term. Looking back, I am pleased that *NWERD* has continued to grow during my tenure, that its authors, reviewers and

readers are all increasing, and that *NWERD* has now been included in the *Directory of Open Access Journals (DOAJ)* and *EBSCO* databases, adding to the visibility of *NWERD*, as well as our *CAERDA* community. Moving forward, Binbin Jiang from Kennesaw State University will take over as the new editor (she has already teamed up with me and worked diligently as the Associate Editor). Indeed, this issue is the fruition of our collaboration. I am confident that *NWERD* will keep on gaining prominence as a refereed professional journal, and I can think of no one more suited for this position than Binbin Jiang!

Wen Ma
Editor

A Comparison of Progress Monitoring Scores and End-of-Grade Achievement

Bob Algozzine

Chuang Wang

Anatoli Boukhtiarov

University of North Carolina at Charlotte

Abstract

In this study, we addressed the need for research demonstrating the extent to which performance on widely-used progress monitoring measures related to end-of-grade achievement on statewide assessments. Specifically, we evaluated the usefulness of STAR Reading and Scholastic Reading Inventory-Interactive for predicting performance on the Florida Comprehensive Assessment Test. We found that scores obtained from regular use of these measures were statistically significantly related to overall end-of-grade achievement markers. We discuss our findings in the context of other similar research.

The No Child Left Behind (NCLB) Act (U.S. Department of Education, 2004) has raised academic standards for all children. Most professionals agree that there has been a correlated increase in the amount of time students spend in assessment-related activities, especially those linked to high stakes testing and similar education reforms (Ananda & Rabinowitz, 2001; Brown & Coughlin, 2007; Simpson, LaCava, & Graner, 2004; Sibley, Biwer, & Hesch, 2001). NCLB also directs that schools use frequent classroom-based assessments to keep track of the progress students are making in reading and other academic areas (Perie, Marion, & Gong, 2007; Schilling, Carlisle, Scott, & Zeng, 2007). Gathering yearly student performance data on local, state, and national indicators has been a part of America's educational accountability efforts for some time and most states require

participation of all students in reading and math assessments during elementary, middle, and high school years (Simpson et al., 2004; Thurlow & Thompson, 1999; Thurlow & Wiley, 2004). Frequent academic progress monitoring has achieved a new level of prominence as a critical feature of Response-to-Intervention (RTI) practices promising to reform the numbers and types of children receiving special education and outcomes for all children (Fuchs & Deshler, 2007). The assumption is that frequently reviewing performance will help teachers identify students who are at-risk and not making adequate progress so that they have a basis for devising suitable plans for instruction. This belief and emerging system is grounded in the existence and use of valid and reliable predictors of students' progress toward the goal of achieving grade-level reading skills (Buck, Torgesen, & Schatschneider, n.d.; Brown & Coughlin, 2007; Fuchs & Deshler, 2007; Hintze, Callahan, Matthews, Williams, & Tobin, 2002; Hintze, Ryan, & Stoner, 2003; Hixson & McGlinchey, 2004; Perie et al., 2007; Roehrig, Petscher, Nettles, Hudson, & Torgesen, 2008).

Literacy skills are fundamental to successful academic performance and frequent assessment and monitoring of them is the foundation for response-to-intervention practices that inform teachers about their students' instructional needs (Fuchs & Fuchs, 2006). In this regard, Coyne and Harn (2006) point out that knowledge of literacy performance directs improved outcomes by providing important

answers that support data-based decision making to improve instruction. In recent years, researchers and other education professionals have expressed concern about the importance this placed on high stakes achievement testing (Wixson & Carlisle, 2005) and there is continuing concern that infrequent, general, end-of-the-year assessments are not useful for making instructional decisions at the classroom level (Roehrig et al., 2008).

Research shows, and virtually all educators agree, that academic improvement requires practice to reinforce skills being learned and continuous monitoring of progress to ensure appropriate areas are targeted for instruction (Fuchs, 1989; Scott & Weishaar, 2003). Unfortunately, the role of practice and progress monitoring is often overlooked and misunderstood. Just setting aside time for student practice is not enough. Similarly, checking performance several times a year provides insufficient evidence for improving skills requiring more frequent attention. Practice must be personalized to each student's individual ability level and immediately followed by informed feedback to ensure a high rate of engagement and success. It must also provide progress monitoring evidence for teachers and other professionals to use to improve instruction and outcomes.

Progress-monitoring assessments must meet several requirements (Brown & Coughlin, 2007; Deno, 1992, 1997, 2003; Deno & Mirkin, 1977; Fuchs & Fuchs, 1999; Perie et al., 2007; Roehrig et al., 2008). First, the content used for keeping track of progress must be representative of the academic performance expected of students at the end of the school year. The measures must also be free of floor or ceiling effects and sensitive to change over a short period of time, over repeated measurements as students gain more skills. The assessment also must be authentic and

have adequate technical characteristics (i.e., validity and reliability). Finally, the outcomes must accurately predict improvements on more generalized assessment measures, such as standardized tests. Progress-monitoring tools that meet or exceed these requirements provide valid and reliable data.

Teachers use progress-monitoring to make decisions about an individual student's instructional needs. For example, based on a student's performance, a teacher may increase the amount and type of instruction, slow the pace of it, or change methods of teaching completely. The use of progress-monitoring instruments has been demonstrated to improve student outcomes in academic areas and has been a widely-accepted evidence-based practice in special education for many years (e.g., Fuchs, 2004; Fuchs & Fuchs, 1986; Madelaine & Wheldall, 2004; Safer & Fleischman, 2005). Relationships between progress monitoring measures and end-of-grade outcomes have also been reported for elementary school students across several demographic groups and statewide assessments (e.g., Barger, 2003; Buck & Torgesen, n.d.; Buck, Torgesen, & Schatschneider, n.d.; Roehrig et al., 2008; Vander Meer, Lentz, & Stollar, 2005; Wanzek, Roberts, Linan-Thompson, Vaughn, Woodruff, & Murray, 2010; Wilson, 2005). This extant knowledge base is grounded in studies illustrating the predictive value of early literacy skills (e.g., oral reading fluency) for success in third grade. In this research, we investigated similarities and differences in performance as well as relationships between scores and predictive accuracy of two widely-used progress monitoring assessments and a state-wide high stakes reading test for students in grades six, seven, and eight. Specifically, we addressed on three research questions with implications for improving summative and

formative assessment practices for at-risk students in middle school:

1. To what extent is performance for sixth, seventh, and eighth grade students on different measures of progress monitoring and end-of-grade reading achievement similar for different demographic groups?
2. To what extent is performance for sixth, seventh, and eighth grade students on different measures of progress monitoring and end-of-grade reading achievement related?
3. To what extent does performance for sixth, seventh, and eighth grade students on different measures of progress monitoring similarly predict performance on end-of-grade reading achievement?

Method

Renaissance Learning offers a computer-adaptive test of general reading ability (STAR Reading) that has good reliability and validity as evidenced primarily by its technical characteristics and correlation with other tests (Renaissance Learning, Inc., 2000, 2006a, b). Scholastic offers a reading comprehension test (Scholastic Reading Inventory-Interactive: SRI-I) that assesses students' reading levels, tracks students' reading growth over time, and helps guide instruction according to students' needs (Scholastic, 2001a, b, 2006). The focus of this project was an analysis of the relationships between scores on STAR Reading, SRI-I, and the Florida Comprehensive Assessment Test (FCAT: Florida Department of Education, 2002; n.d.). Our work addressed the need for research examining the use of interim assessments for improving classroom practice answered and restarted by Roehrig

et al. (2008) and others (Council of Chief State School Officers, n.d.; Perie et al., 2007).

Participants

The sample included a total of 1,077 students with complete assessment information. Of the participants, 514 (48%) were female and 563 (52%) were male. Slightly more than half of the students (53%) were African-American; Caucasian students were the second largest group (23%), and Hispanic students were the third largest group (19%); and, there were 29 (3%) Asian and 18 (2%) multi-racial students in the sample. A total of 580 (54%) were eligible for free or reduced price lunch program. Statistically similar distributions were evident across grades for gender ($X^2 = 0.45$, $df = 2$, $p > .05$), ethnicity ($X^2 = 12.00$, $df = 8$, $p > .05$), and free lunch status ($X^2 = 2.98$, $df = 2$, $p > .05$). Other information about the participants is summarized in Table 1.

Procedures

In early February, all students were administered the *STAR Reading Version 2.0* (STAR Reading: Renaissance Learning, Inc., 2000, 2006a, b) test in a three-week period. The majority of students were tested in a single week. Anyone who was absent or missed the first assessment was followed up during the next two weeks. All students available took the test and there were no special criteria for including or excluding them. In mid-February, all available students were administered the Scholastic Reading Inventory-Interactive (SRI-I: Scholastic, Inc. 2001a, b, 2006) over a two-week period. All students remained in their classroom for all tests and school personnel used laptop carts to complete the STAR Reading and SRI-I assessments. All students took Florida Comprehensive Assessment Test (FCAT: Florida Department of Education, 2002) in

May. Scores for the three reading measures were described and compared to address our research questions.

Measures

STAR Reading is a norm-referenced and criterion-referenced computer adaptive test that is available for students in grades 1-12; we used scaled scores for sixth, seventh, and eighth grade students in this research. The test is timed and usually takes less than 10 minutes to complete. Reading levels are provided relative to national norms which allow teachers to quickly determine appropriate student instructional level. Students who are offered the test are expected to have reading vocabulary of at least 100 words, which corresponds to the beginning reading skills level or above. The test consists of 25 items of multiple choices for all grades. Students of grades 1-2 are offered all 25 items of vocabulary-in-context, whereas students of grades 3-12 are offered 20 items of vocabulary-in-context and five authentic text passages. The test is computer-adaptive; that is, if a student answers one item correctly then the next item will be of increasing difficulty. Conversely, if the student misses the right answer, then the next item will be of lesser difficulty. The STAR Reading 2.x and higher has 1,159 vocabulary-in-context items and 250 authentic text passage items. This makes it possible to use the test as a diagnostic tool to measure students' progress and administer the test to the same group of students five times a year without repeating the items.

According to the STAR Reading Technical Manual (Renaissance Learning, 2006b), each vocabulary-in-context item is a complete sentence that requires students to actually interpret meaning to identify the correct answer. The vocabulary-in-context section is also used to determine the initial difficulty level of authentic text passages.

The test provides grade equivalent, normal curve equivalent, and scaled scores. Additionally, it provides information about the zone of proximal development which indicates the lowest and highest range a student can read. The test software can also generate reports for teachers and parents.

Salvia, Ysseldyke, and Bolt (2006) reported that the test-retest reliabilities of STAR Reading varied from .85 to .95 for scaled scores, and from .79 to .91 for instructional reading level. A total of 34,446 students were tested twice with the interval of about five days between the first and the second test. The validity was established by correlating STAR Reading to other standardized tests. It was found that STAR Reading scores correlate closely to the scores of other reading measures such as: California Achievement Test, Comprehensive Test of Basic Skills, Degrees of Reading Power, Gates-MacGinitie, Iowa Test of Basic Skills, Metropolitan Achievement Test, and Stanford Achievement Test. Some custom-built state tests were also used. They include such states as: Connecticut, Texas, Indiana, Tennessee, Kentucky, North Carolina, and New York (Salvia et al., 2006).

SRI-I is a computer-adaptive test that is designed to assess student's reading comprehension level with texts of increasing difficulty (Scholastic, Inc., 2006). It usually lasts 20-30 minutes. The test ends after enough questions have been answered to compute a Lexile score for the student; we used these scores for sixth, seventh, and eighth grade students in this research. Students can print and view their Recommended Reading reports. The test uses authentic written materials and usually consists of 20-25 questions but no more than 30. The test bank contains more than 4,500 questions, which allows creating a unique test each time. The test measures such reading comprehension skills as referring to

details in the passage, drawing conclusions, and making comparisons and generalizations.

SRI-I is administered to K-12 students and uses the Lexile Framework for Reading (Knutson, 2006; Schnick & Knickelbine, 2000; Scholastic, n.d.). The Lexile measure is criterion-referenced and indicates the reading level of a particular student and that student's reading growth. The Lexile Scale for SRI scores ranges from 0 to 1,700. Comparing to Grade Equivalent, Grade Levels and Lexile Levels can be represented as follows: Grade 1-200 to 400; Grade 2-300 to 600; Grade 3-500 to 800; Grade 4-600 to 900; Grade 5-700 to 1,000; Grade 6-800 to 1,050; Grade 7-850 to 1,100; Grade 8-900 to 1,150; Grade 9-1,000 to 1,200; Grade 10-1,010 to 1,205; Grade 11-1,050 to 1,210; and, Grade 12-1,075 to 1,275. The Lexile measure was associated with other standardized tests such as Stanford 9 (SAT9), the North Carolina End-of-Grade Test, Stanford Diagnostic Reading Test (SDRT). SRI-I has been administered to more than three million students of all grades over the last five years.

SRI-I test-retest reliability was .89 (Renaissance Learning Inc., 2000). Knutson (2006) reported test-retest correlations for grades 3-10 students ranging from .81 to .85. The test was administered first in the fall and then in the spring. It was also administered to second graders in the spring and then to third graders in the fall. Correlation in this case went down to .78. SRI-I criterion-related validity was determined by correlating both spring and fall SRI-I scores to the spring 2002 FCAT-SSS Reading scores. The fall-to-spring correlations for grades 3 through 10 were in the range of .71 to .76, whereas spring-to-spring correlations ranged from .75 to .82.

The FCAT in Reading consists of two parts: criterion-referenced tests (CRT) assessing selected benchmarks in reading

from the Sunshine State Standards (SSS) and norm-referenced tests (NRT) in reading assessing individual student performance in regards to national standards. Multiple choice items are used for grades 3 through 10. Additional short response items are administered at grades 4, 8, and 10. For each grade, the reading scores range from 100 to 500 points. According to the 2004 assessment, internal consistency reliability on the reading test varied from .87 to .91. Criterion-related validity for the same year was determined by correlating FCAT-SSS reading scores with the FCAT-NRT (Stanford-9) scores. According to the 2004 assessment, the correlation between the two tests was in the range of .80 to .84. National percentile rank scores are also available for FCAT. We used current grade FCAT scaled scores in reading in our analyses.

Design and Data Analysis

We used a cross-sectional design to document similarities and differences within and between group performances and relationships between them on two progress monitoring assessments and end-of-grade achievement scores. Multivariate analysis of variance (MANOVA), multiple linear regression, Pearson Correlation, and predictive discriminant analysis were employed in statistical analyses.

Results

We report three types of outcomes. First, we provide descriptive and inferential findings to illustrate levels of performance across the measures between groups of students at each grade level participating in the study. Second, we report simple correlations across measures within grades. Third, we describe predictive analyses of relationships between progress monitoring assessments and statewide achievement test performance.

Descriptive Comparisons

Means and standard deviations for performance of sixth, seventh, and eighth grade students on reading measures across different demographic groups are in Table 1. Since FCAT scores are not comparable

across grades, we completed a series of MANOVAs to document the extent reading performance was statistically similar for sixth, seventh, and eighth grade students on different measures of progress monitoring and end-of-grade achievement for different demographic groups.

Table 1
Means and Standard Deviations across Comparison Group

Grade	Group	Subgroup	n	Test					
				STAR		SRI-I		FCAT	
				Mean	SD	Mean	SD	Mean	SD
Six	Gender	Female	201	601.57	260.89	829.49	257.11	300.30	58.48
		Male	209	625.67	264.60	809.85	282.90	295.98	67.74
	Ethnicity	Asian	8	935.38	246.61	1095.13	211.70	383.75	66.21
		African American	233	558.28	237.20	772.02	259.93	288.42	58.90
		Hispanic	76	580.04	217.78	792.96	257.77	288.83	67.53
		Multi-racial	10	606.90	197.10	933.20	204.20	321.00	35.55
		Caucasian	83	770.70	295.76	936.70	272.85	322.73	61.37
	Lunch	FRL	224	588.89	228.63	802.79	245.88	292.53	58.68
		Non-FRL	186	643.93	296.56	839.58	296.71	304.81	68.06
		Total	410	613.86	262.74	819.48	270.41	298.10	63.33
Seven	Gender	Female	155	711.48	276.64	898.98	245.20	311.00	65.39
		Male	176	754.56	287.40	928.26	244.39	307.14	73.51
	Ethnicity	Asian	14	955.79	255.57	1095.50	148.12	347.57	41.52
		African American	166	643.29	241.53	865.36	229.55	293.11	70.11
		Hispanic	59	696.86	304.62	888.85	311.98	299.20	68.51
		Multi-racial	4	913.00	258.46	1077.50	139.78	345.25	35.68
		Caucasian	88	888.03	22.27	988.36	203.85	337.56	63.48
	Lunch	FRL	165	685.83	250.34	881.15	238.29	297.96	73.84
		Non-FRL	166	783.23	305.12	948.15	247.44	319.99	63.72
		Total	331	734.38	282.81	914.55	244.84	308.95	69.75
Eight	Gender	Female	158	779.95	298.26	953.72	264.60	312.78	48.03
		Male	178	804.10	293.82	912.21	273.55	299.12	58.83
	Ethnicity	Asian	7	1123.43	133.56	1187.14	119.31	358.00	11.86
		African American	175	677.93	244.97	848.77	247.58	284.91	49.46
		Hispanic	67	784.09	280.73	955.76	241.97	315.69	42.58
		Multi-racial	4	1041.50	318.46	1014.50	102.32	337.25	52.34
		Caucasian	83	1001.94	276.84	1061.72	283.78	334.92	44.18
	Lunch	FRL	190	747.91	290.17	910.52	262.81	298.32	52.59
		Non-FRL	146	851.10	293.66	959.33	277.05	314.95	49.20
		Total	336	792.74	295.72	931.73	269.77	305.54	51.56

Gender. Box's test of the assumption of equality of covariance

matrices across gender was non-significant for sixth ($M = 10.02, p > .01$), seventh ($M =$

3.55, $p > .01$), and eighth ($M = 29.90$, $p < .01$) grades. Using Pillai' trace, there was a non-significant difference for gender on STAR Reading, SRI-I, and FCAT performance in sixth grade, $V = .02$, $F(3, 406) = 2.43$, $p > .01$, and seventh grade, $V = .02$, $F(3, 327) = 1.77$, $p > .01$. A statistically significant gender effect was indicated for eighth grade, $V = .06$, $F(3, 332) = 6.68$, $p < .01$; however, univariate follow-up tests revealed non-significant differences between eighth grade girls and boys for STAR Reading, $F(1, 334) = 0.56$, $p > .01$, SRI-I, $F(1, 334) = 1.64$, $p > .01$, and for FCAT, $F(1, 334) = 5.97$, $p > .01$.

Ethnicity. Box's test of the assumption of equality of covariance matrices across ethnicity was non-significant for sixth ($M = 27.25$, $p > .01$) and significant for seventh ($M = 55.95$, $p < .01$) and non-significant for eighth ($M = 47.29$, $p > .01$) grade. Using Pillai' trace, there was a significant difference for ethnicity on STAR Reading, SRI-I, and FCAT performance for sixth grade, $V = .16$, $F(12, 1215) = 1.61$, $p < .01$, for seventh grade, $V = .18$, $F(12, 978) = 5.28$, $p < .01$, and for eighth grade, $V = .28$, $F(12, 993) = 8.60$, $p < .01$.

Univariate follow-up tests revealed differences in STAR Reading scores, $F(4, 405) = 15.16$, $p < .01$, for sixth grade students from different ethnic backgrounds; scores for students from Asian ($M = 935.38$), Caucasian ($M = 770.70$), and multi-racial ($M = 606.90$) ethnic backgrounds were statistically different from their peers from Hispanic ($M = 580.04$) and African American ($M = 558.28$) ethnic backgrounds. Univariate follow-up tests revealed differences in SRI-I scores, $F(4, 405) = 10.47$, $p < .01$, for sixth grade students from different ethnic backgrounds; scores for students from Asian ($M = 73.13$), Caucasian ($M = 57.72$), and multi-racial ($M = 55.90$) ethnic backgrounds were statistically

different from their peers from Hispanic ($M = 44.05$) and African American ($M = 42.66$) ethnic backgrounds. Univariate follow-up tests revealed differences in FCAT end-of-grade achievement scores, $F(4, 405) = 9.65$, $p < .01$, for sixth grade students from different ethnic backgrounds; scores for students from Asian ($M = 383.75$), Caucasian ($M = 322.73$), and multi-racial ($M = 321.00$) ethnic backgrounds were statistically different from their peers from Hispanic ($M = 288.83$) and African American ($M = 288.42$) ethnic backgrounds.

Univariate follow-up tests revealed differences in STAR Reading scores, $F(4, 326) = 16.09$, $p < .01$, for seventh grade students from different ethnic backgrounds; scores for students from Asian ($M = 955.79$), multi-racial ($M = 913.00$), and Caucasian ($M = 888.03$) ethnic backgrounds were statistically different from their peers from Hispanic ($M = 696.86$) and African American ($M = 643.29$) ethnic backgrounds. Univariate follow-up tests revealed differences in SRI-I scores, $F(4, 326) = 7.12$, $p < .01$, for seventh grade students from different ethnic backgrounds; scores for students from Asian ($M = 65.29$), multi-racial ($M = 63.00$), and Caucasian ($M = 55.13$) ethnic backgrounds were statistically different from their peers from Hispanic ($M = 47.32$) and African American ($M = 43.54$) ethnic backgrounds. Univariate follow-up tests revealed differences in FCAT scores, $F(4, 326) = 8.12$, $p < .01$, for seventh grade students from different ethnic backgrounds; scores for students from Asian ($M = 347.57$), multi-racial ($M = 345.25$), and Caucasian ($M = 337.53$) ethnic backgrounds were statistically different from their peers from Hispanic ($M = 299.20$) and African American ($M = 293.11$) ethnic backgrounds.

Univariate follow-up tests revealed differences in STAR Reading scores, $F(4, 331) = 25.77$, $p < .01$, for eighth grade students from different ethnic backgrounds;

scores for students from Asian ($M = 1123.43$), multi-racial ($M = 1042.50$), and Caucasian ($M = 1000.94$) ethnic backgrounds were statistically different from their peers from Hispanic ($M = 784.09$) and African American ($M = 677.93$) ethnic backgrounds. Univariate follow-up tests revealed differences in SRI-I scores, $F(4, 331) = 14.83$, $p < .01$, for eighth grade students from different ethnic backgrounds; scores for students from Asian ($M = 69.57$), Caucasian ($M = 58.99$), and multi-racial ($M = 52.00$) ethnic backgrounds were statistically different from their peers from Hispanic ($M = 47.84$) and African American ($M = 38.80$) ethnic backgrounds. Univariate follow-up tests revealed differences in FCAT scores, $F(4, 331) = 20.42$, $p < .01$, for eighth grade students from different ethnic backgrounds; scores for students from Asian ($M = 358.00$), multi-racial ($M = 7.25$), and Caucasian ($M = 334.92$) ethnic backgrounds were statistically different from their peers from Hispanic ($M = 315.69$) and African American ($M = 284.91$) ethnic backgrounds.

Federal lunch status. Box's test of the assumption of equality of covariance matrices across federal lunch status was significant for sixth ($M = 21.58$, $p < .01$) and non-significant for seventh ($M = 15.73$, $p > .01$) and eighth ($M = 5.34$, $p > .01$) grades. Using Pillai's trace, there was a non-significant difference for federal free lunch status on STAR Reading, SRI-I, and FCAT performance in sixth grade, $V = .01$, $F(3, 406) = 1.61$, $p > .01$, and a significant difference for seventh grade, $V = .04$, $F(3, 327) = 3.90$, $p < .01$ and eighth grade, $V = .04$, $F(3, 332) = 6.68$, $p < .01$. Univariate follow-up tests revealed significant differences for STAR Reading, $F(1, 329) = 10.09$, $p < .01$, between seventh grade students receiving free or reduced lunch ($M = 685.83$) and their peers not receiving free or reduced lunch ($M = 783.23$), for SRI-I,

$F(1, 329) = 7.33$, $p < .01$, between seventh grade students receiving free or reduced lunch ($M = 45.22$) and their peers not receiving free or reduced lunch ($M = 51.70$), and for FCAT end-of-grade achievement, $F(1, 329) = 8.44$, $p < .01$, between seventh grade students receiving free or reduced lunch ($M = 297.96$) and their peers not receiving free or reduced lunch ($M = 319.99$). Univariate follow-up tests revealed significant differences for STAR Reading, $F(1, 334) = 10.33$, $p < .01$, between eighth grade students receiving free or reduced lunch ($M = 747.91$) and their peers not receiving free or reduced lunch ($M = 851.10$), for SRI-I, $F(1, 334) = 3.53$, $p < .01$, between eighth grade students receiving free or reduced lunch ($M = 44.32$) and their peers not receiving free or reduced lunch ($M = 49.08$), and for FCAT end-of-grade achievement, $F(1, 334) = 8.78$, $p < .01$, between eighth grade students receiving free or reduced lunch ($M = 298.32$) and their peers not receiving free or reduced lunch ($M = 314.95$).

Correlation Comparisons

Correlation coefficients for STAR Reading, SRI-I, and FCAT scores for sixth, seventh, and eighth grade students are in Table 2. Relationships were stronger for Grade 6 students than for their peers in Grade 7 or Grade 8. For Grade 6 students, the correlation coefficients were similar with each of the tests explaining about 57% of the variance of the others. For Grade 7 students, (1) STAR Reading explains about 50% of the variance of SRI-I and 53% of the variance of FCAT; (2) SRI-I explains about 50% of the variance of STAR Reading and 49% of the variance in FCAT; and (3) FCAT explains about 49% of the variance of SRI-I and 53% of the variance of STAR Reading. For Grade 8 students, (1) STAR Reading explains about 41% of the variance of SRI-I and 54% of the variance of FCAT;

(2) SRI-I explains about 41% of the variance of STAR Reading and 48% of the variance in FCAT; and (3) FCAT explains

about 48% of the variance of SRI-I and 54% of the variance of STAR Reading.

Table 2

Correlation Coefficients for SRI-I, STAR Reading, and FCAT Scaled Scores across Grades

	<i>STAR Reading</i>	<i>SRI-I</i>	<i>FCAT</i>
Grade 6 (<i>n</i> = 410)			
STAR Reading	--	.76	.75
SRI-I		--	.76
FCAT			--
Grade 7 (<i>n</i> = 331)			
STAR Reading	--	.73	.61
SRI-I		--	.58
FCAT			--
Grade 8 (<i>n</i> = 336)			
STAR Reading	--	.67	.71
SRI-I		--	.68
FCAT			--

Note. All correlation coefficients are statistically significant at $p < .01$.

Concerned with the differences among the subgroups with regard to students' social economic status and ethnicity (Roehrig et al., 2008), we correlated the scaled scores for STAR Reading, SRI-I, and FCAT end-of-grade achievement for each group of participants by school lunch status and ethnicity. The correlation coefficients ranged from .707 to .754 for regular-lunch students; ranged from .708 to .729 for free/reduced price lunch students; ranged from .675 to .704 for African American; ranged from .732 to .761 for Hispanic; and ranged from .680 to .714 for Caucasian. The relationship between STAR Reading, SRI-I, and FCAT were found to be invariant across groups classified by student lunch status and ethnicity using the *t* test suggested by (Bruning & Kintz, 1997).

Predictive Comparisons

Criterion-referenced reading scaled scores were used in multiple linear regressions to provide additional estimates

and predictors of FCAT scores. Standardized coefficients as well as partial correlation coefficients of the independent variables were compared to determine the best predictor of FCAT scores. Since STAR Reading and SRI-I are highly correlated with each other, multicollinearity was examined before each variable was entered into the regression. The variance inflation factor (VIF) was 2.047 and the tolerance value was 0.488. According to Myers (1990), multicollinearity would be of a concern if the VIF is larger than 10. According to Lynch (2003), multicollinearity is a problem when the sample size was small and the model had considerable error. In addition, Lynch (2003) further pointed out three classic symptoms of multicollinearity: (1) significant *F* without significant *t*-ratios, (2) wildly changing estimates when an additional/collinear variable was included in a model, and (3) the estimates of the coefficients were unreasonable. None of these occurred in the current data; therefore,

multicollinearity was not concerned in the following analyses.

We were interested in what variables were good predictors of FCAT scores. Student demographic information (gender, ethnicity, free/reduced price lunch status) as well as their performance on STAR and SRI-I were entered into the regression as independent variables with a stepwise regression procedure. Variables representing student demographic information were recoded into dichotomous variables: gender was coded as 0 and 1 where 1 represents male and 0 represents female. Student eligibility for free/reduced price lunch program was coded as 0 and 1 where 1 represents eligible for this program and 0 represents not eligible for this program. Student ethnicity was coded into three variables African American (1) or not (0), Caucasian (1) or not (0), and Hispanic (1) or not (0). Asians and multi-racial students were not included because the sample size for the interaction effects with these groups of students is extremely small (sometimes less than 1).

The interaction effect between gender and the variable representing African American ethnicity or not was the only statistically significant interaction noted: $t = -4.46$, $p < .001$. As a result, the model was tested for male and female students

separately (Table 3) and STAR Reading was entered into the model first (Model 1). In Model 2, both STAR Reading and SRI-I were included. In Model 3, all three predictors (STAR, SRI-I, and African American ethnicity) were entered into the model. We used this hierarchical approach to examine the amount of variance that was explained by each variable while taking into account the variance already explained in previous models (e.g., how much additional variance can be explained by SRI-I when considering the variance that has been accounted for by STAR). For male students, $R^2 = .50$ for Model 1 when STAR was the only significant predictor; $R^2 = .57$ for Model 2 when SRI-I was also a significant predictor; $R^2 = .58$ for Model 3 when all three variables (STAR, SRI-I, and African American or not) are significant predictors. The change of R^2 was .50 for Model 1, .07 for Model 2, and .01 for Model 3. Each of the change of R^2 was statistically significantly different from zero. For female students, $R^2 = .57$ for Model 1 when STAR was the only significant predictor; $R^2 = .66$ for Model 2 when SRI-I was also a significant predictor. The change of R^2 was .57 for Model 1 and .09 for Model 2. Each of the change of R^2 was also statistically significantly different from zero.

Table 3
Stepwise Estimates of Coefficients for the Multiple Regressions on FCAT

		<i>B</i>	<i>SE B</i>	β	<i>t</i>	<i>p</i>	<i>Partial</i>	<i>Part</i>
Male								
Model 1	STAR	0.148	0.006	0.708	23.64	<.001	.708	.708
Model 2	STAR	0.092	0.008	0.439	11.041	<.001	.424	.307
	SRI-I	0.083	0.009	0.376	9.451	<.001	.372	.263
Model 3	STAR	0.087	0.008	0.415	10.207	<.001	.398	.283
	SRI-I	0.084	0.009	0.377	9.524	<.001	.375	.264
	Black	-8.902	3.541	-0.073	-2.514	.012	-.106	-.070
Female								
Model 1	STAR	0.144	0.006	0.753	25.865	<.001	.753	.753
Model 2	STAR	0.084	0.007	0.439	11.830	<.001	.464	.306
	SRI-I	0.092	0.008	0.438	11.800	<.001	.463	.305

Note. The adjusted R^2 for the final model is .573 for male students and .659 for female students.

The final model fits quite well, $F(3, 555) = 250.11$ for male students and $F(2, 509) = 494.80$ for female students. The adjusted R^2 value was .57 for male students and .66 for female students, suggesting that the percentage of the variance of FCAT that could be explained by the predictors was 57% for male students and 66% for female students. Differences were noted between male and female students: African American male students had a statistically significantly lower performance on FCAT in comparison to non-African American male students ($t = -2.51, p = .01$); however, this difference was not statistically significant for female students ($t = 0.10, p = .92$). All other variables representing student demographic information were excluded because they did not meet the inclusion criterion: Probability of F -to-enter is less than or equal to .05.

The estimates of the standardized coefficients are interpreted for male and female students, respectively. For male

students, a unit increase in STAR Reading scores would result in 0.42 unit of increase in FCAT scores after controlling for SRI-I and student ethnicity of African American or not whereas a unit increase in SRI-I scores would result in 0.38 unit of increase in FCAT scores after controlling for STAR Reading and student ethnicity of African American or not. The partial correlation coefficients indicated that a unit increase in STAR Reading scores would result in 0.40 unit of increase in FCAT scores after removing the linear effect of SRI-I scores on both FCAT and STAR Reading. Similarly, a unit increase in SRI-I scores would result in 0.38 unit of increase in FCAT scores after removing the linear effect of STAR Reading scores on both FCAT and SRI-I. The part correlation coefficients indicated that a unit increase in STAR Reading scores would result in 0.28 unit of increase in FCAT scores after removing the linear effect of SRI-I scores on STAR Reading only.

Similarly, a unit increase in SRI-I scores would result in 0.26 unit of increase in FCAT scores after removing the linear effect of STAR Reading scores on SRI-I only. For female students, a unit increase in STAR Reading scores would result in 0.44 unit of increase in FCAT scores after controlling for SRI-I whereas a unit increase in SRI-I scores would result in 0.44 unit of increase in FCAT scores after controlling for STAR Reading. The partial correlation coefficients indicated that a unit increase in STAR Reading scores would result in 0.46 unit of increase in FCAT scores after removing the linear effect of SRI-I scores on both FCAT and STAR Reading. Similarly, a unit increase in SRI-I scores would result in 0.46 unit of increase in FCAT scores after removing the linear effect of STAR Reading scores on both FCAT and SRI-I. The part correlation coefficients indicated that a unit increase in STAR Reading scores would result in 0.31 unit of increase in FCAT scores after removing the linear effect of SRI-I scores on STAR Reading only. Similarly, a unit increase in SRI-I scores would result in 0.31 unit of increase in FCAT scores after removing the linear effect of STAR Reading scores on SRI-I only. All these coefficient estimates suggested that both STAR Reading and SRI-I assessments are good predictors of FCAT.

Finally, to assess the accuracy of prediction, predictive discriminant analysis (PDA) was employed to measure how well

$$\text{Hit rate} = \frac{TP + TN}{N}; \text{Sensitivity} = \frac{TP}{TP + FN}; \text{and Specificity} = \frac{TN}{TN + FP} \quad (\text{Hosp \& Fuchs, 2005}).$$

Cut-off scores for adequate or inadequate mastery of skills are suggested by the manuals for each test (Florida Department of Education, 2002; Renaissance Learning, Inc., 2006b; Scholastic, Inc., 2006): FCAT (296 for sixth graders, 300 for seventh graders, and 310 for eighth graders); STAR

STAR and SRI-I predicted the students' performance on FCAT based upon FCAT achievement levels. Participants were put into two groups according to the technical report of Florida Center for reading Research (Buck & Torgesen, n.d.): adequate (Levels 3-5) and inadequate (Levels 1-2). True positive (TP), true negative (TN), false positive (FP), and false negative (FN) were counted based upon the results of PDA. TP refers to students who did not master the skill and were predicted as not having mastered the skill. TN refers to students who mastered the skill and were predicted as having mastered the skill. FP refers to students who had mastered the skill but were predicted as not having mastered the skill. FN refers to students who did not master the skill but were predicted as having mastered the skill. Hit rate, sensitivity, and specificity indices were calculated for each PDA to reflect the accuracy of PDA. Hit rate provides an overall indication of how well STAR and SRI-I predicted students' performance on FCAT, sensitivity reflects how well STAR and SRI-I identified students who did not master the skills measured by FCAT, and specificity suggests how well STAR and SRI-I identified students who mastered the skills measured by FCAT. The formulas to calculate these indices were as follows:

(638 for sixth graders, 781 for seventh graders, and 878 for eighth graders); SRI-I (800 for sixth graders, 850 for seventh graders, and 900 for eighth graders). Results in Table 4 indicate that both STAR and SRI-I are accurate in predicting students' performance on FCAT (the average of the

hit rate is 76% across the grades). Specifically, STAR and SRI-I are more accurate in predicting students who are considered “adequate” by FCAT (the average of the specificity is 88% across

grades) than predicting students who are considered “inadequate” by FCAT (the average of the sensitivity is 70% across grades).

Table 4

Hit rates, Sensitivity, and Specificity Indices for STAR and SRI-I Predicting FCAT Mastery

Reading Skill	TP	FP	TN	FN	Hit Rate (%)	Sensitivity (%)	Specificity (%)
Grade 6 (<i>n</i> = 410)							
STAR	187	15	133	75	78	71	90
SRI-I	201	1	87	121	70	62	99
Grade 7 (<i>n</i> = 326)							
STAR	113	9	122	82	72	58	93
SRI-I	73	49	169	35	74	68	78
Grade 8 (<i>n</i> = 335)							
STAR	160	12	116	47	82	77	91
SRI-I	124	48	141	22	79	85	75

Note. TP = true positive; FN= false negative; TN = true negative; FP = false positive.

Hit rate = $(TP + TN)/n$; sensitivity = $TP/(TP + FN)$; specificity = $TN/(TN + FP)$

Discussion

Data-driven accountability has reached a new level under mandates and directives in the NCLB Act of 2004 (U.S. Department of Education, 2004). State and local education agencies are making school personnel test students regularly with screening, diagnostic, progress-monitoring, and high stakes outcome measures (Hasbrouck & Tindal, 2006). As Crawford, Tindal, and Stieber (2001) indicate, the widespread adoption of statewide tests as markers of academic proficiency and an implied link to school quality have made it important that students’ academic *progress* be closely monitored for at least four reasons:

1. Statewide testing programs often involve a format that is difficult for teachers to replicate at the classroom level.
2. Decisions being made are so important that other confirming

information is needed to complement the data.

3. Teachers need other performance indicators related to statewide tests that are available more frequently so that instructional programs can be improved in a timely fashion.

4. Statewide tests may be insensitive to change for low-performing students. (p. 304)

Interest in relationships between progress monitoring measures and high stakes achievement tests has a long and renewing history (Council of Chief State School Officers, n.d.; Coyne & Harn, 2006; Crawford et al., 2001; Deno, Mirkin, Chiang, & Lowry, 1980; Fuchs & Deno, 1981; Linn, 2000; Marston, Deno, & Tindal, 1983; Perie et al., 2007; Roehrig et al., 2008; Schatschneider et al., 2004; Schilling et al., 2007; Sibley et al., 2001; Stecker & Fuchs, 2000).

The purpose of the current study was to examine levels of performance on and relationships between performance on two progress monitoring measures and a statewide end-of-grade achievement test. We reported and compared performances on the STAR Reading, SRI-I, and FCAT end-of-grade achievement across grades, lunch status, and ethnicity. We also examined relationships between and among these measures. Our work adds to extant knowledge by responding to the need for research evaluating the usefulness of monitoring learning progress and predicting high stakes performance as important “social consequences” and by extending prior research on these relationships beyond elementary school students and measures of oral reading fluency (Roehrig et al., p. 362).

Education policy and practice driven by the NCLB Act requires school personnel to disaggregate data on school outcomes by race; and, it is easy to find evidence in government and other reports illustrating significant differences between test scores and other indicators of educational engagement and success for students representing some ethnic minorities and their Caucasian peers (cf. Fang, 2010; Ladson-Billings, 2006; National Center for Education Statistics, 2001, 2007, 2009; Uzzell, Simon, Horwitz, Hyslop, Lewis, & Casserly, 2010; Warikoo & Carter, 2009; Wiggan, 2007). Our comparisons of the performance of sixth, seventh, and eighth grade students on different measures of reading achievement supported these widely-recognized trends. We also documented that two widely-used progress monitoring assessments were good predictors of end-of-grade achievement at three different grade levels and that predictive bias across different demographic groups was minimal. In general, our correlational findings were similar to those of other researchers using different progress

monitoring and outcome measures with younger children (cf. Hintze et al., 2002, 2003; Hixson & McGlinchey, 2004; Roehrig et al., 2008).

Implications for Improvement of Practice

We agree with Roehrig et al. (2008) that “it is essential that educators be provided with precise student achievement data and benchmarks if the rigorous grade-level reading standards set forth in accountability policies are to be met by all students” (p. 362). Our findings indicated that STAR Reading and SRI-I were good predictors of end-of-grade achievement in grades 6, 7, and 8, and the usefulness of our findings was evident across different demographic groups. These outcomes were previously unavailable in research on relationships between progress monitoring measures and statewide assessment outcomes completed in elementary schools in several states (cf. Barger, 2003; Buck & Torgesen, n.d.; Buck, Torgesen, & Schatschneider, n.d.; Hintze et al., 2002, 2003; Vander Meer et al., 2005; Wilson, 2005).

Because our finding of similar predictive relationships had limited practical value in making decisions about which progress monitoring measure to use, we conducted a *post hoc* analysis. According to marketing materials provided by Scholastic, the SRI-I usually takes 20-30 minutes to administer. Administration times were available in the data dictionary provided by the participating schools for STAR Reading but not for the SRI-I. We found that STAR Reading took an average of about seven minutes to administer. Most administrators and teachers believe that despite the value of regularly monitoring student progress there is too much time spent testing or preparing for tests. In our research, a test that required about eight minutes per child to administer did the same job predicting end-of-grade

performance compared with another test that developers believe would take three times as long. Accepting conservative estimates of test administrations every other month, using STAR Reading for progress monitoring would free over 20 minutes per student for critical instructional skills many teachers believe are trumped by testing; and, saving would be greater in schools doing monthly or more frequent progress monitoring.

Conclusion

Although our findings of consistent achievement differences across some demographic groups of middle school students and strong predictive relationships for two different progress monitoring measures and statewide end-of-grade achievement provide support for prior research as well for continued use in school-based decision making, we agree with Roehrig et al. (2008), Wiggan (2007), and Warikoo and Carter (2009) that more research is needed. Areas of clear extension for our study include investigations of similarities and differences in performance as well as relationships between scores on widely-used progress monitoring assessments and high stakes state-wide tests used in elementary schools. Research building on our finding of the potential differential benefits of progress monitoring systems with similar predictive capacity are also warranted and recommended.

Each year millions of students are at risk for serious and continued failure in school and many fail to make acceptable progress, especially when compared to their peers across different demographic groups (National Center for Education Statistics, 2001, 2007, 2009; U.S. Department of Education, 2006). We documented that the persistent and consistent differences evident for reading achievement in elementary school continue into middle school for many

students. We also found that progress monitoring measures administered during the school year predicted end-of-year performances very well. We believe that the value of documenting achievement and predictions of it is not in the magnitude of the differences or relationships that are revealed but in deriving direction for change from them. The best predictor of achievement in elementary school is prior performance in elementary school (Roehrig et al., 2008; Schilling et al., 2007; Wanzek et al., 2010); and, the best predictor of performance in middle school is performance in elementary school (Fang, 2010). Progress monitoring measures administered during sixth, seventh, and eighth grade were found to be strong predictors of end-of-grade achievement. Gaps on two formative measures of achievement for students from Asian and Caucasian ethnic backgrounds and their peers from Hispanic and African American ethnic backgrounds were also evident in summative end-of-grade achievement scores for these students. The implications for improving practice are clear: Continued use of progress monitoring measures such as STAR Reading are powerful tools in efforts to identify students needing assistance to persist and affect high stakes assessments.

References

- Ananda, S., & Rabinowitz, S. (2001). *High stakes assessment innovation: A negative correlation?* San Francisco, CA: WestEd. (Ed 462 446).
- Barger, J. (2003). *Comparing the DIBELS Oral Reading Fluency indicator and the North Carolina end of grade reading assessment (Technical Report)*. Asheville, NC: North Carolina Teacher Academy. Retrieved from https://dibels.uoregon.edu/techreports/NC_Tech_Report.pdf
- Brown, R. S., & Coughlin, E. (2007). *The predictive validity of selected benchmark assessments used in the Mid-Atlantic Region* (Issues & Answers Report, REL 2007–No. 017). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic. Retrieved from <http://ies.ed.gov/ncee/edlabs>
- Bruning, J. L., & Kintz, B. L. (1997). *Computational handbook of statistics*. New York, NY: Longman.
- Buck, J., & Torgesen, J. (n.d.). The relationship between performance on a measure of oral reading fluency and performance on the Florida Comprehensive Assessment Test. FCRR Technical Report #1. Tallahassee, FL: Florida Center for Reading Research. Retrieved from <http://www.fcrr.org/TechnicalReports/TechnicalReport1.pdf>
- Buck, J., Torgesen, J., & Schatschneider, C. (n.d.). Predicting FCAT-SSS scores using prior performance on the FCAT-SSS, FCAT-NRT, and SAT9. (FCRR Technical Report #4). Tallahassee, FL: Florida Center for Reading Research. Retrieved from http://www.fcrr.org/TechnicalReports/fcat_predicting.pdf
- Council of Chief State School Officers (CCSSO). (n.d.). *CCSSO and Renaissance Learning R & D Consortium Using Interim Assessments to Predict Student Proficiency*. Washington, DC: Author.
- Coyne, M. D., & Harn, B. A. (2006). Promoting beginning reading success through meaningful assessment of early literacy skills. *Psychology in the Schools*, 43, 33-43.
- Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment*, 7, 303-323.
- Deno, S. L. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure*, 35(2), 5-10.
- Deno, S. L. (1997). Whether thou goest . . . Perspectives on progress monitoring. In J. W. Lloyd, E. J. Kameenui, & D. Chard (Eds.), *Issues in educating students with disabilities* (pp. 77-99). Mahwah, NJ: Erlbaum.
- Deno, S. L. (2003). Developments in curriculum-based measurement. *Journal of Special Education*, 37, 184-192.
- Deno, S. L., & Mirkin, R. K. (1977). *Data-based program modification: A manual*. Reston, VA: Council for Exceptional Children.
- Deno, S. L., Mirkin, P. K., Chiang, B., & Lowry, L. (1980). *Relationships among simple measures of reading and performance on standardized achievement tests* (Res. Rep. No. 20). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.
- Fang, L. (2010). Are boys left behind? The evolution of the gender achievement gap in Beijing's middle schools. *Economics of Education Review*, 29, 383-399.

- Florida Department of Education. (2002). *Florida Comprehensive Assessment Test (FCAT) for Reading and Mathematics: Technical report for test administrations of FCAT 2002*. Tallahassee, FL: Florida Department of Education.
- Florida Department of Education. (n.d.). *Florida Comprehensive Assessment Test*. Retrieved from <http://fcats.fldoe.org/>
- Fuchs, D., & Deshler, D. D. (2007). What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research and Practice, 22*, 129-136.
- Fuchs, L. S. (1989). Evaluating solutions: Monitoring progress and revising intervention plans. In M. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 153-181). New York, NY: Guilford.
- Fuchs, D., & Fuchs, L. S. (2006). Introduction to Response to Intervention: What why, and how valid is it? *Reading Research Quarterly, 41*, 93-99.
- Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*, 188-192.
- Fuchs, L., & Deno, S. (1981). *The relationship between curriculum-based mastery measures and standardized achievement tests in reading* (Res. Rep. No. 57). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.
- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation on student achievement: A meta-analysis. *Exceptional Children, 53*, 199-208.
- Fuchs, L. S., & Fuchs, D. (1999). Monitoring student progress toward development of reading competence: A review of three forms of classroom-based assessment. *School Psychology Review, 28*, 659-671.
- Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher, 59*, 636-644.
- Hintze, J. M., Callahan, J. E., III, Matthews, W. J., Williams, S. A. S., & Tobin, K. G. (2002). Oral reading fluency and prediction of reading comprehension in African American and Caucasian elementary school children. *School Psychology Review, 31*, 540-553.
- Hintze, J. M., Ryan, A. L., & Stoner, G. (2003). Concurrent validity and diagnostic accuracy of the dynamic indicators of basic early literacy skills and the comprehensive test of phonological awareness. *School Psychology Review, 32*, 541-556.
- Hixson, M. D., & McGlinchey, M. T. (2004). The relationship between race, income, and oral reading fluency and performance on two reading comprehension measures. *Journal of Psychoeducational Assessment, 22*, 351-364.
- Hosp, M. K., & Fuchs, L. S. (2005). Using CBM as an indicator of decoding, word reading, and comprehension: Do the relations change with grade? *School Psychology Review, 34*, 9-26.
- Knutson, K. (2006). *Because you can't wait until spring: using the SRI to improve reading performance*, Scholastic Professional Paper. New York, NY: Scholastic Inc.
- Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U.S. schools. *Educational Researcher, 35*, 3-12.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*, 4-16.
- Lynch, S. M. (2003). *Multicollinearity*. Retrieved from http://www.princeton.edu/~slynch/SOC_504/multicollinearity.pdf
- Madelaine, A. & Wheldall, K. (2004). Curriculum-based measurement of reading: Recent advances. *International Journal of Disability, Development and Education, 51*, 57-82.

- Marston, D., Deno, S., & Tindal, G. (1983). *A comparison of standardized achievement tests and direct measurement techniques in measuring pupil progress* (Res. Rep. No. 126). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities. (ERIC Document Reproduction Service No. ED236198)
- Myers, R. (1990). *Classical and modern regression with applications* (2nd ed.). Boston, MA: Allyn & Bacon.
- National Center for Education Statistics. (2001). *Education achievement and Black-White inequality*. Washington, DC: Department of Education.
- National Center for Education Statistics. (2007). *Status and trends in the education of racial and ethnic minorities*. Washington, DC: Department of Education.
- National Center for Education Statistics. (2009). *The Nation's Report Card: Reading 2009*. Washington, DC: Department of Education.
- Perie, M., Marion, S., & Gong, B. (2007). *A framework for considering interim assessments*. Dover, NH: National Center for the Improvement of Educational Assessment. Retrieved from http://www.nciea.org/publications/ConsideringInterimAssess_MAP07.pdf
- Renaissance Learning, Inc. (2000). *Comparison of the STAR Reading Computer-Adaptive Test and the Scholastic Reading Inventory Test*. Wisconsin Rapids, WI: Author.
- Renaissance Learning, Inc. (2006a). *STAR Reading Computer-Adaptive Reading Test and Database: Software manual*. Wisconsin Rapids, WI: Author.
- Renaissance Learning, Inc. (2006b). *STAR Reading Computer-Adaptive Reading Test and Database: Technical Manual*. Wisconsin Rapids, WI: Author.
- Roehrig, A. D., Petscher, Y., Nettles, S. M., Hudson, R. F., & Torgesen, J. K. (2008). Accuracy of the DIBELS Oral Reading Fluency measure for predicting third grade reading comprehension outcomes. *Journal of School Psychology, 46*, 343-366.
- Safer, N. & Fleischman, S. (2005). How student progress monitoring improves instruction. *Educational Leadership, 62*, 81-83.
- Salvia, J., Ysseldyke, J. E., & Bolt, S. (2006) *Assessment in special and inclusive education*. Boston, MA: Houghton Mifflin Company.
- Schatschneider, C., Buck, J., Torgesen, J., Wagner, R., Hassler, L., Hecht, S., & Powell-Smith, K. (2004). A multivariate study of individual differences in performance on the reading portion of the Florida Comprehensive Assessment Test: A brief report. Tallahassee, FL: Florida Center for Reading Research.
- Schilling, S. G., Carlisle, J. F., Scott, S. E., & Zeng, J. (2007). Are fluency measures accurate predictors of reading achievement? *The Elementary School Journal, 107*, 429-448.
- Schnick, T., & Knickelbine, M. (2000). *The lexile framework: An introduction for educators*. Durham, NC: MetaMetrics.
- Scholastic, Inc. (2001a). *Scholastic reading inventory interactive educator's guide*. New York, NY: Author.
- Scholastic, Inc. (2001b). *Scholastic Reading Inventory: Technical Guide*. New York, NY: Author.
- Scholastic, Inc. (2006). *Scholastic Reading Inventory Educator's Guide: An Overview of SRI Software and Teacher Support*. New York, NY: Author.
- Scholastic, Inc. (n.d.). Accuracy matters: Reducing measurement error by targeted. New York, NY: Author. Retrieved from http://teacher.scholastic.com/products/sri/pdfs/SRI_Accuracy_Professional_Paper.pdf

- Scott, V. G., & Weishaar, M. K. (2003). Curriculum-based measurement for reading progress. *Intervention in School and Clinic, 38*, 153–158.
- Sibley, D., Biwer, D., & Hesch, A. (2001, April). *Establishing curriculum-based measurement oral reading fluency performance standards to predict success on local and state tests of reading achievement*. Paper presented at the Annual Meeting of the National Association of School Psychologists, Washington, DC.
- Simpson, R. L., LaCava, P., & Graner, P. (2004). The No Child Left Behind Act: Challenges and implications for educators. *Intervention in School and Clinic, 40*, 67–75.
- Stecker, P. M., & Fuchs, L. (2000). Effecting superior achievement using curriculum-based measurement: The importance of individual progress monitoring. *Learning Disabilities Research and Practice, 15*, 128-134.
- Thurlow, M. L., & Thompson, S. J. (1999). District and state standards and assessments: Building an inclusive accountability system. *Journal of Special Education Leadership, 12*, 3–10.
- Thurlow, M. L., & Wiley, H. I. (2004). *Almost there in public reporting of assessment results for students with disabilities* (Technical Report 39). Minneapolis: University of Minnesota, National Center on Educational Outcomes.
- U.S. Department of Education. (2006). Proven methods: Early reading first and reading first. Retrieved from <http://www.ed.gov/nclb/methods/reading/readingfirst.html>
- U.S. Department of Education, Office of the Deputy Secretary. (2004). *No Child Left Behind: A toolkit for teachers*. Washington, DC: Author.
- Uzzell, R., Simon, C., Horwitz, A., Hyslop, A., Lewis, S., & Casserly, M. (2010). *Beating the odds: Analysis of student performance on state assessments and NAEP. Results from the 2008-09 School Year*. Washington, DC: Council of the Great City Schools.
- Vander Meer, C. D., Lentz, F. E., & Stollar, S. (2005). *The relationship between oral reading fluency and Ohio proficiency testing in reading* (Technical Report). Eugene, OR: University of Oregon. Retrieved from <https://dibels.uoregon.edu/techreports/ohio.pdf>
- Wanzek, J., Roberts, G., Linan-Thompson, S., Vaughn, S., Woodruff, A. L., & Murray, C. S. (2010). Differences in the relationship of oral reading fluency and high-stakes measures of reading comprehension. *Assessment for Effective Intervention, 35*, 67-77
- Warikoo, N., & Carter, P. (2009). Cultural explanations for racial and ethnic stratification in academic achievement: A call for new and improved theory. *Review of Educational Research, 79*, 366-394.
- Wiggin, G. (2007). Race, school achievement, and educational inequality: Toward a student-based inquiry perspective. *Review of Educational Research, 77*, 310-333.
- Wilson, J. (2005). *The relationship of Dynamic Indicators of Basic Early Literacy Skills (DIBELS) oral reading fluency to performance on Arizona Instrument to Measure Standards (AIMS)* (Technical Report). Tempe, AZ: Assessment and Evaluation Department, Tempe School District No.3. Retrieved from <https://dibels.uoregon.edu/techreports/arizona.pdf>
- Wixson, K. K., & Carlisle, J. F. (2005). The influence of large-scale assessment of reading comprehension on classroom practice: A commentary. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 395-405). Mahwah, NJ: Erlbaum.

Modeling the Effects of Home and Student Factors on Text Comprehension

Qiuying Wang, Oklahoma State University

Ze Wang, University of Missouri

Steven J. Osterlind, University of Missouri

Abstract

This study explored the effects of affective factors and home environments on children's text comprehension from a cultural perspective. About 3,000 British and more than 5,000 Hong Kong fourth grade students participated in the study (data derived from the Progress in International Reading Literacy Study, 2001). The participating students and their parents were administered a number of questionnaires concerning students' home environments, reading self-concept, and attitude toward reading. Students were also assessed in text comprehension. Methodologically, two factors of home environments (Early Home Literacy Activities and Parents' Attitude toward Reading) and two factors of student characteristics (Student's Reading Self-concept and Student's Attitude toward Reading) were modeled factorially, with reading comprehension as the dependent variable. A final model fits the data well, suggesting that students' home environments within their own culture influenced one's text comprehension. However, this influence was indirect and moderated by the student's reading self-concept and attitude toward reading. The findings of this study provide insights on how to improve reading instructions in the Chinese and Western cultures.

Introduction

Focusing education to help children and young adults learn to read is an effort common to most countries. Lost opportunity and consequent deficits in other academic areas, like mathematics, civics, and social and cultural arts, can stem from an inability

to read adequately, regardless of the country a child lives in. In an effort to understand more about reading across many countries, the International Association for the Evaluation of Educational Achievement (IEA) has undertaken an investigation of children's reading literacy and the factors associated with its acquisition. IEA's goals are implemented, in part, through the Progress in International Reading Literacy Study (PIRLS¹) 2001.

PIRLS 2001 focuses on the achievement of young children and identifying what affective experiences they have at home and school that may be associated with learning to read (PIRLS 2001 Database User Guide, p. 2). Affective factors are commonly identified as, but not limited to, reading attitudes, motivation, and reading self-concepts (Kush, Watkins, & Brookhart, 2005). It is held that these indirect and affective components of a child's experience can have a profound influence on their learning how to read (e.g., Bradley, Corwyn, Burchinal, Pipes McAdoo, & Garcia Coll, 2001; MacDonald & Cornwall, 1995; Noble, Farah, & McCandliss, 2006). Further, we know that reading-related affective experiences are different across nations (see Chen, Stevenson, Hayward, & Burgess, 1995; Ogbu, 1981; Wang & Guthrie, 2004).

The purpose of our study was to gain a better understanding of these affective and indirect influences on reading by exploring how they differ across cultures. Two regions of the world, one Western and the other Eastern were selected for the study: Britain and China. The sample from Britain included students from across the island

while the Chinese students came solely from the region of Hong Kong.

Review of the Literature

Reading is a purposeful activity that often involves choice, motivation, and a set of behaviors and attitudes (Guthrie, Wigfield, Metsala, & Cox, 1999; Wigfield & Guthrie, 1997). Ample research evidence suggests that positive attitudes and high motivation are key factors in reading success (Baker, Dreher, & Guthrie, 2000; Snow, Burns, & Griffin, 1998). In fact, a positive attitude toward reading may be among the most important of attributes possessed by a lifelong reader (PIRLS 2001 Database User Guide, p.18).

However, attitude is not without a context for its development. The acquisition of a positive reading attitude and constructive self-concept is certainly influenced by one's cultural experiences. Researchers who explore aspects of literacy between different ethnic heritages have concluded that among Chinese, Japanese, Caucasians, and African Americans, cultural values and beliefs about academic achievement strongly influence their attitudes and behaviors for academic tasks (Chen et al, 1995; Ogbu, 1981; Wang & Guthrie, 2004).

Of course, parents, too, play a fundamental role in transmitting cultural values and beliefs to their children. Baker, Scher, and Mackler (1997) provide empirical evidence that parental encouragement of children's reading is related to the child's attitude toward reading, regardless of socioeconomic status. And, Guthrie and Greaney (1991) suggested that parents who regularly bring their child to a library clearly value literacy, and their beliefs affect the child indirectly through their behaviors.

Children's print-related experiences at home also influence their motivation to

learn to read. For instance, research has consistently shown that school-aged children who have more opportunities to engage in literacy-related activities at home have more positive views about reading and engage in more leisure reading. For example, Hansen (1969) investigated the relations between fourth graders' reading attitudes, home experiences, and independent reading. Results showed that the home literacy environment was significantly related to independent reading. Similarly, Rowe (1991) and Rowe & Rowe (1992) studied 5,600 Australian students ranging in age from 5 to 14 years and found that reading activities at home predicted the reading attitudes of children in each age group (5-6, 7-8, 9-11, and 12-14 years), and the magnitude of the direct effect of home reading activity on achievement, mediated by the indirect effect of attitudes, increased with students' age. In a related study, Neuman (1986) maintained that fifth graders' leisure reading habits were related to parental encouragement of reading, even after controlling for gender and socioeconomic status (SES).

Parents' attitude toward reading and home literacy environments clearly vary across cultures. For example, in the Western culture, personal autonomy and self-reliance are valued (Oyserman, Coon, & Kimmelmeier, 2002). This attitude is manifested in having parents who encourage children to pursue their own goals and make their own choices (Chao, 1996; Hui & Triandis, 1986). In traditional Chinese culture, however, the parents' influence is different. Here, individuals are viewed as interdependent (Markus & Kitayama, 1991), and children's conformity to the expectations of their parents and teachers is most valued (Ho, 1986).

These cultural differences, suggested to us the plausible hypothesis that *family academic values* may be a stronger source of

reading motivation for Chinese students, while *individual recognition* may be more robust for the Western students. Moreover, the complicated nature of written Chinese may make parental input in the teaching of literacy different for the Chinese language than it is for Western languages, like English which is a hybrid of Romance and Latin base. Western languages have more regular grapheme-phoneme connections and a less complicated orthography.

Of particular importance to us is the fact that cross-cultural studies indicate that Chinese parents have higher expectations for academic performance and emphasize academic success more highly than do their Western counterparts (Mau, 1997; Stevenson & Lee, 1996). While this finding is significant to our work, more generally, the literature on cultural comparison across countries is relatively sparse. So far as we know, few published studies examine the cultural effects on relationships among student's reading attitude, self-concept, early home literacy activities, and parents' attitude toward reading and text comprehension in different countries. This study will help to fill in this gap in the literature on reading acquisition.

The Aim of the Study

The aim of this study was to investigate the role of home influences in understanding the commonalities and distinctions in children's reading attitudes and behaviors in relation to text comprehension, and explore cultural effects that may contribute to differences between them. A theoretical structural model is proposed that describes the direct and indirect relationships among early home literacy activities, parents' attitude toward reading, student's reading self-concept, student's attitude toward reading and their reading comprehension for information and for literacy.

Two questions guided our research. They are: (1) Do parental factors affect text comprehension through student characteristics of reading measured by attitude toward reading and reading self-concept? And, (2) Do these relationships vary between Eastern and Western cultures?

Method

Participants

Participants in our study were fourth graders and their parents from two countries: 5,050 from Hong Kong (a Special Administrative Region of China), and 3,156 from England. The data derived from the PIRLS international assessments, year 2001. PIRLS targets fourth graders because typically students at this age (most are 9 and 10 years old) are at an important transition point in their development as proficient readers. Most of the children in the British sample were born in 1990 and 1991 (34 and 66% respectively). Two exceptions were children who were born in 1989 and 1992 respectively. For the Hong Kong sample, the majority of the fourth graders were born in 1990 (12.5%) and 1991 (77.4%), with an additional small percentage of Hong Kong children born between 1986 and 1989, or in 1992 and 1993². About 97% of the students in both country samples reported gender information, with the split between boys and girls at about 50% each.

Measures

Three measuring instruments for study variables were selected from the PIRLS data: the reading test, a required student questionnaire, and a home survey.

Student questionnaire. The PIRLS Student Questionnaire appraises a variety of reading-related aspects of a student's home and school life, including classroom experiences and reading for homework, self-

perceptions, and attitudes toward reading, out-of-school reading habits, computer use, home literacy resources, and basic demographic information. From these PIRLS data, two scales were developed for this study. Each scale is a latent factor assessed with multiple items. The first scale (five items) is a measure of a student's attitude toward reading, and the second scale (three items) is a measure of a student's reading self-concept. The responses were Likert-type coded from 1 to 4 ("Agree a lot," to "Disagree a lot"). Prior to analysis, directionality of responses was addressed with positive statements being reverse coded.

Learning to Read (home) survey. The PIRLS Learning to Read (Home) Survey was completed by each participating student's parents or primary caregivers. This survey investigates child-parent literacy interactions, home literacy resources, parents' reading habits and attitudes, and home-school connections. The original items from the survey were used to develop two scales measuring early home literacy activities and parents' attitude toward reading. Each scale is a latent factor assessed with multiple items. The five items assessing early home literacy activities asked how often the parents or someone else at home did certain activities with the child before she/he began the fourth grade. Parents rated those five items on a three-point scale: often, sometimes, and never or almost never. The scale inquiring into parents' attitude toward reading has five items which were rated on a four-point scale from "Agree a lot" to "Disagree a lot." Prior to the analysis, we reverse coded items as necessary.

PIRLS reading assessment. The PIRLS assessment of reading literacy appraises two reading constructs, regardless

of whether the reading is done in or out of a school. They are: (1) reading for literacy experience or enjoyment, and (2) reading to acquire and use information. These tests are equally long. In reading for literacy experience, the PIRLS framework includes literacy texts that represent the types of narrative structures and language usages most common to nine year-old readers. The main form of literacy text used in the assessment is narrative fiction. In reading to acquire and use information, the PIRLS assessment includes both chronologically and logically structured informational texts, some of which incorporated various types of adjunct aids such as charts, pictures, and graphs.

Because the PIRLS is a long test with a very large number of questions covering a broad domain (i.e., reading comprehension), it is not administered in total to each student. Rather, a given examinee takes only a sample of whole test in a matrix sampling design. Matrix sample is a sampling procedure where many short and overlapping sections of the test are given to a large group of examinees. These short test portions are aggregated by test administrators to compose a whole test. As a consequence of the design, individual scores are not directly produced. Other programs employ a similar matrix sampling methodology, including the largest school-based testing program in the United States, National Assessment of Educational Progress (NAEP [cf. <http://nces.ed.gov/nationsreportcard/assessed> October 13, 2008]).

Still, individual scores can be estimated by use of multiple imputation, or plausible values methodology (see Mislevy, 1991). PIRLS makes use of multiple imputations to estimate individual scores. Because of uncertainty in the imputation process, PIRLS produced five imputed values (labeled "plausible values") for each

student on each measure of reading proficiency. In this study, the first plausible value was used as a measure of a student's reading achievement, consistent with parallel research (cf. Akiba, 2008; Wu, 2005). Two reading achievement scores were used: one that corresponds to each of the two PIRLS content areas, reading to acquire information and reading for literary purpose. The plausible values were divided by 100 so that the standard deviations of the

reading scores were close to that of the scale items.

Results

Descriptive Statistics

The means and standard deviations of survey variables used in the study are presented in Table 1. Recall that these values are on a five-point Likert-type scale.

Table 1

Descriptive Statistics for Survey Variables in the British and Hong Kong Samples

	British (N=3,156)		Hong Kong (N=5,050)	
	M	SD	M	SD
Student's attitude toward reading				
I read only if I have to.	2.58	1.16	2.43	1.09
I like talking about books with other people.	2.33	1.04	2.97	0.98
I would be happy if someone gave me a book as a present.	3.21	0.98	3.45	0.85
I think reading is boring.	3.05	1.13	3.13	1.01
I enjoy reading.	3.13	1.09	3.19	0.94
Student's reading self-concept				
Reading is very easy for me.	3.33	0.78	2.97	0.87
I do not read as well as other students in my class.	2.47	1.07	2.59	0.98
Reading aloud is very hard for me.	2.63	1.15	2.55	1.12
Early home literacy activities				
Read books	2.82	0.40	2.10	0.58
Tell stories	2.59	0.57	1.91	0.65
Sing songs	2.65	0.54	1.92	0.67
Play word games	2.32	0.58	1.68	0.60
Read aloud signs and labels	2.54	0.59	1.93	0.69
Parents' attitude toward reading				
I only read if I have to.	3.48	0.90	2.41	1.10
I like talking about books with other people.	2.95	0.96	2.80	0.96
I like to spend my spare time reading.	3.15	0.96	3.43	0.83
I read only if I need information.	3.29	0.99	2.39	1.05
Reading is an important activity in my home.	3.50	0.76	2.94	0.94
Student reading for information	5.43	0.08	5.38	0.06
Student reading for literacy	5.58	0.09	5.18	0.07

Confirmatory Factor Analysis

Multiple group structural equation modeling was applied to address the research questions. Specifically, a two-step modeling procedure was used to assess the measurement and structural models (see Anderson & Gerbing, 1988). In step one, confirmatory factor analysis (CFA) using maximum likelihood estimation was used to evaluate the measurement model for the adequacy of the hypothesized factor structure in each sample. With CFA, researchers are usually interested in exploring how the observed items correlate. Hence, latent factors and their variances and co-variances of observed variables are usually analyzed.

Our initial hypothesized model did not adequately reflect the data, a common finding in models with many variables and with a large sample size. Therefore, we respecified the model by incrementally adding correlated residuals until an acceptable fit was achieved. Co-variances between the four pairs of measure-variable residuals (i.e., playing word games at home versus reading aloud signs and labels at home; parents read if have to versus parents read only if need information; students read if have to versus students think reading is boring; students like talking about books with others versus students like to have books as presents) were added to the model for British students.

For Hong Kong students, the same four pairs of correlated residuals, plus an additional correlation between a pair of measured-variable residuals (students like talking about books with others versus students think reading is boring), were added to the model. Three paths (Early Home Literacy Activities to playing word games; Student's Attitude toward Reading to students like to have books as presents; Student's Attitude toward Reading to students think reading is boring) were

constrained to be equal across the two groups, following suggestions by Kline (2005). This respecified model was rerun and several fit indices were calculated.

The chi-square statistic for the model was computed first. The chi-square statistic represents the discrepancy between the unrestricted sample covariance matrix and the restricted covariance matrix. The χ^2 value for the respecified measurement model was significant ($\chi^2(252)=1225.42$, $p<.001$) indicating that the observed and specified model differed. Caution is needed when interpreting this finding, however, as a significant chi-square statistic is not uncommon for samples greater than 200, and it does not necessarily reflect a poor fit to the data (Byrne, 2009).

Next, we calculated other, more meaningful fit indicators. We used the comparative fit index (CFI) and the root mean square error of approximation (RMSEA), which are less sensitive to sample size as compared to the chi-square statistic. CFI compares the lack of fit of the hypothesized model with the independent model. The CFI for the modified measurement model was .96, and the RMSEA was .02, indicating a good fit to the data (Hu & Bentler, 1999).

Structural Model Analysis

In step two of our analysis, a structural model was examined to evaluate whether early home literacy activities and parents' attitude toward reading influence student's reading self-concept and attitude, which in turn would have a direct effect on their reading comprehension. It is hypothesized that cultural differences may induce different relationships among our path model. Before analysis, the structural model was modified by constraining the (unstandardized) structural paths to be equal for both groups so as to achieve parsimony. Three structural path constraints (Parents'

Attitude toward Reading to Student's Reading Self-Concept; Student's Reading Self-Concept to Reading for Information; Student's Reading Self-Concept to Reading for Literacy) were later relaxed to have better model fit. The final structural model fit the data acceptably as evidenced by our indicators: $\chi^2(327) = 2887.90$, $p < .001$, CFI=.93, RMSEA=.03. Standardized path

coefficients for the model are displayed graphically in Figure 1. Significant levels were determined by critical ratios on unstandardized coefficients. All coefficients are significant at .001 level, except for the path Parents' Attitude toward Reading to Student's Reading Self-Concept for the Hong Kong group ($p = .22$).

Structural Model of the British and Hong Kong Groups

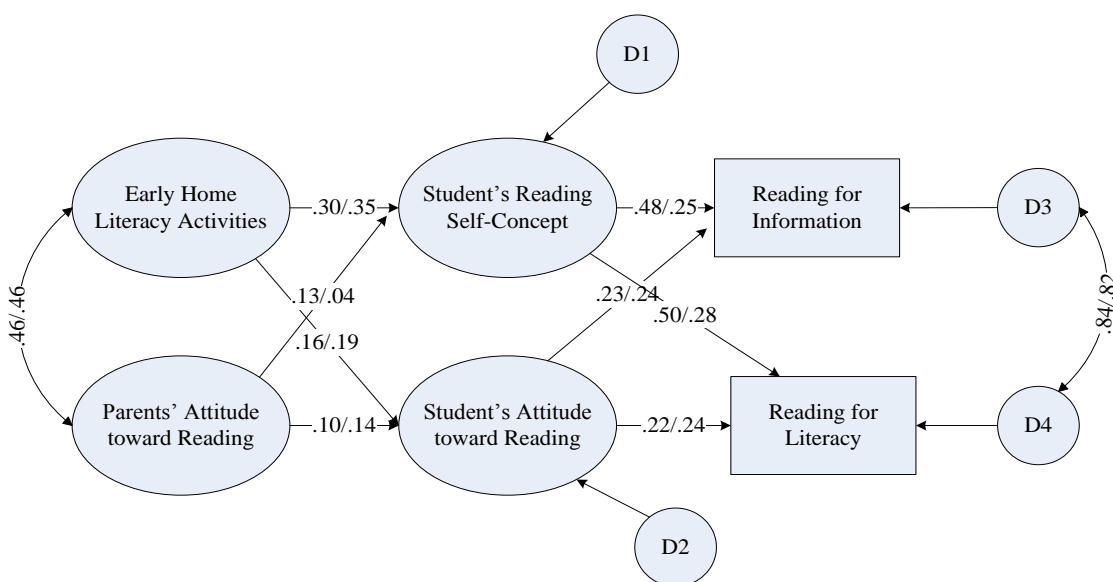


Figure 1. Coefficients for British students are before slashes. For Hong Kong students coefficients are after slashes. Significant levels were determined by critical ratios on standardized coefficients. All coefficients are significant at .001 level, except for the path Parents' Attitude toward Reading \rightarrow Student's Reading Self-Concept for the Hong Kong group ($p = .22$).

In both cultures, Early Home Literacy Activities and Parents' Attitude toward Reading were positively correlated ($r = .46$ for Britain and Hong Kong, $ps < .001$). Again, in both cultures, the Early Home Literacy Activities significantly related to Student's Reading Self-Concept ($\beta = .30$ for Britain and $\beta = .35$ for Hong Kong, $ps < .001$) and Student's Attitude toward Reading ($\beta = .16$ for Britain and $\beta = .19$ for Hong Kong, $ps < .001$). Similarly, in both cultures, Parents' Attitude toward Reading directly

related to Student's Attitude toward Reading ($\beta = .10$ for Britain and $\beta = .14$ for Hong Kong, $ps < .001$). However, Parents' Attitude toward Reading was only positively related to Student's Reading Self-Concept in British sample ($\beta = .13$, $p < .001$), but not in Hong Kong sample. Student's Reading Self-Concept directly predicted their Reading for Information ($\beta = .48$ for Britain and $\beta = .25$ for Hong Kong, $ps < .001$) and Reading for Literacy ($\beta = .50$ for Britain and $\beta = .28$ for Hong Kong, $ps < .001$). Student's Attitude

toward Reading significantly predicted their Reading for Information ($\beta = .23$ for Britain and $\beta = .24$ for Hong Kong, $ps < .001$) and Reading for Literacy purposes ($\beta = .22$ for Britain and $\beta = .24$ for Hong Kong, $ps < .001$).

Discussion

The purpose of this study was to examine the extent to which students' affective characteristics (reading self-concept and attitude toward reading) and home environments facilitate their text comprehension and the possible role of culture in this relationship. Our findings support the hypothesis that both student's reading self-concept and attitude toward reading adequately predict their comprehension of texts. And early home literacy activities have significant effects on children's development of positive reading self-concept and reading attitudes. In addition, parents' attitude toward reading directly influences their children's reading attitude. Those findings hold in the Western culture as well as the Chinese culture and are consistent with existing literature (Whitehurst & Lonigan, 2001; Yaden, Rowe, & MacGillivray, 2000). Thus, it is concluded that, regardless of culture, students' affective characteristics and home environments influence their acquisition of skills needed for comprehending a text. But, the full picture is more complex.

The effects of home environments on a student's text comprehension were moderated by student's reading self-concept and attitude toward reading. Past research indicates that home influences on children's cognitive and linguistic skills may change over time (Chall, Jacobs, & Baldwin, 1990; Rowe, 1991). As children mature and have developed more stable cognitive skills, the effect of home environments may be subsumed, but the effects of their own cognitive skills and self-judgment will

remain salient on their reading comprehension.

It is interesting to notice that Parents' Attitude toward Reading positively influenced Student's Reading Self-Concept in the Western culture ($\beta = .13$, $p < .001$) but not in the Chinese culture. In Hong Kong, the PIRLS study was conducted in Chinese which is the mother tongue for the majority of people and used in daily life. The "reading," therefore, refers to Chinese reading. For the British sample, however, the "reading" refers to English reading. Due to differences in the two language (Chinese and English) acquisitions, it could be the case that the development of reading self-concept differs as well. Our study did not delve into this further question, but our work certainly highlights a need for additional exploration here.

In addition, influences from home, as measured by standardized direct effects, are greater in the Chinese culture than in the Western culture when they exist in both cultures. The Chinese culture in Hong Kong values collectivism where family members may bring merit or demerit to the "face" of the family. Experiences from home are likely to be more crucial for Chinese children due to this awareness than for their Western counterparts

However, Student's Reading Self-Concept is a better predictor of reading proficiency in the Western culture. This may be because children in the Western culture have more exposure outside the family and other factors come into play in their development of reading self-concept. As we mentioned earlier, the acquisition of Chinese and English languages likely differ, and the differences in the effect of Student's Reading Self-Concept on reading proficiency between the two cultures may reflect the differences in language acquisitions as well.

The findings of this study provide insights on how to improve reading instructions in the Chinese and Western cultures. For example, teachers can help students establish more positive attitude toward reading and encourage them to have higher reading self-concept since students' text comprehension is directly related to those constructs. Home environments are important in the development of children's reading self-concept and attitude; therefore, the findings also have implications for offering guidance to parents.

It is noteworthy that, during our modeling, the unstandardized direct effects other than those related to Student's Reading Self-Concept were constrained to be the same across the two groups. Furthermore, three unstandardized direct effects related to Student's Reading Self-Concept (Parents' Attitude toward Reading to Student's Reading Self-Concept; Student's Reading Self-Concept to Reading for Information; Student's Reading Self-Concept to Reading for Literacy) were different across the two groups. This suggests that the difference between the British and Hong Kong samples is indeed in student's reading self-concept. With similar amount of home environments, the two groups of students could develop different amount of reading self-concept; but, with similar amount of reading self-concept, the two groups could have different levels of reading comprehension. Future investigation into why this happens could provide added insight toward our understanding of the development of affective factors related to reading comprehension in the two cultures. For example, reading self-concept may relate to other aspects of self-concept differently across cultures even when home influences are similar.

Meanwhile, there is the possibility that the items in PIRLS 2001 measure slightly different constructs in the Western

culture than in the Eastern culture. For example, the statement, "I read only if I have to," supposedly measuring attitude toward reading for both parents and children may indicate very different meanings in the Eastern and Western cultures. While it is usually considered as a negative statement indicating less self-control in the Western culture, it is likely to be thought of as a measure of "conformance" in the Chinese culture where conformity is valued more than in Britain or most other Western cultures. We call attention to these measurement differences as such sensitivity may influence explanatory interpretations.

There are several limitations of this study. First among the study limits is that we used secondary data from PIRLS 2001 and this does not allow us to include important variables that are not measured. For example, instruction at school likely influences students' reading proficiency but that variable is not available.

Second, PIRLS were administered in different languages in Hong Kong and the United Kingdom. Constructs related to reading may be different because of the different amount of effort required for language acquisition, different components of the languages, etc. Although translating instruments to different languages is common in cross-cultural studies (Tirri & Campbell, 2010), this may affect the external validity of the study.

Third, the Hong Kong sample in PIRLS 2001 was only representative of fourth graders in this special administrative region of China. Because of the century-long influences from Britain when Hong Kong was a British colony, findings from a Hong Kong population may not be generalized to other regions of China.

Future research may be conducted on exploring other factors that affect student's reading self-concept and proficiency. Those may include cognitive factors such as the

importance put on reading, behavioral measures such as time spent on reading, contextual influences such as school climate, etc. Another line of future research is to investigate whether and how home environment and student's self-concept and attitudes affect second language proficiency. For example, in Hong Kong, English is taught as a second language. Will English reading proficiency be affected by the same factors and in the similar way as Chinese reading proficiency? Will English reading proficiency be affected in a similar way in Hong Kong as in Britain? In addition, future research can focus on understanding the mechanism and process of how home environment influence children's reading self-concept, attitudes, and proficiency. For example, interviews with parents and observations in the family may be conducted to answer relevant questions.

conducted using the newer data between the British sample and the Hong Kong sample.

Endnotes

¹ Complete information on PIRLS is available at the U.S. Department of Education, National Center for Education Statistics web site <http://nces.ed.gov/Surveys/PIRLS>.

² Hong Kong was a British colony and became a special administrative region of China on July 1, 1997. The participants in the Hong Kong sample were born before that date and were assessed only four years after the sovereignty handover. The education system in Hong Kong is greatly influenced by that in the United Kingdom, although unlike the UK, the Chinese culture in Hong Kong is low on the cultural value of individualism and high on collectivism. The differences in the two samples can be more readily contributed to differences in cultures than to differences in education systems.

On the other hand, PIRLS 2006 data have been collected and made available to the public. Similar comparisons can be

References

- Akiba, M. (2008). Predictors of student fear of school violence: A comparative study of eighth graders in 33 countries. *School Effectiveness and School Improvement, 19*(1), 51-72.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411-423.
- Baker, L., Dreher, M. J., & Guthrie, J. T. (2000). Why teachers should promote reading engagement. In L. Baker, M. J. Dreher, & J. T. Guthrie (Eds.), *Engaging young readers: Promoting achievement and motivation* (pp. 1-16). New York, NY: Guilford.
- Baker, L., Scher, D., & Mackler, K. (1997). Home and family influences on motivations for reading. *Educational Psychologist, 32*(2), 69-82.
- Bradley, R. H., Corwyn, R. F., Burchinal, M., Pipes McAdoo, H., & Garcia Coll, C. (2001). The home environments of children in the United States. *Child Development, 72*(6), 1844–1886.
- Byrne, B. M. (2009). *Structural equation modeling with Amos* (2nd ed.). New York, NY: Routledge.
- Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Chao, R. K. (1996). Chinese and European American mother's beliefs about the role of parenting in children's school success. *Journal of Cross-Cultural Psychology, 27* (4), 403-423.
- Chen, C. S., Stevenson, H. W., Hayward, C., & Burgess, S. (1995). Culture and academic achievement: Ethnic and cross-national differences. In M. I. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 9, pp. 119-151). Greenwich, CT: JAI Press.
- Guthrie, J. T., & Greaney, V. (1991). Literacy acts. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 68–96). New York, NY: Longman.
- Guthrie, J. T., Wigfield, A., Metsala, J., & Cox, K. E. (1999). Predicting text comprehension and reading activity with motivational and cognitive variables. *Scientific Studies of Reading, 3*(3), 231-256.
- Hansen, H. S. (1969). The impact of the home literacy environment on reading attitude. *Elementary English, 46*(1), 17-24.
- Ho, D. (1986). Chinese patterns of socialization: A critical review. In M. Bond (Ed.), *The Psychology of the Chinese People* (pp. 1-35). Oxford, England: Oxford University Press.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.
- Hui, C. H., & Triandis, H. C. (1986). Individualism–collectivism: A study of cross-cultural researchers. *Journal of Cross Cultural Psychology, 17*(2), 225–248.
- Kush, J. C., Watkins, M. W., & Brookhart, S. M. (2005). The temporal-interactive influence of reading achievement and reading attitude. *Educational Research and Evaluation, 11*(1), 29-44.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford Press.

- MacDonald, G., & Cornwall, A. (1995). The relationship between phonological awareness and reading and spelling achievement eleven years later. *Journal of Learning Disabilities*, 28(8), 523–527.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253.
- Mau, W. (1997). Parental influences on the high school student's academic achievement: A comparison of Asian immigrants, Asian Americans, and White Americans. *Psychology in the Schools*, 34(3), 267–277.
- Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56(2), 177–196.
- Neuman, S. B. (1986). The home environment and fifth-grade students' leisure reading. *Elementary School Journal*, 86(3), 335–343.
- Noble, K. G., Farah, M. J., & McCandliss, B. M. (2006). Socioeconomic background modulates cognition–achievement relationships in reading. *Cognitive Development*, 21(3), 349–368.
- Ogbu, J. U. (1981). Origins of human competence: A cultural-ecological perspective. *Child Development*, 52(2), 413–429.
- Oyserman, D., Coon, H., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analysis. *Psychological Bulletin*, 128(1), 3–72.
- Rowe, K. J. (1991). The influence of reading activity at home on students' attitudes towards reading, classroom attentiveness and reading achievement: An application of structural equation modeling. *British Journal of Educational Psychology*, 61(1), 19–35.
- Rowe, K. J., & Rowe, K. S. (1992). The relationship between inattentiveness in the classroom and reading achievement. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31(2), 349–368.
- Snow, C. E., Burns, M. S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Stevenson, H. W., & Lee, S. Y. (1996). The academic achievement of Chinese students. In M. H. Bond (Ed.), *The handbook of Chinese psychology* (pp. 124–142). Hong Kong, China: Oxford University Press.
- Tirri, K., & Campbell, J. R. (2010). Current trends and dilemmas in cross-cultural research. In D. K. Sharpes (Ed.), *Handbook on international studies in education* (pp. 17–29). Charlotte, NC: Information Age Publishing.
- Wang, J. H., & Guthrie, J. T. (2004). Modeling the effects of intrinsic motivation, extrinsic motivation, amount of reading, and past reading achievement on text comprehension between U.S. and Chinese students. *Reading Research Quarterly*, 39(2), 162–186.
- Whitehurst, G. J., & Lonigan, C. J. (2001). Emergent literacy: Development from pre-readers. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 11–29). New York, NY: Guilford Press.
- Wigfield, A., & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology*, 89(3), 420–432.
- Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2-3), 114–128.
- Yaden, D., Rowe, D., & MacGillivray, L. (2000). Emergent literacy: A matter (polyphony) of perspectives. In M. Kamil, P. Mosenthal, P. Pearson, & R. Barr (Eds.), *Handbook of reading research* (pp. 425–454). Mahwah, NJ: Erlbaum.

A Comparative Study on Pedagogical Content Knowledge of Mathematics Teachers in China and the United States

Xiaobo She, Texas Tech University
William Y. Lan, Texas Tech University
Jennifer Wilhlem, University of Kentucky

Abstract

Using an interview protocol comprised of algebraic questions, U.S. and Chinese middle level teachers' responses were analyzed for differences in pedagogical content knowledge. Four themes were extracted from the analysis: concrete models and simple steps; practical and theoretical approaches; the application of cross-multiplication; generalization of problem and solution types. Our findings revealed U.S. teachers were more likely to use concrete models and practical approaches in problem-solving and promoting students' knowledge skills. However, they seemed to lack deep understanding of mathematical concepts as well as interconnections between concepts. Chinese teachers were inclined to utilize theories and procedures in teaching. By generalizing rules and strategies, Chinese teachers tended to integrate conceptual knowledge points as a conceptual network that made the knowledge applicable within multiple situations.

To meet the goal of competitiveness in mathematics education globally, international comparison studies (ICS) have received increasing attention for the purpose of sharing, discussing, and debating important issues across countries (Robitaille & Travers, 1992). Mathematics education in the United States has benefitted from the findings of ICS that result in initiatives towards improving students' mathematical performance. For example, the report of

the Trends in International Mathematics and Science Study (TIMSS) has shown that both fourth and eighth graders' mathematics scores in the United States have made a significant increase in their 2007 averages when compared to their 1995 scores over the 12-year period (NCES, 2007). Although it is difficult to pinpoint the reasons for the improvement, it is plausible to give credit to focused, cross-nation comparisons, which resulted in educational policy changes, curriculum modification, and the development of teaching practices.

Meanwhile, the international comparative studies trigger our interest in delving into the underlying causes for performance differences between the United States and top-performing countries and districts, such as China-Taipei, Korea, Japan, and Hong Kong (NCES, 1999, 2003, 2007). A variety of research has been processed to identify the essential factors contributing to students' mathematical performance.

Teachers, as one of the most significant factors in mathematics education, do not only influence students on their content knowledge, but also play a critical role in shaping their misconceptions and confusions. However, large-scale investigation in terms of how teachers impact students' academic achievement is still sparse due to the difficulty of implementing among teachers extensively. Hence, small, in-depth studies become especially important and practical in examining the effects of teachers on students'

mathematics learning. In this study, we attempted to use small-group comparisons to examine the characteristics of mathematics teachers' content knowledge and their teaching strategies as two essential components of pedagogical content knowledge (PCK) when teaching algebra. We hope the study will enrich our understanding of teachers' PCK and its impact on the effectiveness of mathematics education.

Theoretical Framework

Although a series of continual large-scale studies have shown U.S. students making significant progress in the international mathematics tests over the last decade (NCES, 2007), the status quo of U.S. students' continuous underperformance as opposed to their eastern Asian peers draws attention and speculation. Current studies mainly focus on identifying the distinctions for students' achievement within various content and competence domains between the top-performing Asian countries and the United States. Specially, Chinese students display superiority over their U.S. peers on base-ten counting (Miller & Stigler, 1987), computation and mental mathematics (Cai, 1997; Geary, Bow-Thomas, Fan, & Sigler, 1993), simple problem solving (Cai, 1995), and representational competence (Brenner, Herman, Ho, & Zimmer, 1999). A number of factors have been suggested as potential contributors for the divergence in specific mathematical areas, such as students' beliefs, attitudes, motivations of mathematics learning, teachers' instructional strategies, and focus of school curriculum, etc. (Chen & Stevenson, 1995; Wang & Lin, 2005).

Undoubtedly, teachers are

considered as one of the most significant factors that affect student learning in mathematics due to the critical role teachers playing in the teaching and learning process. According to the National Council of Teachers of Mathematics (2000), "Effective teaching requires knowing and understanding mathematics, students as learners, and pedagogical strategies" (p. 17). Teachers' mathematics knowledge is essential to effective teaching and student learning (Ball & Bass, 2001; Shulman, 1987). To teach effectively, teachers must possess the knowledge and skills that consists of (a) general ways to present content to students; (b) understanding of students' common conceptions, misconceptions, and difficulties when encountering particular situations; and (c) specific teaching strategies that can be used to meet students' diverse learning needs, which derives from Shulman's original notion of *pedagogical content knowledge (PCK)* (Rowan, Schilling, Ball, & Miller, 2001; Shulman, 1987). A wealth of studies has elaborated on the definition of PCK in accordance with particular attributes and needs within diverse disciplines. For instance, in the domain of mathematics education, PCK has been defined as involving three components: knowledge of content, knowledge of curriculum, and knowledge of teaching as shown in figure 1 (An, Kulm, & Wu, 2004). These three components interconnect and interact intensely in a complex way. Accordingly, profound understanding of mathematical content is likely to reciprocally interact with the effectiveness of teaching (An et al., 2004).

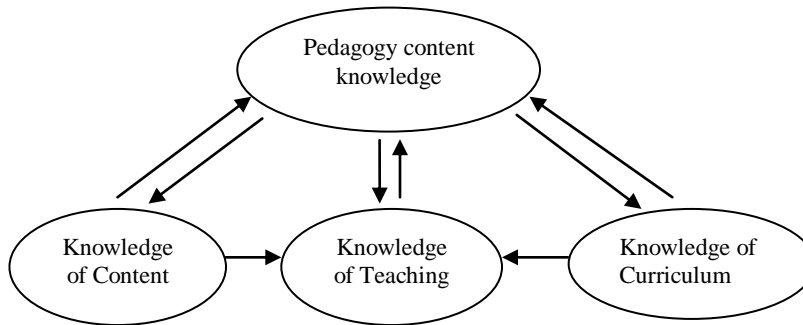


Figure 1. The network of pedagogical content knowledge is adapted from An, Kulm, & Wu, 2004.

Furthermore, effectiveness of teaching is also affected by teachers' epistemological belief on the learning process, that is, learning as knowing or learning as understanding (An et al., 2004). Teachers who hold the belief of "learning as knowing" usually focus on infusing students with specific concepts or procedures without identifying their misconceptions and understanding level. In contrast, teachers with the "learning as understanding" belief tend to encourage students to internalize their newly acquired concepts through the process of integrating prior and current knowledge as a netted whole. As a result, academic outcomes under the epistemological belief of learning as knowing are classified at a surface level while those under the belief of learning as understanding at a mastery level.

Numerous research has been carried out to compare Chinese and American teachers' knowledge of content (e.g., Ma, 1999) and knowledge of effective teaching (e.g., An et al., 2004), but little research examines teachers' pedagogical content knowledge by integrating these two crucial components of PCK. Our current study focused on investigating U.S. and Chinese teachers' PCK differences by comparing their problem-solving

strategies and teaching methods in well-chosen algebraic areas. Our research question was: *What are the differences in pedagogical content knowledge between Chinese and American teachers when observing their problem-solving processes in specific algebraic areas?*

Methodology

Participants of the study were four teachers from the west Texas area of the United States and four teachers from one school in a large city of the Sichuan province in southwest China. In addition, the teachers were teaching at the same middle school level for the same subject area of mathematics in schools with similar enrollment of 500 students. However, there were demographic differences between these two school districts. The U.S. teachers were from a school district with ethnic composition of 49.7% Hispanic, 35.4% Caucasian, and 14.9% African American. On the contrary, all Chinese teachers were from one district with dominantly more than 95% students from one ethnic group – Han.

A set of eight algebraic word problem questions was given to participating teachers to solve within an interview session. All the questions were developed with the purpose of

comparing and contrasting the use of different problem solving solutions and teaching strategies between teachers from two countries on topics of symbols, surface area, proportional reasoning, and patterns (see Figure 2). The questionnaire was prepared identically in both English and Chinese versions.

Teacher participants were encouraged to “think aloud,” a research method that reveals people’s thinking process through language, when attempting the problems. They were also asked to illustrate their teaching strategies if the questions were used in their mathematics courses. Each participant was interviewed and videotaped individually. In addition, their dialogues to themselves as well as to the interviewers during the problem solving processes were transcribed to analyze for similarities and discrepancies in teachers’ content knowledge and teaching strategy aspects of PCK. The lead author conducted all interviews with the Chinese teachers and translated them into English. The second author and two colleagues in the mathematics education program conducted interviews with the U.S. teachers. Each interview lasted approximately 20 to 30 minutes.

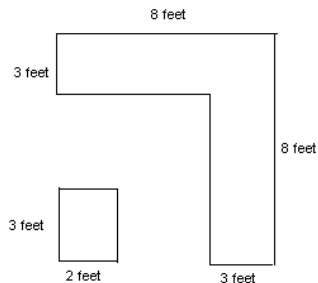
Incorporating an interview method of data collection assists researchers to better understand teachers’ knowledge of content and teaching strategies. The qualitative

approach tends to be more insightful than the frequently utilized quantitative approach, as indicated in the Board on International Comparative Studies in Education:

There is a great need for small, in-depth studies of local situations that permit cross-cultural comparisons capable of identifying the myriad of causal variables that are not recognized in large-scale surveys...much survey data would remain difficult to interpret and explain without the deep understanding of society that other kinds of studies provide...research in cross-national contexts benefits from increased documentation of related contextual information, it would be useful to combine large-scale surveys and qualitative methods. (Gilford, 1999, p. 22)

Data collected from the interviews were coded using norms modified from An et al.’s (2004) coding categories (see Table 1). By using constant comparative data analyses (Glaser & Strauss, 1967), all categories were grouped and thereby formed four themes within the domains of content knowledge and knowledge of teaching strategies.

1. On Friday the low temperature in Nome, Alaska, was -6°F , and the high temperature was 14°F . How much warmer was the high temperature than the low temperature?
2. Mr. Jones wants to install new countertops on his two kitchen counters. The drawing below shows the dimensions of the counters. What is the least amount of material needed to cover the tops of both kitchen counters?



3. A software company employs 450 workers. It plans to increase its workforce by eight employees per month until it has doubled in size. Write an equation that can be used to determine m , the number of months it will take for the company's workforce to double in size and solve this equation.
4. Larry's favorite painting has a width of 30 inches and a height of 24 inches. Larry had a reduced copy of the painting made as a gift for his father. If the reduced picture of the painting was similar to the original painting and the height of the reduced picture was seven inches, what was the width of the reduced picture?
5. Raymond packs boxes for an appliance company. He can pack a large box in 10 minutes and a small box in four minutes. He needs to pack 10 large boxes and 20 small boxes. If 2.5 hours remain before closing time, will Raymond have time to finish the work before closing time if he works without stopping?
6. The table below shows a relationship between x and y . Write an equation that best represents this relationship?

x	y
0	3
1	8
3	18
4	23
6	33

7. Sharon played an electronic game. There were 15 questions, of which she answered three incorrectly. At this rate, how many questions should Sharon expect to answer in correctly if she answers a total of 135 questions?
8. To make a certain shade of orange paint, Calvin must add 20 ounces of yellow paint to every 50 ounces of red paint. If he uses 200 ounces of red paint, what is the number of ounces of yellow paint he should add to get the shade of orange he wants?

Figure 2. Algebraic Ideas Assessment

TABLE 1

Categories for describing teachers' problem-solving and teaching strategies in dealing with Algebraic Ideas Assessment (adapted and modified from An et. al., 2004)

Category	Brief definition
1.	Prior knowledge: Know students' prior knowledge and connect it to new knowledge.
2.	Concept or definition: Use concept or definition to promote understanding.
3.	Rule and procedure: Focus on rule and procedure to reinforce the knowledge.
4.	Draw picture or table: Use picture or table to show a mathematical idea.
5.	Give example: Address a mathematical idea through examples.
6.	Provide students an opportunity to think and respond: Promote students to think problems and give them chances to answer questions.
7.	Manipulative activity: Provide hands-on activities for students to learn mathematics.
8.	Attempts to address students' misconceptions: Identify students' misconceptions.
9.	Use questions or tasks to help students' progress in their ideas: Pose questions or provide activities to increase the level of understanding for students.
10.	Provide activities and examples that focus on student thinking: Create activities and examples that encourage students to ponder questions.
11.	Use one representation to illustrate concepts: Apply repeated addition to address the meaning of fraction multiplication, or use area to address the geometrical meaning of fraction multiplication.
12.	Using more than one representation to illustrate fraction multiplication: Apply both repeated addition and area to address the meaning of fraction multiplication.
13.	Unintelligible response: Provide response that is not relevant to the question.
14.	Incorrect: Provide a wrong answer.

Results

By comparing the teachers' problem-solving strategies and their self-revealing cognitive processes via the "thinking aloud" technique when solving

the eight algebra word problems, we found that both the U.S. and Chinese teachers had extensive content knowledge backgrounds regarding algebraic topics of proportion, rate,

equation with variables, and linear functions. In addition, both groups were equipped with essential skills to teach students the content via various approaches. However, major differences were found in their content knowledge as well as knowledge of teaching strategies between the U.S. teachers and their Chinese counterparts. Four themes were extracted from the differences within the two components of PCK. For Content Knowledge, differences between U.S. and Chinese mathematics teachers were manifested in three themes of *concrete models and simple steps*, *practical and theoretical approaches*, and *application of cross-multiplication*; For knowledge of Effective Teaching,

differences between the two groups of teachers can be found in the theme of *generalization of problem and solution types*.

Content Knowledge

Theme 1 - Concrete models and steps. When attempting question 1, finding the difference between two temperatures with opposite signs, all four U.S. teachers drew a number line in order to make the difference between the temperatures visual whereas all Chinese teachers preferred to use a simple calculation of subtracting the lower temperature from the higher temperature (see Figure 3).

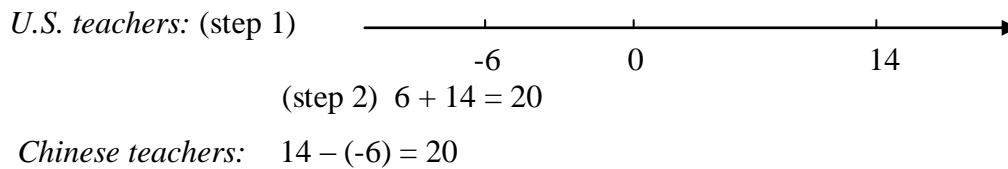


Figure 3. Different solutions by American and Chinese teachers

The U.S. teachers rationalized that using a visualized representation made problem solving easy for students since all they needed to do was to find the distance between two points on the number line. Using concrete or graphic representations has been supported by NCTM (2000) due to its potential to develop meaningful understanding of mathematical concepts. With the number line, the U.S. teachers tended to find the distance from -6°F to 0°F (6 degrees) and the distance from 0°F to 14°F (14 degrees), then add the two distances to get the correct answer of 20 degrees (see Figure 3).

In contrast, Chinese teachers tended to solve this problem by using simple computation steps instead of drawing any explicit graphics. In the

Chinese teacher group, one participant explained her reason of discouraging drawing since generating a graphical representation was time-consuming and distracting. The other two explained that students were allowed to use number lines only at the beginning of learning negative numbers, which helped visualize the relationships between numbers with opposite signs. Once the students were familiar with algorithms of integers, they were no longer encouraged to use number lines and teachers expected them to be able to visualize the number line in their mind rather than on the paper. In other words, Chinese teachers were likely to impose higher expectations on students to develop mental visualization skills in middle school. In addition, Chinese

teachers expressed that students' accumulative exercises in mental visualization would benefit their abilities with spatial visualization and abstract thinking in their future coursework of advanced algebra and geometry.

By comparing their solutions, evidence was found that American teachers were more likely to relate problems to concrete situations for the purpose of visualizing the relationship between numbers whereas Chinese teachers tended to identify the mathematical relationships by looking for key words. For instance, in question 1, all Chinese teachers came to the same solution of subtracting the lowest degree (-6°F) from the highest degree (14°F). In their minds, the word *warm* in the problem implied the solution as a simple subtraction of the smaller number from the larger number. Their problem-solving method was encased with simplicity and efficiency, which was consistent with their teaching strategy. As An et al. (2004) asserted, "the main characteristic of Chinese mathematics is the development and practice of *accurate* and *efficient* means of computation and to apply these in real life" (p. 160).

Theme 2 - Practical and theoretical approaches. Participants' responses indicated that the U.S. teachers were more likely to solve problems with practical and specific approaches whereas Chinese teachers tended to utilize theories or generalized strategies. For instance, given question 6 asking to develop a linear equation based on data of two variables provided in a table, three out of four American teachers adopted the strategy of a "guess and plug in" brute force method. In contrast, Chinese teachers demonstrated varied approaches rather than simply

testing all the numbers in possible relationships. For example, after reading the problem, a Chinese teacher instantly laid out all algebraic possibilities of equations satisfying positive relationships between variables x and y . He also explained pertinent attributes for each equation form in order to identify the appropriate pattern between variables for this question.

There are several types of functions fit for this kind of relationship: the linear function, the proportional function, the quadratic function, the exponential function, and the logarithmic function. Besides the increasing relationship, we should look at the relationships of the variable changes too. Viewing from the angle of parameters, (*this question*) when x increases 1 unit, y increases 5 units; when x increases 2 units, y increases 10 units. This phenomenon tells us that x and y are proportional in relationship and they should be expressed as a linear function...If the growth becomes faster and faster, it is possible to be an exponential function; if the growth gets more and more slowly, it's likely to be a logarithmic function. (Chinese teacher 1)

This reasoning-and-proving method had been adopted primarily by Chinese teachers in practice since it not only ensures the accuracy of problem solving, but also broadens students' mathematical knowledge into an integral structure. Through analyzing diverse possibilities and then sifting to the correct answer, the teacher simplified a complicated problem to a multiple-choice type of problem, that is, to select

a correct response among several options. Nevertheless, this teaching strategy demands a solid understanding of underlying theories of numerous concepts and procedures, sometimes far beyond what is required to solve the problem at hand.

Besides eliminating incorrect options with their extensive knowledge in mathematics, Chinese teachers were inclined to use strategies of generalization. To solve the same question, Chinese participant 2 generalized two solutions by using a linear function and coordinate systems.

Like I just wrote, the tendency of changes in x's is: the second x value is one more than the first one; and the third x value is two more than the second one. It can be found that with the change of x's, the value of y changes correspondingly, that is, with every unit of increase of x is a constant amount of change in y. Therefore, students may conjecture that this is a linear function. Another way we teach students to solve this problem is to use a coordinate system. We usually teach these two methods. (Chinese teacher 2)

In contrast, three out of four U.S. teachers did not show evidence that they had any strategy to employ other than "guess and plug in." Below is a typical dialogue of the way a U.S. teacher explained how the problem should be solved.

They would look to see if there is a pattern in the differences....One-five, two-ten, one-five, two-ten (*points to left column for the first number and right column for second*)....When x is 0, y is 3. So

$x + y = 3$ and that might be one check that they (students) do, which is 3 (writes equation). But $1 + 8$ is 9, so they would back away from that and look for another relationship to what is going on here. (U.S. teacher 4)

U.S. participant 4's explanation showed that she did not see the pattern for the changes of x and y based on the information from the table. Although she found that x and y intervals were aligned with a pattern of "one-five, two-ten, one-five, two tens," she still could not figure out a constant relationship between them as 'x times 5 plus 3 equals to y.' After the interviewer hinted about drawing a line, she finally figured out a linear relationship between x and y. This problem-solving episode implied that U.S. teachers may have not possessed a conceptual understanding of linear functions. Such difficulties experienced by the U.S. teachers would most likely impact their students' learning, interest, and achievement in mathematics.

Theme 3 - The application of cross-multiplication. For those questions involving proportional relationship (Questions 4, 7, and 8), both U.S. and Chinese teachers adopted various approaches when setting up proportional relationships and using a cross-multiplication strategy in solving problems. U.S. teachers manifested divergent attitudes toward using cross multiplication in solving proportion problems. For example, when asked about the application of cross-multiplication, two U.S. teachers expressed frustration and confusion. One teacher stressed that cross-multiplication never made sense to students since it was merely taught as a procedure or as a "short-cut" without helping students to understand why it works. Another

teacher could not interpret cross-multiplication correctly regarding why the method works. Nevertheless, the other two U.S. teachers embraced the use of cross-multiplication while they emphasized how this strategy should be taught to facilitate student understanding of the method. See excerpt below.

Initially when they're being taught this method, they know an equation is about balance between the sides of an equation. And in order to get rid of this denominator you have to multiply both sides by 50 (*denominator*)...eventually they would shorten it because they see you get rid of it (*denominator*) here...I teach them to understand the process....(U.S. teacher 4)

Reversely, Chinese teachers used cross-multiplication (without doubt or hesitation) as if it was one of the most basic algorithmic rules. Due to the emphasis placed on efficiency and accuracy in the Chinese examination system, cross-multiplication has been viewed not only as a procedure but a basic concept needed to be grasped. Generally, Chinese teachers value both conceptual understanding and procedural proficiency equally. Particularly, they believed that procedural proficiency is built upon students' solid understanding of mathematics, which also positively impacts the extent of knowledge mastery inversely. Our finding was consistent with a previous study (An et al., 2004), which reported that far fewer U.S. teachers believed that using procedures and rules were effective in building mathematical ideas than their Chinese peers.

Knowledge of Effective Teaching

Theme 4 - Generalization of problem and solution types. With regard to teaching methods, Chinese teachers were more likely to generalize a problem in terms of different situations and then identified the appropriate solution in response. For example, to solve problems involving decimals (question 3), a Chinese teacher summarized three different types of solutions to round a decimal into an integer and asserted that students should learn to read the problem carefully and find out the key words to match the corresponding situation. He said,

Three methods for it: rounding-up, taking-out, and adding-one methods. This problem needs to use "adding-one" method. Since the result is 56.2, we cannot keep it as 56 months, (*because*) not enough, we should "add one" to 57 months. Another dimension of the meaning in this question is that varied conditions determine dissimilar solutions. (Chinese teacher 1)

By generalizing problems, Chinese teachers tried to reinforce the belief of mathematics as a netted concept web within a learner's mind. In this way, students are able to connect new concepts to their prior knowledge spontaneously as well as to search for solutions from varied angles. For instance, when dealing with the pattern problem (question 6), Chinese teacher 2 described two distinct solutions, "because of the interval as y , we can solve this problem with two ways: the first, visual observation... the second way (is) that students probably think about is the application of function ..."

In response to question 4

involving two similar figures, another Chinese teacher not only reviewed a variety of ratios, but also illustrated their relationships. He stated,

Students need to make sense of one principle: there are different ratios for two similar figures, such as similar ratio, area ratio, and volume ratio. In terms of two specific shapes, the area ratio equals to the square of the similar ratio; the volume ratio equals to the cube of the similar ratio....therefore, identifying what belongs to a similar ratio is crucial, such as the ratio of height, length, median line, perimeter, and so on. All the mentioned ratios are called similar ratio because they are units of length. (Chinese teacher 4)

In contrast, U.S. teachers were more likely to use one method in solving the particular problem scenario without illustrating any larger holistic “big idea” or relating to parallel situations. For example, when solving the same problem 4, a U.S. teacher said,

The height of reduced painting was 7 inches, so, what was the width of reduced picture? OK, whenever you have similar figures...you have similar figures the dimensions, or let's see the scale factor. I guess you can set up proportion. That is what I am trying to say. It is 30 inches as the width (writing), and the height is 24 inches in your original. And the smaller one, the height is 7 inches. We are looking for the width of the reduced picture. You can set up a proportion, you can see 30 inches is to 24 inches, and unknown

width, is to 7 inches. And shortcut is to cross-multiple. (U.S. teacher 2)

Apart from divergent teaching strategies in problem solving depicted above, U.S. and Chinese teachers showed different attitudes and perspectives in teaching it. As mentioned previously in theme 2 of *Content Knowledge*, some U.S. teachers were inclined to use and teach it just as a “short-cut” for the only purpose of simplifying the problem-solving processes in questions. On the contrary, all the Chinese teachers espoused the necessity of procedural practices; they believed that developing students' proficient procedural skills helps to reinforce what they had learned and allowed them to transfer skills easily to new knowledge to novel problem situations (An et al., 2004).

Discussion and Conclusion

Because of the "gatekeeper" status of algebra to advanced mathematics study as well as its significance penetrating K-12 curriculum (Davis, 1985; Oliver, Izsak, & Blanton, 2002), this study used a set of algebraic questions to compare and contrast middle school teachers' relevant content knowledge and knowledge of teaching as two essential components of pedagogical content knowledge. Particularly, these teachers' problem-solving skills and teaching strategies were carefully examined.

Findings in this study may identify the factors that contribute to the discrepancy in mathematics achievement between American and Chinese students that caused by teacher impact. Teachers' pedagogical content knowledge is essential for effective teaching which directly affects students' learning

outcomes. In this study, little evidence revealed an obvious discrepancy in teachers' content knowledge in algebra. Nonetheless, significant differences were noticed in their problem solving strategies and teaching methods. Chinese teachers preferred to tackle problems by looking for key words in order to set up direct relationships while U.S. teachers considered drawing visual representations to be helpful when solving problems. Furthermore, Chinese teachers were likely to sort problems into categories based on wording structure and seek diverse approaches to deal with whereas American teachers were in favor of taking a practical, brute force approach, such as 'trial and error'.

All these phenomena mirrored disparate beliefs and values in two education systems: Chinese teachers treat *accuracy* and *efficiency* as the primary goal for solving problems as well as teaching mathematics while only *accuracy* is emphasized in U.S. (NCTM, 2000). Accordingly, forming a conceptual understanding within a netted knowledge structure becomes the prerequisite to reach this goal. In particular, Chinese teachers were accustomed to teaching students by constantly linking mathematical concepts, which allowed students to review and reinforce concepts and procedures from time to time. Consequently, students are likely convinced that mathematics is a well-structured body of knowledge. In contrast, U.S. teachers heavily relied on practical approaches and external aids, such as graphic representation and visual manipulatives. Despite its advocacy of various forms of representations, NCTM (2000) warned that such representations and manipulatives sometimes are used as if they end in themselves. In other

words, if graphic representations are not used to reach in-depth understanding, students are likely to end up with "learning as knowing" instead of "learning as understanding."

The results of this study support an idea that teaching for understanding is the key for successful math education. Procedural learning can become valuable only when it is based on students' understanding the underlying mathematical logic and reasoning. Chinese teachers value procedural proficiency as an equal weight of conceptual understanding. In their perspectives, procedural proficiency not only results from a genuine mastery of knowledge but also resonates conceptual understanding to some degree. In contrast, U.S. teachers have more ambivalent attitudes toward procedural learning, such as the attitudes expressed by U.S. participants in use of cross multiplication. Some U.S. teachers opposed cross multiplication as conceptual learning since it is merely a shortcut in calculation. In fact, if students have conceptual understanding of fractions, they are capable to deduce why this shortcut works all the time. Using a shortcut should be built on the foundation of understanding.

Evidently, a series of schooling and non-schooling factors should be taken into consideration when exploring teachers' impact on students. For instance, researchers have found that Asian American students are good and even better performers in mathematics when they were exposed to the same curriculum and teaching practices along with other American ethnic peers, which suggests that different educational systems between east and west cannot exclusively explain the significant academic gap internationally. Therefore,

non-schooling factors, such as differences in culture and language, should also be taken into consideration in terms of their effects on teaching-and-learning process. There is no doubt that teaching is a cultural activity. But in what form and to what extent culture influences the teaching practice as well as students' attitude, motivation, and performance are worthy of intensive investigation across countries.

In conclusion, the results of this study suggest that remarkable differences exist in the pedagogical content knowledge between U.S. and Chinese teachers, which may result in dissimilar teaching and learning outcomes. U.S. teachers were more likely to use concrete models and practical approaches in problem-solving and promote students' knowledge skills.

However, they seemed to lack deep understanding of underlying mathematical theories. The Chinese teachers were inclined to utilize theories and procedures for teaching and learning. By generalizing rules and strategies, they were able to integrate knowledge points as a whole network. As An et al. (2004) indicated, different education belief systems produce different attributes of pedagogical content knowledge. In order to really improve teaching, we should invest far more than we do now in generating and sharing knowledge about teaching (NCES, 2007).

Given the fact that this study included a small sample size, its findings cannot necessarily be generalized to all mathematics teachers in the United States and China. Especially, since this study heavily relied on qualitative methods, the findings may not be applied to the population of middle school mathematics teachers in these

two countries. However, these findings do provide some insights on teachers' pedagogical content knowledge through the lens of an international comparative study.

In a future study, we will examine how cultural beliefs influence teachers' pedagogical content knowledge. To improve the levels of generalization and application, a mixed methods approach will be utilized in which both quantitative and qualitative data will be collected and analyzed.

References

- An, S., Kulm, G., & Wu, Z. (2004). The pedagogical content knowledge of middle school, mathematics teachers in China and the U.S. *Journal of Mathematics Teacher Education*, 7, 145-172.
- Ball, D. L., & Bass, H. (2001). Interviewing content and pedagogy in teaching and learning to teach: Knowing and using mathematics. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 83-104). Westport, CT: Ablex Publishing.
- Brenner, M. E., Herman, S., Ho, H. Z., & Ximmer, J. M. (1999). Cross-national comparison of representational competence. *Journal for Research in Mathematics Education*, 30(5), 541-557.
- Cai, J. (1995). *Cognitive analysis of U.S. and Chinese students' mathematical performance on tasks involving computation, simple problem solving, and complex problem solving* (Monograph 7, *Journal for Research in Mathematics Education*). Reston, VA: National Council of Teachers of Mathematics.
- Cai, J. (1997). Beyond computation and correctness: Contributions of open-ended tasks in examining U.S. and Chinese students' mathematical performance. *Educational Measurement: Issues and Practice*, 16(1), 5-11.
- Chen, C., & Stevenson, H. W. (1995). Motivation and mathematics achievement: A comparative study of Asian-American, Caucasian-American, and East-Asian high school students. *Child Development*, 66, 1215-1234.
- Davis, R. B. (1985). ICME 5 report: Algebraic thinking in the early grades. *Journal of Mathematical Behavior*, 4, 195-208.
- Geary, D. C., Bow-Thomas, C. C., Fan, L., & Siegler, R. S. (1993). Even before formal instruction, Chinese children outperform American children in mental addition. *Cognitive Development*, 8(4), 517-529.
- Gilford, D. M., (Ed.). (1993). *A collaborative agenda for improving international comparative studies in education*. Washington, DC: National Research Council, National Academy Press.
- Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York, NY: Aldine.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Child Development*, 2, 279-305.
- National Council of Teachers of Mathematics (NCTM, 2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Center for Education Statistics (NCES, 1999). *Trends in international mathematics and science study, 1999* [Data file]. Washington, DC: Office of Educational Research and Improvement. Available from National Center for Education Statistics Website, <http://nces.ed.gov/timss>
- National Center for Education Statistics. (NCES, 2003). *Trends in international mathematics and science study, 2003* [Data file]. Washington, DC: Office of Educational Research and Improvement. Available from National Center for Education Statistics Website,

- <http://nces.ed.gov/timss>
- National Center for Education Statistics. (NCES, 2007). *Trends in international mathematics and science study, 2007* [Data file]. Washington, DC: Office of Educational Research and Improvement. Available from National Center for Education Statistics Website, <http://nces.ed.gov/timss>
- Olive, J., Izsak, A., & Blanton, M. (2002). Investigating and enhancing the development of algebraic reasoning in the early grades (K-8): The early algebra working group. In D.S. Mewborn, P. Sztajn, D. Y. While, H. G. Wiegel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the twenty-fourth annual meeting of the International Group for the psychology of mathematics education* (Vol. 1, pp. 119-120). Columbus, OH: ERIC.
- Robitaille, D. F., & Travers, K. J. (1992). International studies of achievement in mathematics. In D.A. Grouws (Ed.), *Handbook of mathematics teaching and learning* (pp. 687-709). New York: Macmillan Publishing Company.
- Rowan, B., Schilling, S., Ball, D., & Miller, R. (2001). *Measuring teachers' pedagogical content knowledge in surveys: An exploratory study*. Retrieved from <http://www.sii.soe.umich.edu/documents/pck%20final%20report%20revised%20BR100901.pdf>
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1-22.
- Wang, J., & Lin, E. (2005). Comparative studies on U.S. and Chinese mathematics learning and the implications for standards-based mathematics teaching reform. *Educational Researcher*, 34(5), 3-13.

Faculty Roles in University Governance at Two Universities in the United States and China

Hongchi Jiang, China University of Geosciences

Michael Wei, University of Missouri

Abstract

University governance is a core issue of higher education, and faculty has a critical role in that governance. Based on structured interviews, 20 faculty members (10 each from a Midwest university in the United States and a central university of China) talked about their university's governance structures and their roles in governance. It revealed that the governing structures in the American university are drastically different from those in the Chinese university. Although both universities' faculty members thought they have more or less impact on the governance of curriculum, faculty governance, student governance, and organizational governance, the actual practice, however, is different at the American university and the Chinese university. The Chinese university's faculty members had less of a say in final decisions and their roles were often impinged upon by their administrators. However, all of the interviewees at both universities felt that faculty should participate in university governance and seek more involvement in final decisions. This paper concludes with suggestions on effective governance for both universities.

Introduction

Governance is a political term that refers to the administrative ways of government. Broadly speaking, university governance in higher education refers to the means by which higher educational institutions are formally organized and managed. It is related to the structure and process of authoritative decision-making issues that are significant for external as well as internal stakeholders within a

university (Gayle, Tewarie, & White, 2003). University governance is a core issue of higher education. It concerns the existence, characteristics, and development of universities.

Simply stated, university governance is the mechanism by which universities are operated. The concept of university governance predominantly refers to the internal structure, organization, and management of autonomous institutions. A suitable and meaningful role for faculty in this critical aspect of governance can improve the efficiency of university governance. Since the educational systems of China and the United States are very different, in this case study, we intended to compare two universities, one university in China with a pseudo-name of Huaxin University where the first author teaches, and one university in the United States with a pseudo-name of Clouden University where the second author teaches. Both universities are located in the central part of its respective country. The scope of investigation was the extent that one university differs from the other in terms of university governance. University governance in this paper mainly refers to university internal governance, especially authoritative rights in decision-making.

Literature Review

The issues related to university governance have been hotly debated in past years. Upon reviewing the literature, we narrowed our topics centering on four areas: definition of university governance, governing structure, faculty governance, and academic decision-making rights.

Firstly, there are different definitions for university governance. The American Association of University Professors (AAUP) was the first organization that articulated the importance of faculty rights in university governance. It was concerned about personnel decisions, selection of administrators, preparation of the budget, and determination of educational policies (AAUP, 1940). This statement, correlating with the Yale Report of 1828, was the “first attempt at a formally stated philosophy of education” for universities, emphasized at that time (Brubacher, 1982, p. 5). Recently, Kezar and Eckel (2004) defined university governance at the macro-level of policy decision-making. They suggested that governance should be a multi-level concept including several different bodies and processes with different decision-making functions. In this sense, governance is sometimes defined as the internal management of institutions. Arimoto (2001) argued that in any cases, governance is an extensive group of stakeholders seeking to influence university rules and policies. These stakeholders include higher education associations, funding organizations, the Department of Education, related congressional committees, accrediting institutions, system-level offices, governors, state departments or boards of education, state legislators, students, alumni, local community members, trustees, senior administrators, faculty leaders, and university presidents.

Secondly, governing structures for higher education are highly differed throughout the world. Therefore, different authors studied university governance and governing structures from different angles. For example, Altbach (2005) noted that the different models for higher education throughout the world do share a common heritage. Ehara (1998) compared Japanese and U.S. faculty perceptions of university

governance in one of his early articles. He suggested that the difference in university governance depends on who holds decision-making authority. He further suggested that in a centralized system, all executive decisions would be made by the administration, while, on the other hand, in a decentralized system the faculty would control decision-making.

Han (1993) and Zhou (1989) described the special governing structures at Chinese universities. Han (1993) noted that the Secretary of the University Party Committee regards himself/herself both as an organizer and an educator in the university. The Secretary of the University Party Committee leads a Standing Party Committee. The members of Standing Party Committee hold the key leadership posts in administrative organs of the university. The task of the Standing Party Committee at each university is to ensure that the university follows the Party Committee’s guidelines, and to take responsibilities of political education to university administrators, teachers, and students. Zhou (1989) also observed that the role of the Party Committee in his university was defined as one to support and monitor the university president and the administration in implementing the State’s and the Party’s policies. In this way, the Party Committee Secretary is empowered with overall governance, which is much stronger than that of the university president. The Party Committee’s political supervision is implemented in the university through various Party branches and sub-branches at all levels of the university. Virtually for every academic administrative head, there is a parallel Party Secretary.

In addition to governing structures, Gayle et al. (2003) discuss approaches to effective leadership and strategic management in the 21st century universities. They take an objective look at traditional forms of shared governance and recent

attempts in incorporating them into the university governance system. The major question posed is whether either approach is appropriate to meet the needs of the modern American college and university and the challenges of today's environment. The fundamental premise in the analysis is that governing structures should be evaluated from the perspective of their contribution to teaching and learning, leadership and management, use of technology, and budgeting.

Thirdly, faculty governance is a key issue of faculty role in university governance. The 1970 AAUP survey of faculty governance used by McCormick and Meiners (1988) provided measures of faculty participation in decision-making. They found that faculty control ranged from as high as of 96.58% for decisions concerning academic performance and as low as of 7.02% for decisions concerning long term budgetary planning. On average, faculty members played a greater role in decisions concerning curriculum and faculty governance. Faculty members had less control over decisions involving organizational management, the choice of organizational leaders, and budgetary planning.

Lastly, some scholars studied the decision-making rights of faculty members in academic issues. Benjamin and Carroll (1996) and Brown (2001) studied the areas of university governance where faculty members can play important roles. Benjamin and Carroll found that faculty members remain the critical lynchpin in these areas: (1) deciding what curriculum is taught; (2) choosing the pedagogy to be used; (3) determining what departments and fields to be kept and their size and emphasis; (4) deciding individually what research to be stressed; (5) defining and implementing the criteria and evaluation for determining the quality of faculty; and (6) defining

functionally the standards of admissions and graduation for students.

Brown (2001) defined seven groups of decision-making for faculty governance: appointment, promotion and tenure decisions, curricular decisions, faculty governance decisions, general administration, budgetary decisions, student governance decisions, and individual reward/punishment decisions. He suggested that faculty members in his study have the most control over curriculum decisions and the least control over financial decisions. Faculty members were expected to play an important role in decisions concerning curriculum and faculty governance. Brown (2001) also examined the relationship between faculty participation in university decision-making and university performance. He argued that the optimal level of faculty participation varied by decision-making types. Increased faculty participation may be good or bad; the effects varied by the type of decisions in which faculty participate.

Based on the previous literature review, this paper explores university governing structures and discusses faculty roles in university governance from a comparative perspective. The research questions are the following: (a) what are the governing structures in these two Chinese and American universities, and (b) what are the perceptions of faculty primary roles in university governance between the two universities?

Method

This study is a comparative case study. We used qualitative research methods because qualitative approach is suitable to gain insight into people's attitudes, behaviors, value systems, concerns, motivations, aspirations, culture, or lifestyles. It seeks out the "why," not the "how" of its topic through the analysis of

unstructured information, for instance, interviews, emails, notes, feedback forms, and documents. In our study, we collected such information from the faculty members at Huaxin University and Clouden University because only faculty members know what exact roles they play in their university governance.

The primary purpose of this study was to investigate the university governing structures and faculty primary roles in university governance. The goal was to compare the similarities and differences between universities in the United States and China. To achieve these goals, we chose two universities to compare: one in Midwest America (Clouden University) and one in central mainland China (Huaxin University). The reason to compare these two universities was that both universities are located in the central part of each country, where they share some similarities such as geography, economics, and open-

mindedness of people. Data were collected from structured interviews, follow-up emails, and document analysis. A sample of 20 faculty members was selected: 10 from Huaxin University and the other 10 from Clouden University. The sampling method was purposeful sampling since the first author worked at Huaxin University and the second author works at Clouden University. Upon the approval of SSIRB, recruitment letters were sent to the faculty members at Huaxin University and Clouden University by emails. Consent forms were distributed for their signatures to those who were willing to be part of the study.

Among the respondents who wanted to participate in the interviews, we selected the first 10 faculty members who responded to our recruitment emails. Incidentally, there was a gender balance of respondents in each university. The demographic backgrounds of participants are in Table 1:

Table 1

Demographic Backgrounds of the Selected Faculty Members at the Two Universities

Location	Gender (number of each gender)	Age (number of the age range)	Status (number of academic title)	Disciplines (Number of respondents)
Huaxin University	males (5), females (5)	30-40: 3 40-50: 3 50-60: 4	lecturers (3) associate professors (3) full professors (4)	Political Ideology (1), Geosciences (2), Education (2), Chinese Language (1), Mathematics (2), Economics (2)
Clouden University	males (5), females (5)	30-40: 3 40-50: 3 50-60: 4	assistant professors (3) associate professors (3) full professors (4)	Education (2), Mathematics (2), Economics (2), Psychology (1), Modern Languages (2), Medicine (1)

After the participants were recruited, we interviewed them for about 30 minutes each. During the interviews, we asked each of them nine questions on governing

structures and roles of faculty in their university governance. The nine questions are:

(1) What do you know about the university governance?

(2) What are the faculty's main roles at your university?

(3) Can you make the final decision on the curricular issues at your university?

(4) Who and what decide your appointments and promotions at your university?

(5) What role do you play in student admission, evaluation, and graduation?

(6) Can you have a voice for the selection of major administrative and academic officers, such as the president, dean, and chair?

(7) What's your role in budgetary planning of your school and university?

(8) Do you think it necessary to participate in the university governance and what about your role in your university governance?

(9) What else would you like to say about the faculty role in your university governance?

To clarify the participants' perspectives, we followed up discussions by emails. When analyzing data collected, university documentations about university governance were also reviewed and analyzed.

Results and Discussion

After analyzing data collected from interviews, follow-up emails, and university documents, the themes emerged were university governing structures and primary faculty roles in university governance. The

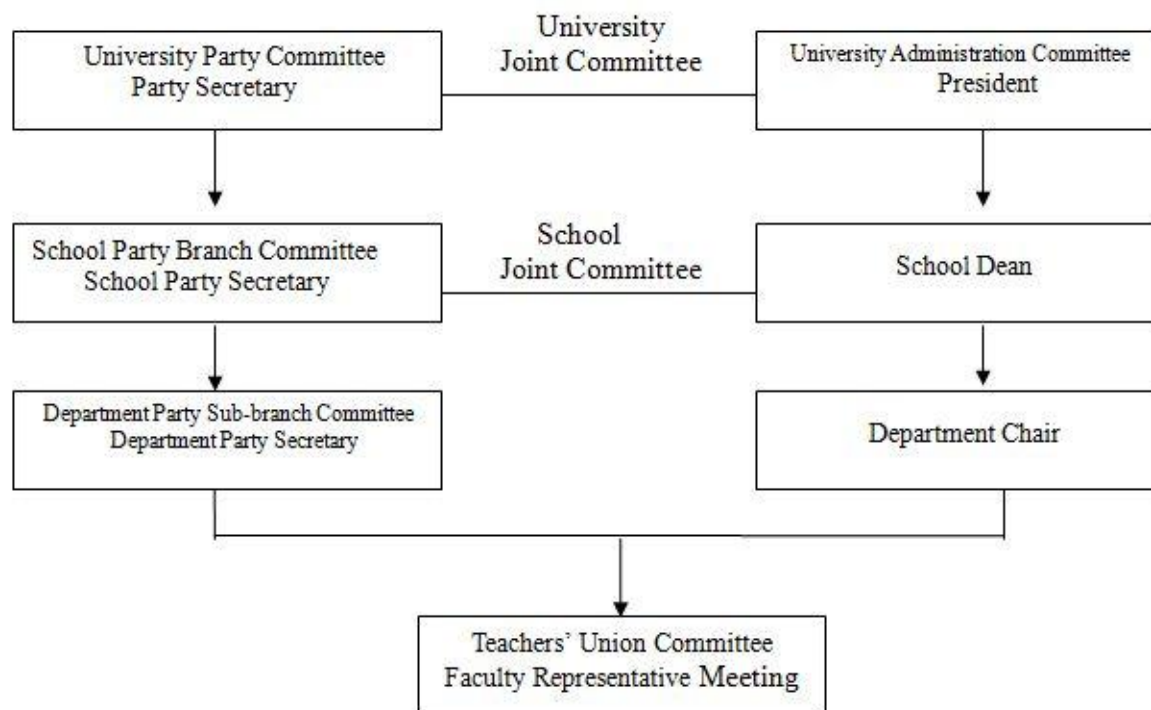
subthemes of faculty roles are curricular decision, faculty governance decision, student governance decision, and organizational governance decision. In this section, we reported different governing structures of these two universities and faculty perceptions of their primary roles in university governance.

University Governing Structures

University governance is determined directly by the governing structures of both universities. Based on the homepage of each university and information from the interviews on governance practice, the university governing structures are very different.

At Huaxin University, one distinctive feature of the governing structure was that there are two parallel governing bodies: political and administrative. The political governing body is parallel to the administrative body at all levels in the university. The political governing body is the Communist Party of China (CPC) University Committee, while the administrative body is the University Administration Committee. This binary governing structure is regarded as dual leaderships. Under this structure, there are two paralleled executive chief officers, and at each academic unit within the university, there are dual leaderships: the Party Branch Secretary (political commander) and the Dean/Department Chair (administrative commander) (see Figure 1).

Figure 1. Governance Structure at Huaxin, the Chinese University



At Huaxin University, the president is the chief executive officer for academics, and at the same time, the president is under the leadership of the University Party Committee, which is described as president-in-charge under the leadership of CPC Committee. However, it appeared that the dean has more power than the Party Branch Secretary at school and department levels. Although issues are normally discussed at the Administration-Party committee meetings and the Party Branch Secretary does have a supervisory role in assuring the academic administrators at school or department levels to implement their authority properly and efficiently, on critical issues, the dean or the chair has the final decision.

In terms of faculty roles in university governance, there are no legislative acts regarding faculty involvement at Huaxin

University. Huaxin faculty members have limited impact within the governing structure of the university. Under this structure, there is a Teachers' Union Committee, consisting of six departments, one office, 24 branch unions, and a membership of more than 2,800. Its main duty is to serve the faculty and staff other than participating in internal government. Every year the Teachers' Union Committee has a Faculty Representative Meeting gathering the six committees. These six committees are Faculty Welfare Committee, Teaching and Research Committee, Budget Committee, Housing Committee, Resolution Committee, and Finance Monitoring Committee. In reality, the Teacher's Union Committee performs as a trade union and under the leadership of the University Party Committee. The role that the Teacher's Union Committee plays in actual university

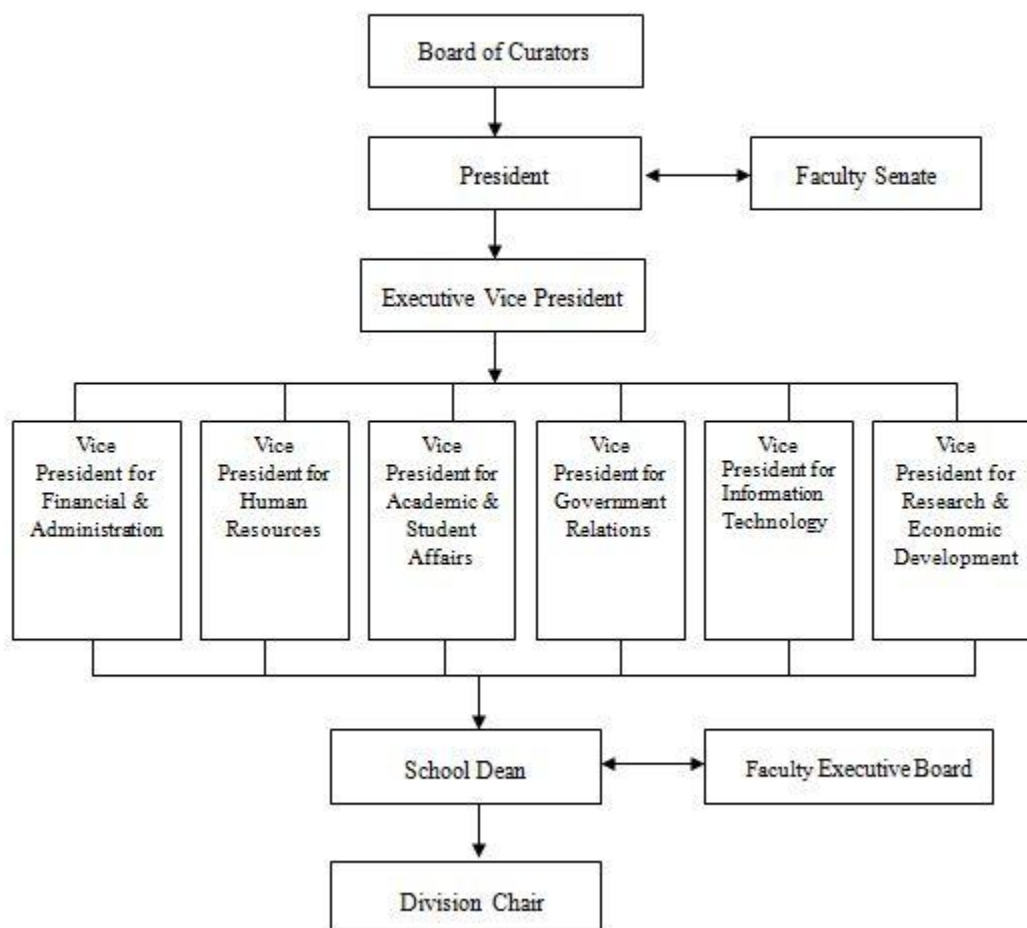
governance is limited. Furthermore, most of the Teacher's Union Committee members hold leadership positions at the university, schools, or departments. Thus, almost no avenue exists for the ordinary faculty to have input into the governing structures at Huaxin University.

Different from the governing structures of Huaxin University, the organization of internal governance at Clouden University is composed of a governing board (board of trustees or board of curators), the university president with a team of administrative president and staff, faculty senates, academic deans, division chairs, and usually

some form of organization for student representation.

At Clouden University, the distinctive feature in the governing structures is that the Board of Curators has the final decision on every important issue. The president is the chief executive officer and stands for the Board of Curators to govern the university. This university also has six executive vice presidents in charge of different aspects of matters. At the school level, the dean is the CEO and responsible to the president's inquiry. The division chair is on the lowest level of the administration (see Figure 2).

Figure 2. Governance Structure at Clouden, the American University



As for the faculty roles in university governance, there are legislative acts regarding faculty. The faculty have proper place within the governing structures at Clouden University. There is a Faculty Senate at the university level representing the faculty members. The Faculty Senate represents the campus as it gives advice to the president and comments on the actions of the president's cabinet (see Figure 2). The Faculty Senate consists of 30 senators from nearly every school and its 24 departments. The 24 departments are involved in nearly every aspect of faculty governance accountabilities such as Campus Promotion and Tenure Advisory Committee, Academic Grievance Hearing Panel, Standing Committee on Research Dishonesty, Campus Faculty Committee on Tenure, Program Evaluation Committee, University Budget Committee, IT Privacy Oversight Committee, the Parking and Traffic Committee, and so on. At the school level, there is also a Faculty Executive Board to serve the faculty governance. The senators of Faculty Senate and the members of the Faculty Executive Board are all tenured faculty members. Both the Faculty Senate and the Faculty Executive Board can write recommendations about the governance issues they discussed, but they do not have the power to sign the final decision paper. All final decision papers must be signed by the president or the dean.

Even as is, the role faculty can play in actual university governance is substantial at Clouden University. For example, in 2005, the faculty senate of Clouden University successfully forced the president of the university to resign. According to the university record, the president held a grossly unrealistic vision for the university and had committed some major injustices to the faculty membership. Although having held several prestigious positions such as the

Council of Advisors on Science and Technology (PCAST) appointed by former president George W. Bush, the Nebraska Power Review Board by Governor Bob Kerrey, the Research and Education Advisory Panel to the U.S. General Accounting Office by then-Comptroller General of the U.S. Charles Bowser, the Clouden President still had to answer the Faculty Senate's questioning about her hiring of a friend. The then-President hired her friend, a California consultant to organize a breakthrough team to "transform" the campus. This team had the idea of making the university "better than Harvard." Two of the goals of this breakthrough team included enlarging the university endowment by \$500 million and installing a computer data port at every desk in every classroom. But two weeks into the semester, many students still did not have textbooks and the number of book problems that year was unusually high (Blackwood, 2001). Some deans and professors chafed under the transformation effort, which cost the university hundreds of thousands of donated dollars and took up time that they thought might have been better spent counseling students, writing papers, or researching scientific questions (Blackwood, 2001). However, the president thought of these deans and professors as "terrorists." Finally the president lost the confidence of many professors, who accused her of unfairly cutting them out of the decision-making process. The Faculty Senate's confidence motion towards the president demonstrates that there was an actual avenue for Clouden University faculty members to administrate influence into the university governance.

The Faculty Primary Roles in University Governance

The principal responsibilities of the faculty are teaching, researching, and

service. An important additional responsibility of the faculty is to ensure that the university fulfills its educational mission. The faculty therefore must be involved in the generation and implementation of policies that affect the university's mission. On matters primarily affecting the academic mission of the university (curriculum, subject matter and methods of instruction, degree requirements, faculty scholarship, faculty status, and faculty service), the principal responsibility for formulating and evaluating ideas lies with the faculty.

The themes that emerged from data regarding faculty primary roles of these two universities are curricular decisions, faculty governance decisions, student governance decisions, and organizational governance decisions. In this section, we report faculty members' perceptions regarding these four primary roles in each university.

Curricular decisions. The term "curriculum" carries with it a multiplicity of meanings. No matter how it is conceptualized, curriculum at a college or university is the fixed series of studies required for graduation or qualification in a major field of study in the various schools within the university system.

At Clouden University, the curricular goals and standards are often determined by outside authorities such as the State educational authority and national professional organizations. All syllabi must include these goals, standards, and learning outcomes that are congruent to goals and the standards. The faculty members are entrusted to make important curricular decisions. Many faculty members see curriculum development and revisions as their exclusive domains and one of the primary responsibilities.

At Clouden University, faculty members collectively (not individually)

make decisions about curriculum. A participating professor from the education department stated:

Curricular issues and formulations are usually first discussed by a concerned faculty group or committee in a particular discipline or field of study, and a proposal is written expressing the curricular concerns and recommendations. Once the proposal is developed by the faculty group or committee, the proposal is presented to a larger group such as the faculty of a Division of a School for its discussion and vote of approval. Once the Division votes its approval for the proposal, the proposal is submitted to the School's democratically elected curriculum committee for its vote of approval, which is called the Committee for Academic Programs and Standards (CAPS). Once approved by this curriculum committee, the proposal is submitted to the School faculty for its vote of approval. If approved by the school faculty, the proposal is going to be forwarded for official approval of the dean at the School and then to the Office of the Provost and President of the University. The newly adopted curricular issue or formulation becomes part of, or is reflected in, the university's catalogue and is published in hard copy and electronically online. (January 10, 2009)

Although accrediting bodies indirectly have more power over faculty concerning curricular issues, the faculty can submit applications to revise, modify, or create a new course, and the faculty are still regarded as the "experts" on curriculum development

and are usually allowed to decide what to teach and how to teach it. A mathematics professor stated proudly, “I am the boss about curricular decisions!”

At Huaxin University, the curriculum goals and standards are also determined by outside authorities, and the faculty syllabi must also include these goals, standards, and learning outcomes that are congruent to goals and the standards. Faculty members are limited in authority to make important curriculum decisions. Some curricula, such as political ideology, are formulated by the National Ministry of Education. The faculty members do not have any choice in the matter of what and how to teach. Just as an associate professor of political ideology said,

Political ideology is a very important course. It relates to the main ideology of our country. So I must teach the students according to the selected textbooks strictly and mustn't speak freely, especially against our government. That is, I just do what I am told to do. (January 15, 2009)

As for other curricula, the university usually has a curricular list for faculty members to choose from. But at the department level the faculty members usually have less choice for the curricula. The fact is that the curricula are assigned to the faculty by the dean or the chair. In selecting textbooks, the National Ministry of Education compiled a recommended list of textbooks for universities to choose from, and the faculty members are only free to choose which textbooks to use from that list. Although they are free to decide the way that they wish to teach, that “We have some limited freedom to choose what to teach and how to teach it” is a common voice when the Huaxin University faculty members were asked about the curricular decisions.

Faculty governance decisions.

Faculty governance includes many aspects. At Clouden University it is concerned mainly with new faculty appointment, faculty promotion, and faculty status. As stated by a psychology professor, the process of appointing a new faculty at Clouden is like this:

When appointing new faculty, this faculty position should be advertised, and a Search Committee of faculty members is convened to screen the applicants and to determine those candidates who best fit in to the position. The committee appointed by the dean consists of faculty members in the division who will be working with the new faculty member, and another faculty member outside the division. Students are also invited to participate. The faculty members screen the applicants and determine the best ones to visit campus (usually three are invited). All faculty members in the school are invited to attend presentations by the applicants and to give feedback to the search committee about the best candidate for the position. The search committee recommends and ranks the applicants. The dean makes the selection, but the dean almost always respects the recommendation of the faculty on the search committee. (January 11, 2009)

Faculty promotion and tenure at Clouden University are determined by a committee at the school level and by a committee at the campus-wide level. Both of these committees are democratically elected by faculty members within the various schools and colleges of the university. The data submitted for promotion and tenure by a faculty member must relate to the faculty

member's past performance in the areas of research, teaching, and service. A professor of medicine said that:

An extremely heavy emphasis in promotion and tenure decisions is placed on the faculty member's research – particularly the publishing of results of the faculty member's research, and peer-reviewed journals. The dossier submitted by the faculty member include such things as student ratings of teaching; letters from independent reviewers outside the university; and documentation of the faculty member's publications. However, it is never a completely black and white criteria because different faculty in different divisions have different responsibilities; for example, faculty in the fine arts are expected to put on performances, while faculty in sciences are expected to get grants and do research in laboratories. There are committees in each school to help determine if the faculty member has met the criteria for promotion and tenure and the suggestion of promotion and tenure committee from the school is the most influential. The promotion and tenure committee recommends faculty for tenure to the larger school committees. (January 14, 2009)

The decisions of promotion and tenure committee and the approvals of administrators must comply with bylaws, policies, and procedures approved by the university governing board. Deans and administrators approve the decisions recommended by the promotion and tenure committee. In rare cases, the administrators will overrule faculty in decisions relating to promotion and tenure.

As to the faculty status, politics play a big role at Clouden. An assistant professor in modern languages stated that:

Faculty of color have a more difficult existence at the university because while they may be content experts, it's hard for their colleagues to see them as such because the first things that many people see is the color of your skin before they see that you're actually an expert with their same qualifications and knowledge. At the same time, this is a difficult issue because people can also forget to see you as a faculty of color that is interested in issues that affect communities of color, and they may not understand how to be respectful of the cultural understandings and contributions that you bring to the university. (January 20, 2009)

At Huaxin University, the university personnel department is in charge of appointing new faculty. A Chinese language professor stated this:

When appointing new faculty, this faculty position is advertised, and a Search Committee appointed by the school dean is convened to screen and interview the applicants. They determine who best fits in to the position. The committee consists of the administrators in the university personnel department, the head of the school (dean, associate dean, school party secretary, and deputy party secretary, etc.) and some professors in the division. The search committee makes the selection and submits the decision to the university personnel department for approval. (January 18, 2009)

There is no tenure system in China. Faculty promotion at Huaxin University is

determined by an academic committee at the school level and by a promotion committee at the university level. But most of the committee members are administrators. Thus, the faculty has a very weak voice for promotions. A geosciences professor of Huaxin University said that:

One's promotion usually depends on three aspects: the availability of a vacant position, performance, and personal relationship. Having a vacant position is a very important issue because the number of positions in the university is determined by the Ministry of Education. The faculty performance can be measured by the prescriptive rules made by the university. The faculty member's research or scholarly activity including peer-reviewed journals, publications, and grants are extremely emphasized in promotions. The relationship aspect is also important and necessary in faculty promotions because all the committees in this university are led by the party and administration, all the decisions are made by the heads of the university (including party secretary, president, etc). Having a good personal relationship with the leadership of the university is very helpful for the promotions. (January 22, 2009)

At Huaxin University, there is little discrimination based on religious beliefs or ethnic backgrounds among the faculty members, according to all interviewees. Generally speaking, faculty members (especially for the ones who are not full professors) have less power and a much weaker voice about faculty governance decisions than the administrators. Most of the participating professors stated the similar opinion that:

In this university, the faculty members feel that they are inferior to the administrators when it comes to faculty governance decisions, and are eagerly waiting for their status to be improved. (Economics professor, January 24, 2009)

Student governance decisions.

Student governance concerns student academic life. It includes both student affairs and academic affairs. The student affairs are concerned about the out-of-classroom student services and programs, while academic affairs relate directly to the educational process and the granting of degrees. Here we limit our discussion of student governance decisions to student admission, evaluation, and graduation.

At Clouden University, the Office of Student Affairs is under the authority of the Vice President for Student Affairs. The student academic affairs are governed by the faculty members. Typically, each school establishes criteria for admissions. The staff at the Admissions Office review undergraduate applications and faculty members decide the graduate admissions. An assistant professor of education explained the graduate admissions process like this:

If you're a graduate and doctoral faculty, you're allowed to teach at that level, and you have a say in admissions of graduate students. For graduate students, first they must meet the graduate admission requirements of the university, and then the applications are forwarded to the schools so that the faculty in the respective divisions can accept or deny the applicants. (January 21, 2009)

Once students are admitted to a school, faculty members and department set evaluation standards in their programs to

assess the students' performance. The faculty members evaluate students with coursework, program of study, and Master's thesis or doctoral dissertation. A psychology assistant professor commented that:

Graduation criteria are usually based on successful completion of coursework and their theses or dissertations. The students' progress is usually tracked by the department. But faculty members are ultimately responsible for determining who meets criteria for graduation. (January 26, 2009)

At Huaxin University, admissions are under the charge of the Student Affairs Department and student academic affairs are overseen by the faculty members and departments. The admissions quota is set by the National Ministry of Education. It requires that all students pass the college entrance examinations for admissions. Admissions for both undergraduate and graduate students mainly depend on the students' scores on the entrance examinations. Just as an associate professor of economics said that:

Scores are the most important after all. The faculty's role in the student admissions is limited; but now the professor is playing a more important role in the doctoral student admissions process. At some universities, prestigious professors can waive entrance examination requirements for particularly talented graduate students. (January 28, 2009)

Similar to that at Clouden University, the evaluation of student academic performance at Huaxin University is determined by faculty members. But the graduation of students is determined by many factors, including the student's academic and non-academic performances during the course of their university study. The evaluations are

administered by different departments inside or outside the campus. For instance, their foreign language performance is assessed by the College English Test, a proficiency test administered by the National Ministry of Education.

Organizational governance decisions. Here we narrowly define organizational governance as some important decisions of a university, namely, the selection of primary administrators, academic officers, and budgetary planning.

At Clouden University, the final decision on the selection and appointment of administrative and academic officers is almost always made by administrators. The governing board (Board of Curators) has the final word in selecting presidents. Usually search committees are charged with finding, interviewing, and selecting two or three candidates for final consideration. The search committee includes faculty members and perhaps other persons such as administrators, students, community members, etc. Administrators usually ask for input in the form of questionnaires, rating scales, etc. from faculty members and other concerned parties. But the chief administrator makes his or her appointment. An associate professor of education commented that:

The amount of collaboration involved in making an appointment depends, of course, on the type of appointment, the publicity surrounding the appointment, and the democratic nature or democratic orientation of the particular administrator making the appointment. (January 22, 2009)

There are some differences in the levels of publicity and faculty input regarding the selections and appointments of administrative and academic officers between the Chinese and American

universities. Generally, the American university has more publicity about the candidates who are going to fill the positions. However, the degree of publicity varies among different positions at Clouden University. Regarding this, the education professor further concluded that:

Usually, the more local the appointment, the more likely that faculty will have a voice. The faculty members in the division have a direct vote on the selection of a chair. For the selection of a dean, the faculty may or may not get on the search committee that hires the dean, but the faculty may get to fill out a survey and provide their opinions and assessment based on their interaction with the candidate. For the selection of a president, usually there is one or two faculty members from each school that get to sit on the search committee to select the president. (January 22, 2009)

In the 2008 presidential selection of Clouden University, the Board of Curators entrusted an international firm to manage the search for the university president. This international firm took care of advertising candidate recruitment, collecting feedbacks, and nominating candidates. After the firm submitted the candidates list to Clouden University's Board of Curators, the Board named a search committee to interview candidates and to provide a list of unranked recommendations. The 16-member search committee represented the university faculty, staff, students, administrators, alumni, and the community. Through these procedures, the new president was selected. During this presidential selection, the faculty could put forth their suggestions as they like.

At Huaxin University, the chief administrators, including the president, vice president, party secretary, and deputy party

secretary, are all named by the Ministry of Education. At the school level and department level, the immediate upper level of administration is in charge of the selection and appointment. The search process includes promulgating the advertisement of candidate recruitment, collecting the feedback, nominating the candidates and organizing the interviews. The interview committee is composed of heads of schools and university administrators. After the interviews, the joint Administration-Party Committee discusses and makes the selection. Then the particular department publicizes the selection outcomes and leaves three to six days for feedback (but it is usually just a routine procedure). After this period, the particular department makes the appointments. A lecturer of mathematics commented that "During the whole process, faculty members have very weak voice for it. Some process is just showing the routine."

In addition to faculty/administrators selection and appointment, budgetary planning is another important organizational governance investigated in this study. At Clouden University, the school budgets are determined by the number of student enrollments and programs. The faculty role in the budgetary planning mainly happens at the school level. But most Clouden faculty members felt that they have played no roles in budgetary planning unless they're elected to sit on the faculty executive board, which consists of about two faculty members per division. In the faculty executive board, they can interact with the dean's office more closely on budgetary matters. The Faculty Senate at the campus level has a direct and influential role in budgetary planning of the entire university. Two or more faculty members are selected from each school to form the Faculty Senate.

As for budgetary planning of Huaxin University, it is the same as faculty/administrator selection. Faculty members have no power at all. The university budgetary planning is controlled by the financial department, which is also led by the university president. The school budgetary planning is controlled and led by the dean, who is consulted along with other heads of the school and perhaps some professors. The notion that, "I know nothing about our university and school budgetary planning, neither do I have ways to know about it," is common when the Chinese faculty members answered the questions of their roles in university and school budgetary planning.

Conclusion

This paper investigated university governance in terms of governing structures and faculty primary roles at one Midwest American university and one central Chinese university. There appeared distinctive governing structures in Clouden University (the American university) and in Huaxin University (the Chinese university). The governing structures in the two universities are quite different because of their different leaderships. At Clouden University, the Board of Curators has the final decision on every important issue. The university president is the chief executive officer and stands for the Board of Curators to govern the university. At Huaxin University, the governance structure is described as the president-in-charge under the leadership of the CPC committee. Political and administrative leadership are two parallel governing bodies. All important governing issues are decided by the joint Administration-Party Committee at the university levels and most school/department levels. The dual leadership governance structure of the Chinese university made the university a

much more bureaucratic organization resulting in far less academic autonomy.

The different governing structures of these two universities made the primary faculty roles concerning university governance quite different in the areas of curricular decisions, faculty governance decisions, student governance decisions, and organizational governance decisions. Compared with the faculty at Clouden University, the faculty members at Huaxin University seemed to have less power, and they played less important roles in university governance. However, there are some factors that both universities have in common, that is, the evaluation of student academic performance and weak voice on the budgetary planning. The faculty members of both universities have no final decisions in many university governance decisions. Especially in China, the faculty roles in university governance are always invaded and even substituted by administrators, and all the faculty members of the two universities think that it is necessary for faculty members to participate in the university governance and both groups seek more say in final decisions concerning university governance.

The perspectives of the faculty members at both Clouden and Huaxin revealed that university governance is a very complicated process. Effective university governance generally requires that all the stakeholders, such as the faculty, professional and supporting staff, students, and external constituents, be represented on budgetary, policy, and procedure decision-making entities. Appropriate representations of these groups are normally obtained through the university's council and committee structures. Elected and appointed representatives should, as far as possible, be selected specifically for the roles in which they will serve. That is, shared governance

is an ideal governance model for the university.

Shared governance needs all of the stakeholders participating to realize their responsibility for decisions made in the governing process. It only works where there is a high level of participation from all the stakeholders. It seems that there is a long way to go for both universities, but it is particularly suggestive that the Chinese university develops authentic shared governance.

University shared governance logically should follow the university development with respect for the faculty members, and should endow them with substantial authority in the university governance system. Further, we need more transparency with regard to all levels and types of university governance. We also need to develop policies and procedures that will assure transparency and collaboration in university governance at every school and college in the university system.

For both Huaxin University and Clouden University, the road seems long in developing authentic shared governance, especially at Huaxin University. It is necessary to develop policies and procedures that assure transparency and collaboration in university governance. An authentic, ideal, “shared governance” university system will never be possible as long as some individuals in the university community lack integrity and a democratic nature or democratic orientation.

Limitations of the Study

In the current study we investigated the university governing structures and faculty members’ primary roles on university governance decision-making. In addition to the distinctive structural differences, four faculty roles in curricular decision, faculty governance, student governance, and organizational governance emerged from the

data collected. However, university governance is very complicated and includes more than those investigated. A sample of 20 faculty members at the Huaxin University and Clouden University were not sufficient to generalize our findings to other universities in China and the United States. The findings may not be suitable for interpreting faculty roles in other universities either in China or in America. Future research in the topics alike should enlarge the participant pool as well as disciplines and schools. Focus of analysis in faculty perceptions of university governance should also include the variables such as age, gender, race, and academic status, etc.

Acknowledgement

The project was supported by the Humanities and Social Sciences Research Fund of Chinese Ministry of Education (09YJC880099) and the Special Fund for Basic Scientific Research of Central Colleges, China University of Geosciences, Wuhan (CUG090119). The authors would like to express their heartfelt gratitude to the editor and the reviewers in improving the paper.

References

- American Association of University Professors (AAUP). (1940). *1940 Statement of principles on academic freedom and tenure*. Retrieved from <http://www.aaup.org/AAUP/pubsres/policydocs/contents/1940statement.htm>
- Altbach, G. P. (2005). Patterns in higher education development. In P.G. Altbach, R.O. Berdahl, & P.J. Gumport (Eds.), *American higher education in the Twenty-first century: Social, political, and economic challenges* (pp. 16-18). Baltimore, MD: The Johns Hopkins University Press.
- Arimoto, A. (2001). *University reforms and academic governance: Reports of the 2000 three-nation workshop on academic governance*. Hiroshima, Japan: Research Institute for Higher Education, Hiroshima University.
- Benjamin, R. & Carroll, S. (1996). Impediments and imperatives in restructuring higher education. *Educational Administration Quarterly*, 32, 705–719.
- Blackwood, K. (2001). Brain wash. *Pitch Weekly*, 21, pp.15-21.
- Brubacher, J.S. (1982). *On the philosophy of higher education*. San Francisco, CA: Jossey-Bass Publishers.
- Brown, W. O. (2001). Faculty participation in university governance and the effects on university performance. *Journal of Economic Behavior & Organization*, 44(2), 129-143.
- Ehara, T. (1998). Faculty perceptions of university governance in Japan and the United States. *Comparative Education Review*, 42(1), 61-72.
- Gayle, D. J., Tewarie, B., & White Jr., A. Q. (2003). *Governance in the Twenty-first-century university: Approaches to effective leadership and strategic management: ASHE-ERIC Higher Education Report*. San Francisco, CA: Jossey-Bass Publisher..
- Han, Y. (1993). A comparison of organizational structures in universities in Australia and China. *Journal of Educational Administration*, 31(2), 63-83.
- Kezar, A. & Eckel, P. D. (2004). Meeting today's governance challenges. *The Journal of Higher Education*, 75(4), 371-398.
- McCormick, R. & Meiners, R. (1988). University governance: A property rights perspective. *Journal of Law & Economics*, 31(2), 432-442.
- Zhou, N. (1989). *Higher education reforms in post-Mao China (1978-1988)*. Unpublished doctoral dissertation. The State University of New York at Buffalo, Buffalo, New York.

Chinese as a Foreign Language Learners' Strategy Use and Writing Achievement

Ko-Yin Sung, Utah State University

Abstract

This study reports on the relationship between four college-level Chinese as a foreign language learners' strategy use and their writing achievement. Data included the learners' responses to Oxford's (1990) Strategy Inventory for Language Learning (SILL) survey, learner interviews, classroom observations, and learners' writing samples. This study found that the two learners whose instructor regularly trained them to use specific writing strategies, such as repeating and translating, tended to be at a higher writing level compared to the other two learners whose instructor emphasized the practice of oral skills. This finding implies that language instructors need to help develop learners' four language skills equally and explicitly teach learning strategies that help develop them.

Introduction

Language learning strategies, defined as "specific actions taken by the learner to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations" (Oxford, 1990, p. 8), provide learners means of active involvement, which are necessary for developing second language proficiency (Oxford, 1990). Research shows that learners who intentionally select and combine strategies relevant to a given language task show improved proficiency in the target language (Grenfell & Macaro, 2007). Most importantly, strategies are

teachable to learners; hence, studying learners' strategy use provides language educators data regarding which strategies students are unaware of, which helps language educators determine what strategies need to be explicitly introduced.

The learning phenomenon of strategy use is important to investigate as strategies are associated with successful learning (Grenfell & Macaro, 2007). However, studies of foreign language learning, specifically language learning strategy use, have predominantly focused on learning English as a foreign or second language (Peacock & Ho, 2003; Poole, 2005; Shmais, 2003; Yongqi, 2005). Currently, there is a lack of research on strategy use in learning Chinese as a foreign language. This study of Chinese language learners' strategy use and its relation to language achievement provides valuable findings and suggestions which add to the existing knowledge of the Chinese language teaching and learning field.

Theoretical Framework

Ellis' Strategy Framework

Language learning strategies are positioned by Ellis (1994) as having the "mediating role" between learner factors and learning outcomes (p. 529). Figure 1, adopted from Ellis (1994), illustrates the relationship between individual learner differences, situational factors, learning strategies, and learning outcomes.

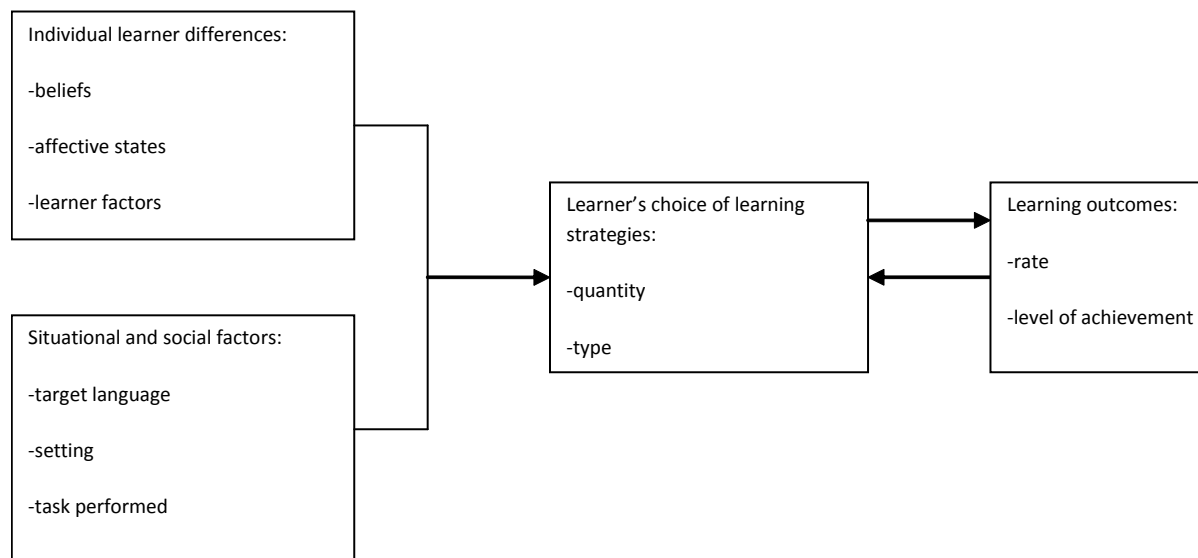


Figure 1. Strategy Framework: The Relationship between Individual Learner Differences, Situational Factors, Learning Strategies, and Learning Outcomes. From Ellis (1994, p. 530).

Ellis (1994) defines strategy as “a mental or behavioral activity related to some specific stage in the overall process of language acquisition or language use” (p. 529). He explains that individual learner differences, together with different social factors, influence learners’ strategy use. For instance, individual learner factors such as second language (L2) learners’ motivation in learning the target language with social factors such as teaching instruction received, may affect the learners’ strategy use. Ellis (1994) points out that learners’ choices of strategies affect the degree of success in language learning in terms of the rate of acquisition and the level of achievement. For example, certain strategy use in a given language task may result in higher L2 performance while certain strategy use may not be as efficient in the same type of language task. Finally, Figure 1 shows that the level of success experienced and the level of L2 proficiency affect learners’ strategy use. This study focused on the last segment of Ellis' (1994) framework in which

learners' choice of language learning strategies in relation to their writing achievement were investigated.

Oxford's Taxonomy of Strategies

Researchers in the second language acquisition field have attempted to identify learners’ strategy use and categorized them into different types of strategies. For example, Oxford (1990) proposes six categories of strategies: cognitive, metacognitive, memory, compensation, affective, and social strategies, and explains how each category aids the development of communicative competence. Oxford (1990) explains that the six categories of strategies are grouped into two different types: direct and indirect. Memory, cognitive, and compensation strategies are direct strategies which are used by learners for immediate response to language tasks. They are considered direct strategies because all three categories require “mental processing of the language” (p. 37). Each of the three types does the processing in different ways for

different purposes. For example, Oxford (1990) defines memory strategies as having the function of helping learners group comprehensible input and retrieve information. Memory strategies such as using flash cards or grouping words of similar functions help learners remember newly learned vocabulary. Next, cognitive strategies, defined as “manipulation or transformation of the target language by the learner” (Oxford, 1990, p. 43), have the function of helping learners to understand and produce L2. For instance, learners use the cognitive strategy of skimming to locate the main idea in a text. Last, compensation strategies, “intended to make up for an inadequate repertoire of grammar and, especially, of vocabulary” (Oxford, 1990, p. 47), allow learners to use clues for guessing, leading to more comprehension.

On the other hand, metacognitive, social, and affective strategies are indirect strategies learners use to prepare for language tasks and help learners gain more control of their learning. Metacognitive strategies are those used to organize learning so that learners may more easily coordinate the second language acquisition process. For example, learners set goals and identify the purpose of a language task. Next, affective strategies, defined as strategies which help learners be more active in language learning, include strategies such as giving oneself a valuable reward for a good language performance and writing a diary to keep track of one’s personal experience in the language learning process. Social strategies are used to learn the target language through interaction with others. Together these strategies aid the development of communicative competence. Among the many categorizations of strategies created by scholars, Oxford’s (1990) taxonomy of strategies is considered the most comprehensive (Ellis, 1994). Hence, Ellis’ (1994) language learning strategy

framework and Oxford’s (1990) taxonomy of strategies were used to guide the study in investigating language learners' strategy use and its relation to their writing achievement.

Literature Review

Chinese Language Learning Strategies

Language learning strategies have been seen as one of the most important factors in predicting foreign or second language performance (Oxford, 1990). Several studies (Shen, 2004; Wang, 1998) on Chinese language learning strategies primarily focused on a specific aspect of Chinese learning skill and Chinese character learning. For instance, Shen (2004) investigated the effects of three encoding strategies for Chinese character learning among students learning Chinese as a foreign language. The three encoding strategies identified in the study were rote memorization, student self-generated elaboration, and instructor-guided elaboration. Rote memorization refers to the use of repetitive rehearsal strategies such as listening to the instructor present the sound, shape, and meaning of words repeatedly without elaboration, and have been observed to be related to shallow processing in the memory. Students’ self-generated elaboration strategies, in which learners use any elaboration strategy they prefer, and instructor-guided elaboration strategies, in which the instructor explains the meanings or gives examples of the lesson taught, both resulted in a deeper processing in the memory, resulting in a significantly higher retention of sound and meaning of a character than shallow processing (Shen, 2004).

In another study of Chinese language learning strategy use, Wang (1998) investigated 15 first-year Chinese language learners in an American university. Results indicated that 80% of the learners used repetition strategy in practicing writing

Chinese characters. Ninety-three percent of the learners reported that they used memorization strategy. In examining students' cognitive strategy use, only 20% of the learners used a categorization strategy to study Chinese characters. Sixty percent of the class used a read aloud strategy to study Chinese, and 73% of them used a translation strategy. Wang (1998) concluded that the low percentage of certain Chinese character learning strategies could be due to the instructions learners received. The Chinese instructor did not spend time teaching Chinese characters, but placed more emphasis on listening and speaking skills. Wang (1998) suggested that even though the learners have used a wide variety of strategies in learning Chinese, they could have benefited more if they were provided with more Chinese character learning strategies in class. This suggestion may imply that teachers' instructions and beliefs about effective strategy use may affect learners' strategy use.

Although language learning strategies have long been identified as important factors in learning a second language, research on Chinese language learning strategies is still in its infancy. Current studies on Chinese language learning strategies are scarce. Hence, this study is needed in hopes that it will contribute to the current literature.

Strategy Use and Language Achievement

Current studies on strategy use and language achievement seem to suggest that certain strategies lead to more successful learning. In other words, among language learners who use a variety of learning strategies, some learners are more successful in learning the target language than others. For example, participants who identified their own ways of learning English in Yongqi's (2005) study on English vocabulary learning strategies researched different degrees of language achievement

in class. A participant who used the strategy of making frequent contact with native speakers of the target language reached higher achievement on vocabulary tests than a participant who spent the majority of her time on memorizing word lists. Moreover, Andreou, Andreou, and Vlachos' (2004) study showed that the more successful students used combinations of strategies more frequently compared to others. The interpretation of the finding is that common strategies students used alone were not adequate to move learners to higher proficiency levels. It is the diverse combinations of strategies the successful learners used to promote high achievement in language learning. Although these studies suggest that certain strategies lead to more successful learning and others yield minimum language achievement, the studies only reviewed English as a foreign language. Chinese, a logographical language, on the other hand, has distinct features compared to English. Hence, my study adds value to the literature with a comprehensive description of language learning strategy use and language achievement with an emphasis on American learners of Chinese language.

Methods

Research Sites

This study involves two first-year Chinese classes in two research sites, Santos and Triangle Universities, both of which are higher education institutions located in a large metropolitan city in Texas. The following paragraphs briefly describe the Chinese language programs in the two universities.

Santos University: Santos University in Texas is one of the state's fastest-growing public schools, with more than 20,000 students in graduate and undergraduate programs in 2007. Starting in 2006, the College of Education and the

College of Business at Santos University cooperatively developed an elementary Chinese language course in response to increasing student needs for learning Chinese language skills. The demand from students for learning Chinese is evident as the Chinese language course expanded from one class to two in 2007.

Triangle University: Triangle University, a private institution, is one of the only two universities in Texas to offer an undergraduate degree in Chinese. The program has existed since 1990, and has continued to expand.

Participants

In order to eliminate the effect of writing level difference on the measurement of writing achievement, this study intended to only look for novice learners in Chinese writing with CHLL or non-CHLL backgrounds. The participants involved in the study were four first-year Chinese language students in their second semester of study, which represented the majority of student backgrounds in the first-year Chinese classrooms. The four participants included one Chinese heritage language learner (CHLL) and one non-CHLL from Santos University, and one CHLL and one non-CHLL from Triangle University. Tim (non-CHLL) and Luke (CHLL) were enrolled in Professor Le's class at Santos, and Jenny (non-CHLL) and Yaoming (CHLL) were enrolled in Dr. Fu's class at Triangle. The two non-CHLLs, Jenny and Tim, are native speakers of English who had no Chinese language experience prior to the Chinese class. One of the CHLLs, Yaoming, whose father immigrated from Hong Kong and mother from Beijing, was born in the United States. At home Yaoming's parents' language of communication was Cantonese, but they talked with Mandarin-speaking friends in Mandarin. Growing up Yaoming always spoke English with his brother and

parents even though his parents spoke Cantonese to the children. The other CHLL, Luke, whose parents are Mandarin speakers from Ningpo, China, came to the United States when he was 3 years old. Luke's parents sometimes spoke in Mandarin Chinese, sometimes in English, and sometimes mixed the two languages. However, Luke only responded to them in English. Despite the family domain, Luke's friends were American; therefore, he never used Chinese in the school domain. Both Yaoming and Luke were sent to Mandarin Chinese Saturday schools when they were in fifth and sixth grades, but neither of them had interest in learning Chinese and withdrew after a couple lessons. Based on the previous Chinese language experiences described above, the two CHLLs who did not have any informal or formal training in Chinese writing at home or in school prior to enrolling in the classrooms under study fitted the criteria for the participants in this study.

Data Collection and Analysis Methods

Data for this study were collected in the two research sites from February to May in 2008. I used several data collection methods to investigate strategy use and language achievement. In particular, the learners' writing samples were used as their language achievement measure, and their responses in the SILL surveys and pair interviews were used to summarize their strategy use. The classroom observation data served as the learning context in this study. The following paragraphs discuss how each data collection instrument was designed and used.

Oxford's strategy inventory for language learning (SILL) survey. In order to measure language learning strategy use, Oxford's (1990) SILL for native speakers of English learning another language was used in this study. The SILL was used because it

is a widely used language learning strategy survey with high reliability and validity. The SILL for native speakers of English learning another language comprises of a five-point Likert-scale that assesses the frequency of learners' use of strategies. The scores range from 1, being "never or almost never true of me" and 5, being "always or almost always true of me." The SILL contains 80 items grouped into the six categories of strategies in Oxford's (1990) taxonomy of language strategies. The four participants' average score in each of the six strategy categories were calculated.

Samples of learners' writing. The participants were asked to complete a writing assignment consisting of two writing prompts toward the end of the school semester. I decided to use the learners' writing and not other language skills as part of the language achievement data due to the limited writing activities and assignments given in class. The participants' only writing assignments were character writing practices in which they copied new characters learned multiple times. Thus, it would be difficult to evaluate the participants' writing skill based on such writing practices. Also, making my own evaluation of the learners' writing would ensure consistency, as it is possible that the instructors would not apply the criterion the same way. As a result of these factors, I designed a writing assignment containing two writing prompts, each of which allowed the learners to use vocabulary and language structures learned throughout the semester in the class in completing the writing assignment. The learners' writing samples were evaluated and levels determined in order to investigate their writing levels and strategy use. The writing proficiency section in the ACTFL language proficiency guidelines were used to guide the evaluation of the learners' writing samples. The guidelines helped

determine the Chinese writing level of the learners in regard to the amount of expressions, characters, vocabulary, and grammar utilized and the accuracy of their writing. The ACTFL guidelines are divided into four levels: Novice, Intermediate, Advanced, and Superior. The Novice, Intermediate, and Advanced are further subdivided into Low, Mid, and High, e.g., Novice-Low, Novice-Mid, and Novice-High. The analysis took the narrative form in which the learners' writing samples or segments of the writing samples were described using the criteria of the rubric and thus helped determine their writing levels in the ACTFL guidelines. In order to increase the inter-reliability of the result, I asked two Chinese speaking colleagues in the field of second language learning to help evaluate the writing samples. When the results of the two graders' evaluations were not consistent, I acted as the third grader to determine the writing level of the learners.

Pair interviews. A total of two interviews were conducted toward the end of the school semester in May 2008. Each of the interviews lasted 90 minutes and was conducted with the participants from the same class. The intention of pairing up the participants from the same class is based on the belief that in a strategy use interview, interviewees in small groups tend to build on the response provided by others by adding strategies of their own (O'Malley & Chamot, 1990). In the semi-structured pair interviews, the results of the SILL survey were presented to the participants. In reviewing the results, the participants were given the opportunity to elaborate and explain the strategies they tended to use, and to describe how effective they thought these strategies were. The interviews were open-ended and sought to elicit any information relevant to the study.

Classroom observations. I observed the classes 50 minutes each time, three times a week over a 10-week period from February to April in 2008. The observation intended to capture the learning context of the four learner participants in order to give readers rich background of this study. Specifically, when I observed the classes, I attempted to document Professor Le's and Dr. Fu's weekly teaching routines in my notes. All teaching themes were identified and coded using Nvivo software.

Findings

Weekly Routines in the Classrooms

The classroom observation data indicated that the two instructors had different emphases in teaching Chinese. Professor Le emphasized vocabulary, reading, and verbal translation in each lesson. For example, she frequently asked her learners to read the vocabulary and the textbook aloud after her. In addition, she wrote new words on the board, explained the meaning of each character or word, and placed each new character or word in different phrases or sentences orally. Next,

she asked learners to translate verbally between English and Chinese by asking the “how do you say....in Chinese/English?” questions. On the other hand, Dr. Fu's class was highly structured in which every day in a week had specific tasks. For example, Dr. Fu began with a quiz over the previous segment on Wednesday and focused on the new segment's grammar and new vocabulary, which included character learning. On Fridays, the class continued with vocabulary drills, and practiced the newly learned language structure and vocabulary in games and in role-playing. On Mondays the class continued with games and role-playing exercises with a focus on practicing and memorizing the textbook dialogue. Dr. Fu discussed the translation exercises they did before class every Tuesday.

The types of language tasks practiced in each of the focal classrooms are introduced in Table 1. This table illustrates the two instructors' teaching routines from February to April in 2008, including number of language learning tasks and total minutes observed for each task.

Table 1

Language Learning Tasks in Professor Le's and Dr. Fu's Class

Professor Le's Instruction			Dr. Fu's Instruction		
Language Tasks	# of Tasks Observed	Total Minutes	Language Tasks	# of Tasks Observed	Total Minutes
Vocabulary	21	433	Structure and Vocabulary	18	413
Reading	13	236	Games and Role-Playing	7	152
Conversation Practice	5	87	Dialogue Practice	9	194
Tasks categorized not	Not Applicable	394	Translation Exercise	9	352
All tasks		1150	Tasks categorized not	Not Applicable	439
			All tasks		1550

The tasks not categorized in each class include quizzes, exams, student presentations, video watching, and any other learning tasks which were not regularly practiced in class. For example, Dr. Fu's class took weekly quizzes which usually lasted 20 to 30 minutes. The total number of minutes for weekly quizzes is included in the "tasks not categorized" column.

Learners' Writing Achievement

The learners' writings were sampled by asking them to complete a writing assignment toward the end of spring 2008. The learners had one week to complete the writing assignment. The writing assignment consisted of two questions. The first question asked the learners to write a paragraph using at least five of the 10 vocabulary words listed. The 10 words included were basic pronouns, verbs, time expressions, and adjectives which the

learners were taught in the classes. The second question asked the learners to write a letter introducing themselves to a Chinese pen pal. The directions specifically instructed the learners to describe personal preferences and daily routines.

In order to increase the reliability of the results, I asked two Mandarin Chinese speaking colleagues who received higher education in the Mandarin Chinese-speaking country, and have received professional research and teaching training in the field of second language learning, to evaluate the writing samples using the ACTFL guidelines. When the results of the two graders' evaluations were not consistent, I acted as the third grader to determine the writing level of the learners. The learners' writing levels would be the level identified by at least two of the three graders. Table 2 shows the evaluation from each grader and the final writing level of each learner.

Table 2
Learners' Writing Levels

Name	Grader 1	Grader 2	Grader 3	Final
Jenny (non-CHLL) (Fu's class)	Intermediate-High	Advanced	Intermediate-High	Intermediate-High
Yaoming (CHLL) (Fu's class)	Intermediate-Mid	Intermediate-High	Intermediate-High	Intermediate-High
Tim (non-CHLL) (Le's class)	Intermediate-Low	Novice-High	Novice-High	Novice-High
Luke (CHLL) (Le's class)	Intermediate-Mid	Intermediate-High	Intermediate-Mid	Intermediate-Mid

Table 2 illustrates that even though the evaluation for each learner from Graders 1 and 2 are not consistent, the levels identified are always back to back (e.g. intermediate-

high vs. advanced for Jenny, or intermediate-low vs. novice-high for Tim).

Learners' Writing Samples

The learners' writing samples are shown along with the evaluations from the three graders, which are summarized below.

Jenny - intermediate high level. Jenny's writing level was determined to be

at the intermediate high level according to the ACTFL guidelines. Jenny's writing for the first question illustrated that she was able to correctly use the time expression words to express time and tenses (See Figure 2).

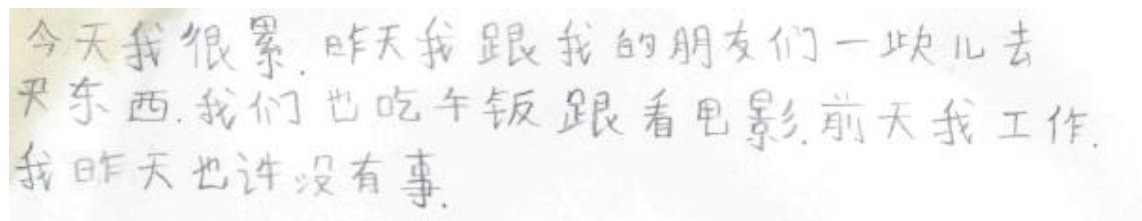


Figure 2. Jenny's writing sample 1.

Translation: [*I am very tired today. I went with my friends to go shopping yesterday. We also ate lunch and saw a movie. I worked the day before yesterday. I might be free yesterday.*]

In the paragraph, Jenny used the time expression words, "today," "yesterday," and "day before yesterday" to describe events in each day. Although the last sentence, "I

might be free yesterday" was not semantically correct, and Jenny made writing errors on two characters, 块 and 许, the rest of the paragraph was comprehensible to readers.

Figure 3 is Jenny's answer to writing question 2, writing a letter to introduce oneself to a pen pal.

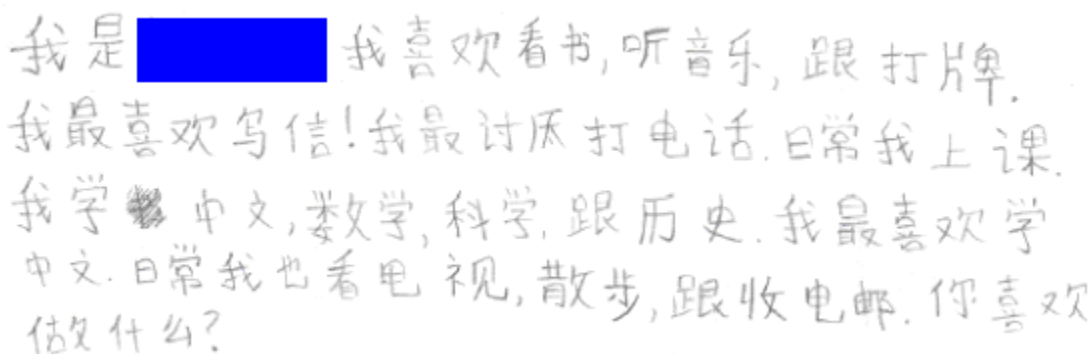


Figure 3. Jenny's writing sample 2.

Translation: [*I am Jenny. I like to read, listen to music, and play cards. I like to write letters the most! I hate making phone call the most. I go to school every day. I learn Chinese, math, science, and history. I like to learn Chinese the most. I also watch*

TV, take a walk, and receive emails every day. What do you like to do?]

Jenny briefly expressed her likes and dislikes, and her school experience in the paragraph. She knew how to use different

verbs such as *to read*, *to play*, *to write*, *to make a phone call*, and nouns such as *music*, *card*, and *letter* to indicate her likes and dislikes. In addition, she was able to use the word, 最 (/zui/the most) to indicate the level of preferences and dislikes. The paragraph has conscious organization in which Jenny first talked about her likes and dislikes followed by her school experience. At the

我常常买东西。
今天我买苹果。
明天我吃苹果，苹果很好。
我不有苹果。

Figure 4. Tim's writing sample 1.

Translation: [I often go shopping. I buy apples today. I eat apples tomorrow, apples are very good. I no have apples.]

Tim tried to describe what he did yesterday and today. He correctly used the time expression words, *today* and *tomorrow*. However, his vocabulary seemed limited. Throughout the paragraph, the only object he described was *apple* and the only adjective he used was *very good*. In addition, Tim's character writing

end of the paragraph, Jenny posed a question in order to find out her pen pal's likes and dislikes. Jenny's letter shows that her ability to narrate in paragraphs was emerging.

Tim - novice high level. Tim's writing was determined to be at the novice high level. Figure 4 shows Tim's writing for writing question 1.

performance was inconsistent. For example, the word, *apple* written in the last sentence was partially wrong, while the same word in the other sentences were correct. Finally, Tim used the grammatically incorrect word, 不有 (/bu you/no have) instead of 没有 (/mei you/do not have) to express that he did not have apples in the last sentence. Tim's letter to a pen pal shows the same writing pattern (See Figure 5).

很好，我姓 [redacted]，叫 [redacted]
星期，我去十点汉语上课，
四十点经济上课，
五十点化学上课。
星期二，我去七点数学上课，
四十点文化上课。

Figure 5. Tim's writing sample 2.

Translation: [Very good. My surname is Smith, and my full name is Tim Smith. Monday, I go 11 the Chinese to class, At 40 o'clock, to the Economics class, At 50 o'clock, to the Chemistry class. Tuesday, I go at 7 o'clock Math to class, At 40 o'clock to the Culture class.]

Tim tried to write 你好 (/ni hao/how do you do?) to greet his pen pal at the beginning of the

星期二	七点	我	去上	数学课
Tuesday	7 o'clock	I	go to	the math class
(Date)	(Time)	(Subject)	(Verb)	(Object)

Sometimes Tim omitted the first character of the verb 去 in 去上 (/qu shang/, to go) and misplaced the second character of the verb 上 in the object. He also misplaced the time word and the verb. For example, he wrote, 星期二我去七点.... (/xing qi er wo qu qi dian/Tuesday I go to 7 o'clock...) Instead of 星期二七点我去... /xing qi er xi dian wo qu/ (Tuesday at 7 o'clock I go...). The earlier sentence is considered grammatically incorrect in Chinese. In Chinese, the time expression words have to be written together either before or after the subject and should not be separated. Finally, Tim did not

letter, but it seems that he mistakenly wrote 很好 (/hen hao/very good) instead. He was able to use the sentence structure learned in class, "My surname is _____, and my full name is _____." in the next sentence. In the rest of the paragraph, Tim tried to describe the times and dates, and the classes he went to; however, he failed to correctly use the sentence structure to describe them. The correct sentence structure should be:

successfully indicate the times of the class. He wrote 40 o'clock and 50 o'clock, times that do not exist. Tim was not able to compose simple sentences and was determined to be at the phrase writing level, which is identified as the novice-high level in ACTFL.

Yaoming – intermediate high level. Yaoming's writing assignment was considered at the intermediate high level. In the first writing question, Yaoming explained why he decided to buy Chinese food (See Figure 6).

今天，我覺得很餓，我要吃法國飯，可是太貴了。
明天好一點兒，我有十塊錢。現在我買中國飯。

Figure 6. Yaoming's writing sample 1.

Translation: [Today I feel very hungry. I love to eat French food, but it is too expensive. It will be better tomorrow. I have 10 dollars. I buy Chinese food now.]

The content of Yaoming's paragraph was well-organized. He first expressed that he felt hungry and his preference of having French food. Later he explained the reason he decided to buy Chinese instead of French

food. Yaoming used three time expression words: today, tomorrow, and now. However, his statement about being better tomorrow in the paragraph was not clear as to what it

referred to. Besides the odd sentence, Yaoming had enough vocabulary and language structure knowledge to make the paragraph comprehensible to his readers.

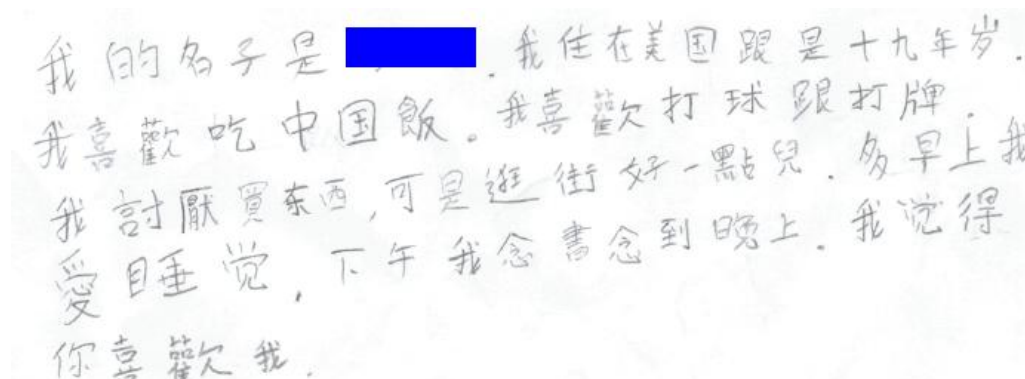


Figure 7. Yaoming's writing sample 2.

Translation: [My name is Yaoming. I live in the U.S. and am 19 years old. I like to eat Chinese food. I like to play ball and cards. I hate shopping, but going window-shopping is better. I love to sleep in the morning. In the afternoon I study until the evening. I feel that you like me.]

In Figure 7, Yaoming's letter provided information about his name, age, home country, his likes and dislikes, and his daily routine. Occasionally, he made character writing and grammar errors. For example, he mistakenly wrote the character, 子 instead of 字 in the word 名字 (/ming zi/name) in the first sentence. In regard to grammar, he incorrectly used the conjunction word, 跟 (/gen/and) in the second sentence to connect two clauses while the conjunction is only used to link nouns. The structure of the sentence with the misused conjunction word appeared to resemble literal translations from English to Chinese. Yaoming was able

to successfully describe activities he liked and disliked in the next few sentences. In terms of describing daily routine, Yaoming described the events he did in the morning, the afternoon, and the evening. He used the character, 多 (/duo/more), in front of the sentence, *I love to sleep in the morning*, which created confusion as to what information he tried to add to the sentence. On the other hand, Yaoming was able to use the "...until..." structure, in the sentence, 下午我念书念到晚上 (/xia wu wo nian shu nian dao wan shang/In the afternoon, I study until the evening), which is usually taught in the second year Chinese class. Overall, Yaoming was able to produce Chinese writing that is mostly cohesive and comprehensible.

Luke – intermediate mid level. Luke's writing assignment was considered to be at the intermediate mid-level. Figure 8 shows Luke's answer for the first writing question.

今天我去商场买饼。我很喜欢吃面。我买了以后我回家。明天我有课。

Figure 8. Luke's writing sample 1.

Translation: [I went to the market to buy (?). I like to eat noodles very much. After I bought it, I went home. Tomorrow I have a class.]

Luke described series of events he did and will do. He correctly used three time expression words: today, after, and tomorrow. The content of the paragraph followed a time sequence; however, Luke made a couple errors. He invented the last character in the first sentence, and created confusion as to what object he indicated buying in the market. Moreover, in the sentence, 我买了以后我回家 (/wo mai le yi

hou wo hui jia/After I bought it, I went home), the second 我 (I) in the sentence is usually unnecessary and omitted in Chinese. Despite the minor errors, Luke's paragraph was organized and showed his emerging ability to narrate events.

Figure 9 illustrates Luke's letter to his pen pal. He knew to use the letter format, having the greeting word, 你好 (/ni hao/how do you do) in the first line. He provided information in regard to his name, school name, class information, his likes and dislikes, his hometown, and his desired occupation.

你好,
我的名字是 [redacted]。我上的大学叫 [redacted]。
我每天上课。我有五节课。
我喜欢生物学。我要当医生。在美国用功很重要。我喜欢打球。我从 Ningbo 来 [redacted] 所以我喜欢吃海鲜。
你喜欢甚么?

Figure 9. Luke's writing sample 2.

Translation: [How do you do. My name is Luke. The University I go is University A. I go to class everyday. I have 5 classes. I like biology. I want to be a doctor. In the U.S., sports are very important. I like to play balls. I am from Ningbo, so I like to eat sea food. What do you like?]

The information Luke provided in the paragraph was clearly written. He was able to correctly use different verbs and nouns to describe the aforementioned information. However, the paragraph as a whole did not seem cohesive. For example, he wrote that he liked Biology and that he wanted to be a doctor. The relation between

Biology and doctor was implied, but not explicitly linked in the paragraph. Luke also made character writing errors. For example, He wrote 用动 instead of 运动 (/yun dong/to exercise) and 五们课 instead of 五门课 (/wu men ke/ five classes). The character writing errors and the loosely connected ideas

illustrated that Luke's writing is at the intermediate-mid level.

The Relationship between Learners' Strategy Use and Writing Achievement

Table 3 summarizes strategy use reported by the learners in the SILL survey.

Table 3
Learners' Strategy Use Reported in the SILL Survey

Categories of Strategies	Jenny (Mean)	Tim (Mean)	Yaoming (Mean)	Luke (Mean)
Memory	3.07	3.33	3.07	3.07
Cognitive	2.92	2.92	3.44	3.36
Compensation	2.75	2.88	3.88	4.63
Metacognitive	3.31	2.88	3.69	3.50
Affective	2.29	2.71	2.43	2.71
Social	3.00	2.92	2.78	4.67
Total Average	2.89	2.94	3.22	3.66

Jenny and Yaoming, who had the same writing level, appeared to use different types of strategies. Jenny tended to use Metacognitive strategies ($M = 3.31$) more while Yaoming tended to use compensation ($M = 3.88$) and Metacognitive strategies ($M = 3.69$). Luke, whose writing was at the intermediate-mid level tended to use social ($M = 4.67$) and compensation ($M = 4.63$) strategies. Tim, who had the novice-high writing level tended to use memory ($M = 3.33$) strategies more. In regard to the average use of all six categories of strategies, Jenny ($M = 2.89$) and Tim ($M = 2.94$) used the strategies less frequently in general compared to Yaoming ($M = 3.22$) and Luke ($M = 3.66$). The SILL survey provided the general idea of the learners' categories of strategy use. However, it is difficult to determine if there is a relationship between their strategy use

reported and their writing achievement as not all the strategies are relevant for learning Chinese writing. This study takes on another method to investigate the potential relationship between strategy use and writing achievement. In the following paragraphs, the learners provided in-depth information in the interviews about the strategies they used that may have helped increase their writing achievement and their confidence in writing Chinese.

Learner interviews. When asked how Chinese writing was practiced, both Jenny and Yaoming mentioned character writing and responded that they used the repeating strategy. They both would write things over and over again. Jenny indicated using another strategy, the associating strategy. She would make association between characters and things that look

similar to the characters. These strategies might have increased Yaoming's and Jenny's knowledge of Chinese vocabulary and writing levels. Both Yaoming's and Jenny's writing samples showed that they used a fair number of vocabulary words and made only minor character writing errors that would not affect readers' comprehension of the writings. Another aspect of writing is sentence structure. In Dr. Fu's class, the learners did weekly translation exercises that emphasized on the sentence structures and vocabulary taught in class. Yaoming and Jenny both agreed that the weekly translation exercises were helpful in terms of developing the knowledge of sentence structures. Yaoming stated,

I use it (the workbook) a lot for the translation exercise. It gives you a general idea of the sentence structures. They give a lot of examples too.

The translation exercises also helped the development of Chinese grammar. Jenny said,

It kind of forces you to try to figure out how to work things out, the grammar works, so it helps you understand why that isn't the way you thought it was. Like by doing it, you see a mental note of what's going on.

In evaluating Yaoming's and Jenny's sentence structures used in their writing samples, the graders found that the majority of the structures used were correct with only a couple words misplaced. Furthermore, I noticed that the sentence structures they used were practiced in the translation exercises. Hence, there seemed to be certain level of influence of strategy use on the learners' writing achievement.

In the interview with Luke and Tim, I found that neither of them was confident about Chinese writing. Tim stated,

Writing is by far the hardest for me. Going from English to Pinyin or Pinyin to English is easy, but going from English to character I think it's the hardest thing for me to do. Pinyin to character is hard too.

Tim expressed that Chinese writing was difficult for him; however, he did not spend time practicing writing to improve it. He said,

I practice reading a little bit, especially when I know there is a quiz or tests. I will practice before those. But I don't practice writing that much and it's really hard to write characters for me. But I am also really bad at writing in general.

Luke seemed to be in a similar study habit in terms of Chinese writing. Luke stated,

I don't really practice writing that much cause...yeah I don't practice writing at all. It's confidential right? Yeah I don't practice writing. I only write when Professor Le tells me I have an assignment to do.

Luke expressed that Chinese writing was difficulty for him compared to speaking when asked about the difficulty of the Chinese course he was taking. Luke explained,

Sometimes it (the class) can get really easy, sometimes it can get really hard. For me like speaking, it's like, she (Professor Le) asks how do you say high school? I already know it, so in that aspect it's very easy. But then she will be like how do you write 体育馆 (*gym*), ok, I have no idea. Reading and writing are super hard for me. I don't know why.

Tim and Luke did not seem to have specific strategies in mind for practicing writing. As a result, they did not spend any time practicing writing after class. When the learners tried to find the reasons they encountered difficulties in Chinese writing,

Tim blamed his generally bad writing skills and Luke expressed that he did not know the reason. The classroom observation data informed me that Professor Le only stressed the teaching of writing at the character level. She would show students how to write characters, but the rest of the class time would be spent repeating vocabulary or phrases orally. The effect of not having any strategies specific for practicing writing has shown in Tim's and Luke's writing. Both of them had lower writing levels compared to Jenny and Yaoming at Triangle University. In particular, Tim's writing used limited vocabulary and had few and repeated sentence structures. Luke's writing was more advanced compared to that of Tim's; however, the organization of his writing was weaker compared to that of Jenny's and Yaoming's.

Based on the learners' writing levels and their strategy use, it seems that there was a relationship between their strategy use and writing achievement. Jenny and Yaoming's use of the repeating strategy to practice character writing and the translating strategy to practice using different sentence structures taught in class seemed to help increase their writing ability. On the other hand, Tim and Luke, who did not use any specific strategy to practice writing in or after class seemed to have lower writing achievement.

Conclusion

The study found the influence of strategy use on the learners' writing achievement based on the classroom observations, the SILL survey results, and learners' writing samples and interview data. Dr. Fu's learners, Jenny (non-CHLL) and Yaoming (CHLL), who used the *repeating* and *translating* strategies for writing, tended to be at higher writing level compared to Professor Le's learners, Tim (non-CHLL) and Luke (CHLL), who were not able to

identify specific strategies to practice writing. This finding confirms Ellis's (1994) strategy framework in which learners' choice of strategies influence learners' rate of acquisition and language achievement. In addition, Professor Le's learners, Luke and Tim, who expressed that they did not use specific strategy to practice writing, did not seem to have a clue as to why they felt Chinese writing was difficult. According to Ellis's (1994) strategy framework, "the success that learners experience and their level of L2 proficiency can also affect their choice of strategies" (p. 529). In other words, the influence goes both ways between strategy use and L2 achievement. In this study, since Luke and Tim never had to write in Chinese except for character writing practices, they did not have opportunities to try any strategies for writing. Their lack of successful writing experience in Chinese could not help them confirm effective writing strategies which lead to a high level of writing achievement. On the other hand, Dr. Fu's learners, Yaoming and Jenny, who achieved higher writing proficiency, confirmed the effective use of the writing strategies adopted from the weekly translation exercises.

One other possible reason that Jenny and Yaoming wrote better in Chinese than Tim and Luke could be attributed to the type of institution where they studied. Jenny and Yaoming were in a private institution where the environment, teaching style, and learning resource distribution may be different compared to those in a public institution where Tim and Luke studied. In other words, factors other than the writing strategies taught by the instructors may have influenced the participants' writing strategy use and their writing achievement. Drawing upon Ellis's (1994) strategy framework in which both individual and social factors may affect a learner's choice of strategies and language achievement, the current study is

limited in terms of the number of factors (teaching environment) investigated and the type of language achievement (writing samples) analyzed. Future research is needed to find more factors which explain language learners' strategy use and confirm the relationship between strategy use and language achievement in all four language skills. Nonetheless, the finding about Jenny and Yaoming using specific writing strategies learned in class to reach high writing achievement implies that language instructors need to help develop learners' four language skills equally and to explicitly teach learning strategies that would help develop the language skills. In other words, language instructors should help learners identify effective strategy use in learning each of the four language skills and encourage them to try combinations of effective strategies. Successful or failed learning experiences resulting from experimenting with the use of different strategies will help direct learners toward finding the most effective and suitable strategies for themselves.

References

- Andreou, E., Andreou, G., & Vlachos, F. (2004). Studying orientations and performance on verbal fluency tasks in a second language. *Learning and Individual Differences, 15*, 23-33.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford, England: Oxford University Press.
- Grenfell, M. & Macaro, E. (2007). Language learner strategies: Claims and critiques. In A.D. Cohen & E. Macaro (Eds.), *Language Learner strategies: 30 years of research and practice* (pp. 9-28). Oxford, England: Oxford University Press.
- O'Malley, J. M. & Chamot, A. U. (1990). *Learning Strategies in Second Language Acquisition*. Cambridge, England: Cambridge University Press.
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. New York, NY: Newbury House Publishers.
- Peacock, M. & Ho, B. (2003). Student language learning strategies across eight disciplines. *International Journal of Applied Linguistics, 13*(2), 179-200.
- Poole, A. (2005). Gender differences in reading strategy among ESL college students. *Journal of College Reading and Learning, 36*(1), 7-20.
- Shen, H. (2004). Level of cognitive processing: Effects on character learning among non-native learners of Chinese as a foreign language. *Language and Education, 18*(2), 167-182.
- Shmais, W. (2003). Language learning strategy in Palestine. *Teaching English as a Second or Foreign Language-EJ, 7*(2).
- Wang, S. (1998). A study on the learning and teaching of Hanzi-Chinese characters. *Working Papers in Educational Linguistics, 14*(1), 69-101.
- Yongqi, P. (2005). *Vocabulary Learning Strategies in the Chinese EFL Context*. New York, NY: Marshall Cavendish Academic.