

Reforming Curriculum in A Centralized System:  
An Examination of the Relationships between Teacher Implementation of Student-Centered  
Pedagogy and High Stakes Teacher Evaluation Policies in China

Mei Luo

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy  
under the Executive Committee  
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

© 2012

Mei Luo

All rights reserved

## ABSTRACT

### Reforming Curriculum in a Centralized System:

An Examination of the Relationships between Teacher Implementation of Student-Centered Pedagogy and High Stakes Teacher Evaluation Policies in China

Mei Luo

Past research has shown poor implementation levels of a classroom-level curriculum strategy that forms the centerpiece of China's new national education reform policies, namely, *Student-Centered Pedagogy* (SCP). This dissertation set out to investigate the influences of selected school and teacher background characteristics, classroom-level SCP implementation variables, and a high stakes teacher evaluation policy variable, on SCP implementation levels reported by high school teachers in a selected school district in China.

The overall aim of the research was to study relationships among a number of factors hypothesized to affect teachers' SCP implementation levels, guided by a theoretically-grounded, conceptual framework. The study particularly examined the potential adverse influence of an output-driven teacher evaluation policy on SCP implementation levels. The teacher evaluation policy is tied to secondary school students' performance on the high stakes, national college entrance examination in China, the *Gaokao*.

Eight contextual and reform-related factors derived from a review of literature were tied together in the conceptual framework suggesting direct, mediating and moderating influences on SCP implementation. Based on the framework, paths by which the variables could affect SCP implementation levels directly or indirectly, were tested in stages.

Data were collected and analyzed using survey research methodology. The first part of the analyses involved the design and validation of a bilingual teacher survey (English and Chinese), tapping the key variables. The second part of the analyses involved a series of hierarchical regression models to test hypothesized pathways and relationships among the measured variables.

The theoretical premise of the study was that the large size and highly centralized structure of the Chinese educational system led it to adopt an output-control mechanism in the form of the high-stakes teacher evaluation policy tied to student performance on *Gaokao*. The adoption of such an output-control mechanism resulted in a mismatch between the philosophy underlying the newer SCP reforms and the pre-existing teacher evaluation policy, which in turn led to poor implementation levels of SCP in classrooms by teachers. Previously, researchers in China have overlooked the importance of policy incompatibility issues in examining effects of reforms at the classroom level.

The study found that, consistent with the literature, *teacher beliefs in SCP* and *teacher self-efficacy in practicing SCP* had consistently positive, statistically

significant influences on *SCP implementation* (for Beliefs in SCP,  $t(224)=3.745$ ,  $p=.000$ , standardized  $\beta=.22$ ; for Self Efficacy,  $t(224)=3.387$ ,  $p=.001$ , standardized  $\beta=.23$ ). Also consistent with expectations, the influence of the survey measure tapping perceived *control by the output-driven teacher evaluation policy* on SCP implementation, was negative and statistically significant ( $t(224)=-1.982$ ,  $p=.049$ ; standardized  $\beta=-.12$ ).

*Perceived support* for SCP implementation, including resources, professional development programs, support from principals and colleagues was a statistically significant predictor in initial models, but the factor was found to lose statistical significance when combined with the variable tapping perceived *control by the output-driven teacher evaluation policy*. With all the specified independent and mediating variables in the regression model, the cumulative variance explained on SCP Implementation levels was 20% ( $R^2 = .199$ ). The overall model was statistically significant ( $F[7,224]=7.935$ ,  $p=.000$ ). Together, these results confirmed the main hypotheses of the study.

Contrary to the literature, an omnibus *school* factor and individual *teacher background characteristics* (Gender, Teaching Experience, and Educational Degree) were not found to be statistically significant predictors of SCP implementation levels. Furthermore, the moderating effects of Grade level and Class size were not found to be statistically significant either.

Policy implications of the results for China are discussed, along with limitations and contributions to theory on educational reforms. Recommendations are made for future research.

## TABLE OF CONTENTS

|  |    |
|--|----|
| CHAPTER I INTRODUCTION.....  | 1  |
| Aims of the Study.....   | 1  |
| Research Context: China’s New Curriculum Reforms .....   | 4  |
| Government Support for Reform Implementation.....  | 5  |
| Implementation Problems with SCP.....  | 6  |
| Statement of the Problem: Gaps in the Existing Literature .....  | 7  |
| Bridging the Gaps.....   | 7  |
| Theoretical Justification .....  | 8  |
| Relevance of Organizational Theory in Studying China’s Educational Reforms.....                                | 8  |
| Policy Mismatches.....   | 10 |
| Parallel Evidence from the U.S. Educational Reforms Context.....   | 11 |
| Making a Case for the Present Study.....   | 13 |
| Research Objectives and Questions.....   | 14 |
| Objectives.....  | 14 |
| Specific Research Questions.....   | 15 |
| Introduction and Justification of Research Methodology.....  | 16 |
| Significance of the Study.....   | 17 |
| Current Reform Policy and Practices in China.....  | 17 |
| Research Base on Educational Reforms.....  | 17 |
| Definition of Terms.....   | 18 |
| Organization of the Study.....   | 20 |
| CHAPTER II REVIEW OF THE LITERATURE.....   | 21 |
| New Curriculum Reforms in China and Gaokao-based Evaluation Policies: Policy Incompatibility and Tensions..... | 22 |
| New Curriculum Reforms in China.....   | 22 |

|  |    |
|--|----|
| Test-related Teacher Evaluation Policies of the Chinese Educational System.....  | 25 |
| A Regional Case: Teacher Evaluation Policies in Jingyang District.....   | 27 |
| Definitions of Student-Centered Pedagogy: Consistency in China and the West...   | 29 |
| How China Defines Student-Centered Pedagogy.....   | 29 |
| Other Definitions of Student-Centered Pedagogy.....  | 30 |
| Organizational Theory, Centralized Structures and Output Control Mechanisms in<br>Large Bureaucracies: Interpreting China's Teacher Evaluation Policy..... | 34 |
| Control Mechanisms from Organizational Theory.....   | 34 |
| Degree of Centralization.....  | 35 |
| Output Controls Evidenced in China's Secondary Education Bureaucracy.....  | 37 |
| Factors Affecting Classroom Reform Implementation levels in Large Education<br>Systems: Empirical Research on Reforms in China.....                        | 38 |
| Factors Affecting Classroom Reform Implementation Levels in Large Education<br>Systems: Lessons from the U.S.....  | 39 |
| Factors at the School Level.....   | 39 |
| Factors at the Classroom Level: Teacher Background Characteristics.....  | 40 |
| Teacher experience.....  | 40 |
| Gender.....  | 41 |
| Educational degree.....  | 41 |
| Factors at the Classroom Level: Specific SCP-Relevant Variables.....   | 42 |
| Self-perceived knowledge and capacity for reforms.....   | 42 |
| Professional development.....  | 43 |
| Teacher attitudes, values, and self-efficacy.....  | 44 |
| Resources.....   | 45 |
| Social supports and network.....   | 47 |
| High stakes testing, accountability system and SCP practices.....  | 48 |
| Systems-based and Multivariate Influence.....  | 51 |



|   |    |
|---|----|
| Summary of the Review of Literature: Constructs, Variables, and A Conceptual Framework..... | 52 |
| Constructs and Variables.....   | 52 |
| Conceptual Framework.....   | 56 |
| Research Questions and Hypotheses.....  | 58 |
| CHAPTER III METHODOLOGY.....  | 64 |
| Research Site.....  | 64 |
| Sampling Design.....  | 65 |
| Unit of Analysis.....   | 65 |
| Sample Size and Statistical Power Estimation.....   | 65 |
| Random Sampling Procedure.....  | 66 |
| Data Collection Procedure.....  | 66 |
| Sample Composition and Representativeness.....  | 67 |
| Sample Representativeness on Gender and School Membership.....                              | 69 |
| Instrument Design and Validation.....   | 71 |
| Specification of Survey Domains.....  | 72 |
| Domain 1.0: Perceived support for implementing SCP.....                                     | 73 |
| Domain 2.0: Teachers' beliefs in SCP.....   | 73 |
| Domain 3.0: Teacher reports of self-efficacy in practicing SCP.....                         | 74 |
| Domain 4.0: Perceived control by the teacher evaluation policy.....                         | 74 |
| Domain 5.0: Teacher reports of SCP implementation.....                                      | 74 |
| Pre-Pilot Content Validation of Items.....  | 75 |
| Back-translation Method for Designing a Bilingual Survey.....                               | 75 |
| Pilot Study.....  | 76 |
| Instrument Refinement Following Pilot.....  | 78 |
| Empirical Validation of Final Instrument.....   | 78 |
| Exploratory Factor Analyses (All Items).....  | 79 |

|   |     |
|---|-----|
| Exploratory Factor Analyses (Positively-oriented Items).....                              | 80  |
| Final Scales: Descriptive Statistics and Reliability Estimates.....                       | 87  |
| Revised Conceptual Framework Based on EFA Results.....                                    | 88  |
| Data Analysis Plan.....   | 90  |
| Differences by Grade and Class Size.....  | 90  |
| Bivariate Correlation among Survey Measures.....  | 90  |
| Coding Procedure for the School Variable.....   | 90  |
| Hierarchical Multiple Regression Models.....  | 91  |
| Testing Statistical Assumptions in the Data Set.....                                      | 97  |
| CHAPTER IV RESULTS.....   | 98  |
| Grade Level Differences on Survey Construct Measures.....                                 | 98  |
| Class Size Differences on Survey Construct Measures.....                                  | 102 |
| Pearson Correlations.....   | 106 |
| Influence of School Characteristics on SCP Implementation by Teachers.....                | 109 |
| Influence of Teacher Background Characteristics on SCP Implementation .....               | 111 |
| Influence of Classroom Level SCP-relevant Variables on SCP Implementation ....            | 113 |
| Influence of Perceived Control by Teacher Evaluation Policy on SCP<br>Implementation..... | 116 |
| The Moderating Influences of Grade Level and Class Size.....                              | 119 |
| Summary of Significant Variable Relationships.....  | 123 |
| CHAPTER V DISCUSSION.....   | 125 |
| The Validated Conceptual Framework.....   | 125 |
| Discussion of Results by Research Questions and Hypothesis.....                           | 129 |
| Results on Research Questions 1.0-2.0.....  | 129 |
| Results on Research Question 3.0.....   | 131 |
| Results on Research Question 4.0.....   | 133 |
| Results on Research Question 5.0.....   | 134 |

|   |     |
|---|-----|
| Implications of Results for Policy, Practice and Theory.....  | 137 |
| Policy Levers in the Context of China’s Education Reforms.....  | 137 |
| Recommendations for Change.....   | 139 |
| Reforming the high-Stakes test, <i>Gaokao</i> .....   | 140 |
| Towards decentralization.....   | 141 |
| Contributions to Theory and Research Base on China’s New Curriculum Reforms.                                  | 142 |
| Implications for Reform Policy and Theory in U.S. Education.....  | 144 |
| Limitations of the Study.....   | 144 |
| Generalizability.....   | 144 |
| Measurement of Variables.....   | 145 |
| Analytic Models and Multiple Regression Procedure.....  | 147 |
| Suggestions for Future Study.....   | 148 |
| REFERENCES.....   | 150 |
| APPENDICES.....   | 167 |
| <i>Appendix A:</i> Excerpts of the Teacher Evaluation Policy by Jingyang District,<br>Deyang City, China..... | 167 |
| <i>Appendix B:</i> Domain Specifications and Indicators for the Teacher Survey<br>Questionnaire.....          | 174 |
| <i>Appendix C:</i> Item-Total Statistics for Three Domains (the Pilot Version).....                           | 183 |
| <i>Appendix D:</i> The Teacher Survey Questionnaire (Post-Pilot Version, in English).                         | 187 |
| <i>Appendix E:</i> The Teacher Survey Questionnaire (Post-Pilot Version, in Chinese).                         | 195 |
| <i>Appendix F:</i> Examination of Model Fitness of Regressions: Casewise<br>Diagnostics for Outliers.....     | 203 |
| <i>Appendix G:</i> Examinations of Basic Assumptions of Multiple-Regression<br>Models.....                    | 205 |

## LIST OF TABLES

|   |     |
|---|-----|
| Table 1. <i>A Contrast between the Student-Centered and Teacher-Centered Pedagogy: Classroom Characteristics under the Two Conditions</i> .....     | 31  |
| Table 2. <i>Constructs and Variables for Study</i> .....  | 54  |
| Table 3. <i>Background Characteristics of Sample on Education, Grade and Subject Taught</i> .....   | 68  |
| Table 4. <i>Background Characteristics of Sample on Teaching Experience and Class Size</i> .....  | 69  |
| Table 5. <i>A Comparison between the Population and the Sample: Gender</i> .....  | 70  |
| Table 6. <i>A Comparison between the Population and the Sample: School of Teaching</i> .....  | 71  |
| Table 7. <i>Descriptive Statistics and Reliability Coefficient of the Three Domains in the First Iteration</i> .....                                | 77  |
| Table 8. <i>Percentage of Variance Explained for Extracted Factors in Exploratory Factor Analysis</i> .....   | 80  |
| Table 9. <i>Results of Promax-Rotated Factor Pattern Showing Standardized Loading and Structure Coefficients</i> .....                              | 82  |
| Table 10. <i>Interfactor Correlation Matrix</i> .....   | 85  |
| Table 11. <i>Modified Domain based on Factor Structure: Items Defining Scales</i> .....   | 87  |
| Table 12. <i>Descriptive Statistics and Cronbach's Alpha Estimates for Seven Survey Scales</i> .....  | 88  |
| Table 13. <i>Grade Level Differences: Perceived Support from Colleagues and Professional Development Programs</i> .....                             | 99  |
| Table 14. <i>Grade Level Differences: General Beliefs Regarding SCP</i> .....   | 99  |
| Table 15. <i>Grade level Differences: Beliefs Regarding Teachers' Role in SCP</i> .....   | 100 |
| Table 16. <i>Grade level Differences: Self-Efficacy in Practicing SCP</i> .....   | 100 |
| Table 17. <i>Grade level Differences: Perceived Control by the Output-based Components of Teacher Evaluation Policy (Gaokao-related)</i> .....      | 101 |
| Table 18. <i>Grade level Differences: Perceived Control by the Process-based Components of Teacher Evaluation Policy (Non-Gaokao-related)</i> ..... | 101 |
| Table 19. <i>Grade level Differences: Teachers' Implementation of SCP</i> .....   | 102 |

|           |  |     |
|-----------|--|-----|
| Table 20. | <i>Class Size Differences: Perceived Support from Colleagues and Professional Development Programs</i> .....                         | 103 |
| Table 21. | <i>Class Size Differences: General Beliefs Regarding SCP</i> .....   | 103 |
| Table 22. | <i>Class Size Differences: Beliefs Regarding Teacher’s Role in SCP</i> .....   | 104 |
| Table 23. | <i>Class Size Differences: Self-efficacy in Practicing SCP</i> .....   | 104 |
| Table 24. | <i>Class Size Differences: Perceived Control by the Output-based Components of the Evaluation Policy (Gaokao-related)</i> .....      | 105 |
| Table 25. | <i>Class Size Differences: Perceived Control by the Process-based Components of the Evaluation Policy (Non-Gaokao-related)</i> ..... | 105 |
| Table 26. | <i>Class Size Differences: Teachers’ Implementation of SCP</i> .....   | 106 |
| Table 27. | <i>Pearson-Correlation Matrix</i> .....  | 108 |
| Table 28. | <i>Teachers’ Reports of SCP Implementation Regressed on School Context</i> ...   | 111 |
| Table 29. | <i>Teachers’ Reports of SCP Implementation Regressed on Teacher Background Characteristics</i> .....                                 | 113 |
| Table 30. | <i>Teachers’ Reports of SCP Implementation Regressed on SCP-relevant Variables</i> .....   | 115 |
| Table 31. | <i>Teachers’ Reports of SCP Implementation Regressed on Perceived Control by Teacher Evaluation Policy</i> .....                     | 118 |
| Table 32. | <i>Moderating Effects of Grade Level</i> .....   | 121 |
| Table 33. | <i>Moderating Effects of Class Size</i> .....  | 123 |

## LIST OF FIGURES

|            |   |     |
|------------|---|-----|
| Figure 1.0 | Conceptual Framework of the Study.....  | 57  |
| Figure 1.1 | Excerpt of Figure 1.0 Relevant to Hypothesis 1.0.....   | 58  |
| Figure 1.2 | Excerpt of Figure 1.0 Relevant to Hypothesis 2.0.....   | 59  |
| Figure 1.3 | Excerpt of Figure 1.0 Relevant to Hypothesis 3.0.....   | 60  |
| Figure 1.4 | Excerpt of Figure 1.0 Relevant to Hypothesis 4.0.....   | 61  |
| Figure 1.5 | Excerpt of Figure 1.0 Relevant to Hypothesis 5.0 and Hypothesis 6.0...                                    | 63  |
| Figure 2.0 | Revised Conceptual Framework with Variables Tapped by Scales.....   | 89  |
| Figure 3.0 | The Validated Conceptual Framework and Statistically Significant<br>Variable Relationships and Paths..... | 129 |

## ACKNOWLEDGEMENTS

Support from many individuals enabled me to complete this dissertation. I would like to thank members of my dissertation committee, Madhabi Chatterji, Luis Huerta, Craig Richards, Francisco Rivera-Batiz, and Denise Burnette for their guidance, support, and honest criticism throughout my years at Teachers College, which were fundamental to my work. Throughout my dissertation research, Professor Chatterji's willingness to serve as the Sponsor and ask challenging questions kept my work grounded and meaningful. Her careful attention to the rigor of my methodological approach and analysis pushed and enhanced my work and I thank her for her candor, patience, and responsiveness. Professor Huerta and Professor Batiz's inspiration and guidance helped me to press on. Professor Richards and Professor Burnette's constant encouragement gave me the courage to carry on whenever encountered difficulties. I thank all the committee members for their enthusiasm around my work, for their efforts, and for their belief that I could get this done.

No words are sufficient to express my appreciation to my parents. They were beacons of light throughout this process and I owe them a tremendous amount. Their love made all the difference in helping me to attain this. Their constant care to the progresses I made had helped me get through the darkest times.

Finally, my heart and sincere appreciation goes out to my husband and my soul mate, Ling Zhou. You have always been there no matter I needed you or not. Your endless inspiration, keen insight, and unfaltering support over the past thirteen years have meant more than you'll ever know. "Thank you" is too pale to express my feeling at this

moment. Your patience and sense of humor, your encouragement and positive attitude, your understanding and persistent support.....above all, your unconditional love and belief in me has been with me every step of the road, which sometimes was seemingly so impossible to pursue. I want you to know, like thirteen years ago when you entered into my life, you are a magical rainbow; because of you, without knowing it, I am vaporized by the happiness you bring me. Love forever!



REFORMING CURRICULUM IN A CENTRALIZED SYSTEM:  
AN EXAMINATION OF THE RELATIONSHIPS BETWEEN TEACHER  
IMPLEMENTATION OF STUDENT-CENTERED PEDAGOGY AND HIGH STAKES  
TEACHER EVALUATION POLICIES IN CHINA

**CHAPTER I INTRODUCTION**

**Aims of the Study**

The aim of the present research was to study factors affecting the implementation of a classroom-level curriculum strategy that forms the centerpiece of China's new curriculum reforms, referred to as *Student-Centered Pedagogy* (SCP) (Zhang, 2008; Zhong, 2008). In the context of China's education system and new reform policies, SCP implies that students are encouraged to be independent-minded knowledge seekers, developing personal meaning about the physical world around them through direct experience and dialogue with teachers and others in their educational environment (Zhang, 2008; Zhong, 2008). China's conceptualization of SCP draws largely on a social constructivist perspective found in the Western literature (Deboer, 2002; Piaget, 1963). SCP is deemed critical to the success of the curriculum reforms in China since the approach touches upon the technical core of the educational process, or teaching and learning in the classroom (Zhang, 2008; Zhu, 2008). Despite calls for more use of SCP, however, Chinese scholars have noted poor implementation levels of SCP strategies in China's schools and classrooms (Gao & Deng, 2008; Liu, 2011).

The study asserted that a main barrier to successful implementation of SCP in secondary schools in China is the incompatibility between the two national-level policies that are currently in effect today. One policy is used for evaluating teachers and requires competitive student performance on the *Gaokao*, the national college entrance examination in China. In contrast, the new policy calls for an emphasis on SCP during teaching (see *Guidelines for New Curriculum Reforms* [the Guideline hereafter], The State Council of People's Republic of China, 2001).

China's secondary education system has long placed a heavy emphasis on the high stakes, National Higher Education Entrance Examination for students, the *Gaokao*. Schools and teachers are held accountable for annual student performance on *Gaokao*. Results on the *Gaokao*, also serve as a major indicator for conducting teacher performance reviews in several regional jurisdictions (Chen & Li, 2007; Han & Yang, 2008; Zhang, 2007; Zhao, 2007). An added aim of the present study was to examine the potential negative influence of this "output-control mechanism" (Ouchi, 1977, p. 97) on reported levels of classroom implementation of SCP by teachers.

Organizational theory suggests that very large organizations tend to develop highly centralized governance structures and adopt "output-control mechanisms" (Ouchi, 1977, p. 97) similar to the *Gaokao*-based staff performance evaluation policies prevalent in China today (Evans, 1975; Ouchi, 1977, 1978, 1979; Williamson, 1971). Borrowing from that literature, organizational "control" in this study refers to bureaucratic control, indicating the processes used for monitoring the work of employees with rules, policies, a hierarchy of authority, reward systems, and other formal mechanisms to manage member behavior and assess performance (Ouchi, 1977, 1978, 1979). In particular provinces and districts, student performance on the *Gaokao* is tied to specific teacher evaluation criteria and merit pay schedules. Given the

immense size of national and regional education systems in China, this study asserted that the Chinese educational system adopted its own version of an “output-control mechanism” (Ouchi, 1977, p. 97), characterized by the high-stakes teacher performance evaluation policies.

Parallels can be found in the history of standards-based reforms in the public education system in the United States (U.S.). School accountability policies associated with the standards-based reform movement in the U.S. were enforced through national legislative actions like No Child Left Behind Act of 2001 (20 U.S.C.A. § 6301 et seq., West 2003). According to some (Bakers & Richards, 2008; Meyer & Rowan, 2006; Spillane & Burch, 2006), there has been a movement towards greater centralization in public school systems of the U.S. since the 1960s. Many state-level public education systems in the U.S. today are large and hierarchical in organization. The current reform movement in education in the U.S., like China, has two features that seem to be in opposition. One is an emphasis on rigor and accountability through high-stakes testing of students, and evaluation of teachers and schools based on student test scores. The other is the effort to develop student-centered approaches to teaching and learning (Deboer, 2002).

This study attempted to examine empirically how selected, research-supported factors at the school and classroom levels affected secondary teachers’ implementation of SCP by taking a systems-based conceptual approach. As parts of a larger education system, classroom activities are influenced by multiple and interrelated forces. It is thus necessary to investigate relevant factors together by adopting a multivariate conceptual model. Classroom-level factors selected for study were based on an extensive review of literature that is elaborated in Chapter II of this dissertation. The output-control variable that is of key interest in this study, was operationally defined through an individual teacher’s lens, as perceived levels of control exercised by the prevailing teacher performance evaluation policy (Cooper, Slavin, & Madden, 1997). SCP,

implementation at the classroom level served as the dependent variable, or desired outcome, in the conceptual framework.

The first chapter frames the research study in the larger context of education reforms in China. The chapter provides a theoretical justification for the study, introduces the research objectives, questions and methods employed, and points to the significance of the study. To conclude, the chapter defines salient terms used in the work and describes how the remaining chapters are organized.

### **Research Context: China's New Curriculum Reforms**

In 2001, China initiated a series of centralized educational reforms to address new demands for public education that emerged with the evolution of the “information age” and knowledge-based economies around the world (Zhong and Yang, 2002, p. 14). The information age is characterized by the ability of individuals to transfer information freely, and to have instant access to information that would have been difficult or impossible to secure in earlier eras (Beniger, 1986). This shift from the traditional industrial models to economies based on the manipulation of information, according to Zhong and Yang (2002), resulted in new demands for education and life-long learning in China. As life-long learners, students must not only master basic learning tools, such as reading, writing, verbal communication, and problem solving; they must also develop moral values and a worldview that can adapt to an ever-changing modern society. China initiated its new curriculum reforms to help its populace keep up with rapid societal and global changes. The conventional methods of schooling were viewed to have failed in developing students that are ready to meet requirements of the new information age (Zhong & Yang, 2002).

Zhong and Yang (2002) summarized China's new curriculum reforms as having three major goals: a) to reform the outdated textbook system by including local governments and schools in decision making processes, b) to reconfigure the structure of outdated courses of study across all educational levels, and c) to reform teaching and learning processes by implementing SCP. Of the three, SCP was deemed to be the most critical component of China's curriculum reforms (Zhang, 2008; Zhu, 2008).

Through SCP policies, China's education leaders attempted to reform and phase out classroom pedagogical practices of teachers that were conventional and highly teacher-centered. More current, student-centered instructional approaches were now endorsed. According to Zhu (2008), the success of the new reforms hinged completely on whether SCP could be successfully implemented and sustained in classroom level activities. Otherwise, fundamental changes would fail to occur in the educational system as a whole.

### **Government Support for Reform Implementation**

Approaches to large-scale policy implementation in China have drawn on strategies learned from past experience in both the east and the west. Such strategies have focused on providing SCP-related professional development programs to teachers and on providing financial and material supports to schools to help bring about teaching changes. Muju Zhu (2004), the deputy director of the Basic Education Division in the Ministry of Education in China, stated recently that the financial resources devoted by the central government towards the new curriculum reforms reached 70 million Yuan in 2004. Because the central government is responsible for only a small share (around 12%) of all school expenditures (Lv & Pang, 2002), the amount of reform funding dedicated by local governments was estimated to be higher.

Reports through 2010 suggest that SCP-related professional development programs have reached approximately 92% of all elementary, middle-high, and high school teachers nationwide (Liu, 2011).

### **Implementation Problems with SCP**

Since SCP constitutes a fundamental part of China's curriculum reform agenda, much attention by the Chinese research community has been directed towards investigating the question as to whether teachers are authentically implementing SCP in their classrooms. To date, the research suggests that SCP-related policy efforts have failed to yield desired outcomes. Teacher practices related to SCP appear to be symbolic rather than actual, and little or no assimilation of SCP principles have occurred at the classroom level (Li, 2008; Ma & Tang, 2002; Xia, 2008; Yan & Zhou, 2008; Zhong, 2005b). Only 3.3% of teachers in a national news report were satisfied with the current status of SCP implementation in their classrooms (Liu, 2011).

A number of factors have been identified by Chinese researchers to explain the implementation problems related to SCP. Several relate directly to beliefs and behaviors of teachers in the classroom and those who work at the front lines of the educational system. For example, teachers' knowledge and beliefs regarding the new reform policies, perceived adequacy of professional development opportunities, perceived availability of resources, and their individual abilities to carry out reforms are some important factors that were found to affect implementation of educational reforms initiated at higher levels of the governmental hierarchy (Li, 2008; Ma & Tang, 2002; Yan & Zhou, 2008).

### **Statement of the Problem: Gaps in the Existing Literature**

Despite attempts at research, the existing literature from China provided limited explanations with regard to barriers and issues surrounding SCP implementation in China. If the efforts to solve the problems were on the right track, after over-10 years of implementation and repeated rounds of remedial action, classroom-level SCP practices should have become more widespread. There should have been evidence of substantive aspects of SCP implementation from classrooms, instead of symbolic SCP gestures (Li, 2008; Ma & Tang, 2002; Xia, 2008; Yan & Zhou, 2008; Zhong, 2005b)

The research base from China on SCP policy implementation was limited. First, a majority of articles reviewed were found to be limited mostly to conceptual discussions of SCP and lacking in empirical evidence. Second, potential causal or correlational factors were not discussed or examined in a systematic manner. Third, beyond the resource and professional development factors identified, Chinese researchers failed to locate or investigate other factors that could be at the root of the implementation problems. The limited explanatory power of the existing literature in China limits any possible follow-up policy actions for improving SCP practices. It also limits the value of the theoretical knowledge base on how classroom-level reforms can succeed in China.

### **Bridging the Gaps**

The present study set out to fill the above gaps. It hypothesized that, the root of the SCP implementation problems lay in the long-standing “output-control mechanism” (Ouchi, 1977, p. 97) of the Chinese educational system in the form of *Gaokao*-related staff evaluation policies. Adopting a systems-based approach, the study attempted to empirically examine the mediating

influences of the said control mechanisms on classroom-level implementation of SCP, along with a number of other relevant exogenous and mediating factors at the school and classroom levels that dealt with core teacher beliefs and behaviors surrounding SCP reforms .

The premise that output-control mechanisms inherent in the Chinese educational system could be one source of the SCP implementation problems is supported by views of some Chinese researchers as well. This perspective holds that since the two standardized tests mandated at the 9<sup>th</sup> grade (*Zhongkao*) the 12<sup>th</sup> grade (*Gaokao*) have led to a “teach-to-test” phenomenon (*Ying Shi Jiao Yu*) in China, the test-related policies present a principal barrier of classroom-level implementation of SCP (Gao & Deng, 2008). This view identifies the ostensible tensions between the standardized test and SCP-oriented instruction (Gao & Deng, 2008).

Standardized tests are, by themselves, neutral by nature; they are but tools designed to collect information on the status of student learning. How the test scores are used makes the testing and test-related policies controversial and a possible barrier to student-centered instruction. This study hypothesized that the main reason why *Gaokao* and *Zhongkao* have heavily influenced classroom activities and led to “teach-to-test” practices instead of SCP, is closely tied to the output-control mechanisms of the Chinese secondary educational system and teacher evaluation policies.

## **Theoretical Justification**

### **Relevance of Organizational Theory in Studying China’s Educational Reforms**

What are the main output control mechanisms of the Chinese educational system, and what can be learned from organizational theory to understand the potential effects of such mechanisms on reform implementation in a centralized system? Is there evidence from other



large education systems in the world that suggests similar policy implementation issues related to higher degrees of centralization and control? To answer these questions, this study drew on a few concepts of organizational theory on the structure of organizations, control mechanisms, evaluation policies of personnel (Evans, 1975; Ouchi, 1977, 1978, & 1979; Williamson, 1971). It also drew on the research base on reform implementation in large education systems like the U.S..

Organizational control refers to bureaucratic control used to influence member behaviors and assess performance (Ouchi, 1977). The control may be enforced through rules, policies, the hierarchy of authority, reward systems, and other formal mechanisms. Behavioral control is one of two forms of bureaucratic control mechanisms found in an organization (Ouchi, 1977, 1978, & 1979). It refers to the direct evaluation of production *process* behaviors. Output control is the second form of bureaucratic control mechanisms found in an organization (Ouchi, 1977, 1978, & 1979). Output control refers to an evaluation of the *results* of production process behaviors.

A related organizational concept in this study is *centralization*, which refers to the authority system found in organizations. In typical centralized systems, members are organized through multiple horizontal departments and vertical hierarchies. The power flow within a centralized organization is usually unilateral with policy-making located at the top of the organization's hierarchy. While the degree of top-down control can vary, highly centralized structures have a pyramid-shaped chain of authority (Bray, 2003). In the context of the Chinese educational system, this study claimed that the large and centralized structure is a primary determinant of the output control mechanisms found China's teacher evaluation policies.

As the literature review in Chapter II will show, due to its large and highly centralized structure and the readily available standardized testing systems at the existing grades (*Gaokao* at the 12<sup>th</sup> grade and *Zhongkao* at the 9<sup>th</sup> grade), the Chinese educational system has relied on output-driven teacher evaluation policies as a primary organizational control mechanism in K-12 education (Chen & Li, 2007; Han & Yang, 2008; Jiang, 2008; Zhang, 2007; Zhao, 2007). Such output-driven teacher evaluation policies are based on student test performance, which are treated as objective and reliable *output* indicators of teacher quality (Han & Yang, 2008; Jiang, 2008).

### **Policy Mismatches**

The observed contradiction between SCP and *Gaokao*-oriented instruction in China is, in fact, a manifestation of the contradiction between the philosophy and rationales underlying the two national policies highlighting SCP implementation versus system outputs. The contradiction reflects two layers of mismatch. The first is a mismatch of policy intentions: SCP reforms are intended to change the processes of teaching and learning whereas the output-control mechanisms of the older system emphasize solely educational outputs in terms of students gaining admission to colleges. The second is the mismatch of philosophy: SCP reforms encourage inquiry-based and open-ended teaching and learning approaches (Deboer, 2002; Grant & Hill, 2006; Knowlton, 2000; Passman, 2000; Pedersen & Liu, 2003; Piaget, 1963; Zhong, 2008), whereas the current teacher evaluation policies are tied to standardized tests of students' rote knowledge of the curriculum, as evidenced on national, college entrance examination scores (Chen & Li, 2007; Han & Yang, 2008; Jiang, 2008; Zhang, 2007; Zhao, 2007).

The *Gaokao* policy cultivates a high-stakes environment for schools and teachers that favors demonstrable outputs and rote knowledge accumulation over student-oriented teaching processes and open-ended inquiry. Since the new SCP reforms are tied to very low stakes with no consequences for teachers who do not follow the policies, the *Gaokao*-related output control mechanism takes precedence within the centralized structure of the system. To what extent this observed incompatibility between the new reform movement and the pre-existing organizational control mechanisms affects SCP implementation is still unknown and therefore warrants investigation.

### **Parallel Evidence from the U.S. Education Reforms Context**

The basic premises and need for the study are also supported by parallel evidence from the U.S. context. According to Deboer (2002), the current reform movement in education in the U.S. has two opposing features as well: the emphasis on rigor and accountability through high-stakes testing of students, and the effort to develop student-centered approaches to teaching and learning. As indicated, there has also been greater centralization of the system over time (Meyer & Rowan, 2006; Spillane & Burch, 2006).

The increase of state policy activity since the middle of last century (Baker & Richards, 2008) has gradually transformed the organizational structure of the U.S. educational system from a fragmented, decentralized state to one of a greater centralization (Meyer & Rowan, 2006; Spillane & Burch, 2006). Responding to this new organizational structure of greater centralization, the central authorities of the U.S. educational system have adopted various forms of output-control mechanisms, with increasing use of accountability-related student testing and school evaluation policies.

Today, the school and teacher accountability system that has evolved in the U.S. is also characterized by a push for utilization of students' standardized test data. The U.S. "output control mechanisms" (Ouchi, 1977, p. 97) are reflected in legislation like the federal No Child Left Behind [NCLB] Act of 2001 (20 U.S.C.A. § 6301 et seq., West 2003). This law stipulated that the evaluation of schools would be based on a rigid, top-down performance monitoring system where schools would be held to standards of Adequate Yearly Progress (AYP) tied to students' performance on state-endorsed standardized tests (Deboer, 2002; Meyer & Rowan, 2006; Ravitch, 2010; Spillane & Burch, 2006). Schools that failed to meet the AYP standards would face severe punitive consequences, such as, taking "corrective actions" (Ravitch, 2010, p. 97) if a school missed its targets for any subgroup for four consecutive years, and "restructuring" (Ravitch, 2010, p. 98), if it missed targets for five consecutive years. Corrective action indicates possible changes in curriculum, staff, or the length of school day or year. Teacher evaluation systems were also affected by NCLB policies in the U.S., with increased use of value added evaluation models that incorporated student test scores. Recently, the Obama administration signaled to American teachers that the new administration is open to the idea of linking teacher pay to student performance on state-administered standardized tests (Meckler, 2011).

Both observations and research evidence in the context of U.S. educational reforms suggest two contradictory lines of reform policy, just as in China. The first is the use of accountability-driven, high stakes evaluation policies. The second is an emphasis on student-centered teaching approaches at the classroom level (Deboer, 2002). According to some, teacher changes with regard to SCP-related reforms have been adversely affected by the high stakes personnel evaluation policies (Deboer, 2002; Meyer & Rowan, 2006; Passman, 2000; Pedersen & Liu, 2003; Spillane & Burch, 2006). Constraints on the curriculum and narrowing of teaching

to match the external tests have been widely documented to be detrimental (Nichols & Berliner, 2007; Passman, 2000; Pedersen & Liu, 2003; Watanabe, 2007).

The U.S. literature on SCP implementation in education also suggests that the degree of teacher practices differ significantly depending on grade levels that are the focus of high-stakes testing (Nichols, Glass, & Berliner, 2006). More student-oriented teaching is observed in grades without state-administered standardized tests. Teachers at the grade levels at which the test is given are particularly vulnerable to the pressures of “teaching to the test” (Amrein & Berliner, 2002).

Class size is also a factor potentially influencing variability of levels of teachers’ SCP-related practices. Studies focusing on investigating the effects of class size on teaching and learning in the U.S. showed that reduced class size significantly affected teaching methodologies (Molnar, Smith, Zahorik, Palmer, Halbach, & Ehrle, 1999). More SCP-oriented teaching behaviors appeared more often in smaller classes (Molnar et al., 1999). Based on these findings, it is logical to wonder if, in China’s case, teachers’ SCP practices are truly affected by the output-driven evaluation policies, and whether such a relationship is moderated by these two factors: Grade Level and Class Size.

### **Making a Case for the Present Study**

Classroom-level curriculum reforms in China are based on the assumption that teachers will be using SCP principles in their day-to-day practices (Zhong, 2008). However, China’s new curriculum reforms face similar reform implementation problems as those observed in the U.S. context. Because evidence from the U.S. shows that high-stakes, test-driven accountability measures can have adverse influences on the implementation of student-centered educational

approaches in classrooms, there is a need to study if and how teachers' SCP practices in the Chinese context are influenced by existing teacher evaluation policies. Unlike the U.S. educational system, where output-focused sanctions are currently placed merely at the school level, results of high-stakes standardized tests measuring student achievement are directly tied to individual teacher evaluations and merit pay in China (Chen & Li, 2007; Han & Yang, 2008; Jiang, 2008; Zhang, 2007; Zhao, 2007). Without a satisfactory examination of factors related to the observed policy contradictions in the Chinese education system, reforms calling for a full-scale implementation and institutionalization of SCP are likely to fail.

### **Research Objectives and Questions**

#### **Objectives**

Based on the above theoretical rationale, this study had three broad objectives:

- To examine to what extent the factors identified in the existing educational reforms literature at the school and the classroom levels affect teacher implementation of SCP in secondary schools in China;
- To examine empirically the basic premise of the study that controls enforced in a large and highly centralized Chinese educational system through the *Gaokao*-related teacher evaluation policies will adversely affect levels of classroom implementation of SCP by secondary teachers;
- To investigate whether the relationship between the system controls enforced by *Gaokao*-related teacher evaluation policies and secondary teacher implementation of SCP would be moderated by two factors, class size and grade level.

## Specific Research Questions

This study attempted to answer the following research questions:

1. To what extent do *school* characteristics, as predicted by educational reform and policy implementation literature, affect SCP implementation at the classroom level?
2. To what extent do selected *teacher characteristics*, as predicted by educational reform and policy implementation literature, affect SCP implementation in the classroom after taking school-level differences into account?
3. To what extent do selected SCP-relevant constructs (*teacher perceived support*, *teacher beliefs*, and *teacher self-efficacy in practicing SCP*), as predicted by the educational reform and policy implementation literature, affect SCP implementation in the classroom, accounting for school and teacher characteristics?
4. To what extent is *control* enforced by the organization's *Gaokao*-related teacher evaluation policy for secondary teachers, a negative predictor and a significant mediating variable for SCP implementation in classrooms?
5. To what extent is the relationship between *control* enforced by the organization's *Gaokao*-related teacher evaluation policy for secondary teachers and reported SCP implementation levels in the classroom, moderated by *class size* and *grade level*?

### **Introduction and Justification of Research Methodology**

The study was carried out using cross-sectional, survey research methodology where teachers served as the population units (Babbie, 1990). The sample consisted of 232 randomly selected high school teachers from a target population of Grade 10 to Grade 12 teachers at Jingyang District of the City of Deyang, Sichuan Province, China. Teachers in these higher grades were expected to be affected most by *Gaokao*-related evaluation policies, as students enter college after Grade 12.

The study was conducted in three stages. To start, a conceptual framework was proposed based on a comprehensive literature review given in Chapter II. The conceptual framework helped tie together all the selected variables and suggested pathways by which they would influence the dependent variable, SCP implementation in the classroom. Next, a multi-construct survey instrument was designed and validated using an iterative procedure (Chatterji, 2003). This work helped derive valid and reliable survey-based measures of the selected variables from teacher responses to the survey questionnaire. Finally, a series of hierarchical multiple regression models were run to test relational hypotheses designed to answer the research questions.

Survey methodology was deemed appropriate for the study as this approach permitted data-gathering directly from reform implementers at the classroom level. Surveys also constitute suitable instruments for tapping the constructs of teacher behavior, attitude and perception that were central to the purposes of the study (Babbie, 1990; Chatterji, 2003). A survey is also an efficient data-gathering tool for the limited time and resources that were available to the researcher for the study (Babbie, 1990; Robson, 2002). The methods are further detailed in Chapter III.



## Significance of the Study

### Current Reform Policy and Practices in China

Due to its long-standing dominance in the system, the potential adverse impact of the existing output-control mechanism was overlooked during the early stages of the new curriculum reforms in China. Questions were asked but quickly brushed aside, and attention focused mostly on providing resources for reforms. By examining a new construct, teacher perceptions of *control* of the *Gaokao*-related evaluation policy (the output-driven organizational control mechanism in the Chinese educational system), the study broke new ground. An investigation as to the potential impacts of the output-control mechanism itself could help inform policy makers in China about the root of the current SCP implementation problems. Policy implications of findings are discussed in Chapter V.

### Research Base on Educational Reforms

The results of the study, detailed in Chapter IV, add to the empirical research base on Chinese educational reforms. Most importantly, research efforts on SCP-related implementation have never looked at the issues from the angle of organizational control in centralized bureaucratic structures. Since the inception of the new curriculum reforms, there have been a plethora of articles dealing with factors influencing implementation outcomes from the perspective of implementers at the bottom of the organizational hierarchy in China (Li, 2008; Ma & Tang, 2002; Yan & Zhou, 2008). However, many of these studies were piece-meal research efforts, and results were contradictory and inconsistent (Ryan, 1996). In addition, most were conceptual discussions, lacking empirical evidence.

The present study and accompanying results inform the current literature by showing whether the factors identified are statistically significant and substantial predictors of SCP, when examined via a comprehensive conceptual model supported by the literature. The results delineate the relative importance of different factors on SCP implementation by teachers. This research looked particularly for the missing links in the empirical relationships among a regionally enforced evaluation policy, school factors, teacher and classroom factors, and teacher-reported SCP policy implementation in classrooms. Chapter V discusses next steps in research based on the detailed results shown in Chapter IV.

### **Definition of Terms**

Selected terms are now defined as used in this study to facilitate common interpretation.

***Constructivism***—Constructivism is a theory of knowledge that argues that humans generate knowledge and meaning about an object from social interactions with others and direct experience in relation to the object combined with their own ideas about the object (Bransford et al., 2000; Duffy & Jonassen, 1992; Piaget, 1963; Vygotskii, 1978)

***Student-Centered Pedagogy (SCP)***—SCP in this study indicates associations with constructivism and manifests itself in the active involvement of students in the teaching and learning process. A variety of instructional approaches fit beneath the umbrella of SCP, including case-based teaching/learning, project-based scenarios, and problem-based teaching/learning (Grant & Hill, 2006; Knowlton, 2000; Pedersen & Liu, 2003; Zhong, 2007, 2008).

**Organization**—In this study, the definition of an organization draws largely on the perspective of the classical organization theory on bureaucracy (Weber, 1947), in which an organization was broadly defined as a unit of people, structured by a well-defined line of authority, pursuing the desired collective goals by following the formal rules and regulations. The present study treated the entire Chinese educational system as one organization, in which members are structured in a pyramid-shaped authority hierarchy with teachers at the bottom; teachers' behaviors are monitored and adjusted via organizational control mechanisms (i.e., formal evaluation policies) to produce desired educational outcomes.

**Organizational Control**— In this study, organizational control refers to bureaucratic control. More specifically, it refers to the use of rules, policies, hierarchies of authority, reward systems, and other formal mechanisms to influence member behaviors and assess performance (Ouchi, 1977). In the context of the Chinese educational system, organizational control is mainly realized by means of staff evaluation policies tied to student performance on *Gaokao*.

**Behavior Control**—Behavioral control is one of the two bureaucratic control mechanisms found in an organization (Ouchi, 1977, 1978, & 1979). In this study, it refers to direct observation of teaching-related “processes” of teachers.

**Output Control**—Output control is the other bureaucratic control mechanism in an organization besides behavioral control (Ouchi, 1977, 1978, & 1979). In this study, output control refers to formal staff evaluations tied to performance of students on external examinations like the *Gaokao*.

**Centralization**—In this study, centralization refers to the abstract, pyramid-shaped authority system (Bray, 2003). In the centralized structure, members are organized through multiple

horizontal departments and vertical hierarchies. The power flow within a centralized organization is unilateral with policy-making located at the top of the organization's hierarchy. In the context of the Chinese educational system, the large and centralized structure is the primary determinant of the output control mechanisms (Ouchi, 1977).

### **Organization of the Study**

Chapter II provides a detailed discussion of a relevant body of literature and derives a conceptual framework and specific research hypotheses to guide the study and analysis plan. In Chapter III, the research methodology and procedures are described in more detail, including information on sample composition, the sampling design, results of the pilot study on the survey, final instrumentation and measures, and data analysis plan. Chapter IV presents results of both descriptive and relational analyses tied to hypotheses and research questions. Chapter V discusses the results of the study with reference to the originally proposed conceptual model and the implications of results for reform policy, theory, and research practice in China and elsewhere. Chapter V also presents a discussion of the limitations of the study and makes suggestions for future research.

## CHAPTER II

### REVIEW OF THE LITERATURE

This chapter discusses in detail the literature bases used to develop the conceptual framework that guided the overall study, along with the specific research questions and hypotheses tested. In separate sections, it discusses the literature on:

- China's new curriculum reforms and existing teacher evaluation policies tied to the *Gaokao*;
- Student Centered Pedagogy (SCP) as a new reform strategy in China and SCP definitions found in the literature;
- Organizational theory and control mechanisms found in large bureaucracies; and
- Factors documented to affect pedagogical reform implementation by teachers in large education systems of China and the U.S, where high-stakes testing and accountability systems are employed in grades Kindergarten-12.

At the end of the literature review, links are drawn between the different literature bases to present a comprehensive conceptual framework. Variables are selected for the investigation and the theoretical construct domains of the survey questionnaire are identified. The chapter concludes by discussing the path diagram showing variable relationships leading to the dependent variable, levels of SCP implementation reported by teachers. Six hypotheses, aligned with respective research questions and conceptual framework, are then presented.

## **New Curriculum Reforms in China and *Gaokao*-based Evaluation Policies:**

### **Policy Incompatibility and Tensions**

#### **New Curriculum Reforms in China**

In 2001, the Ministry of Education of the People's Republic of China issued a new education policy initiative, generally known as "new curriculum reforms" (The State Council of People's Republic of China, 2001, p. 1). The document detailing the initiative is titled "Guidelines to Curriculum Reforms in Basic Education (Guidelines here after)" (The State Council of People's Republic of China, 2001, p. 1). The reforms signified the eighth major overall effort of the Chinese education system since 1949 (Wu, 2005). According to discussions among Chinese scholars, the eighth reform effort was called upon to address the conflict between China's traditional methods of schooling and the new demands for public education stemming from the emergence of the "information era" (Zhong, 2005b; Zhong & Yang, 2002, p. 14).

In the context of Chinese education, traditional schooling meant three things. First, it implied the use of a rigid centralized textbook system, where textbooks were disseminated and administered by the national government. Recently, Zhong and Yang (2002) criticized traditional textbooks as being "difficult, over-complicated, irrelevant, and outdated" for local educators and students (p.14).

Second, traditional schooling implied an obsolete course system, in which numerous courses were narrow and single subject-oriented. Courses were found to be lacking comprehensiveness, especially at the elementary level, because they failed to draw cross-disciplinary connections on overlapping topics (Zhong & Yang, 2002). The new curriculum reforms attempted to fade the boundaries between disciplines by combining related content areas

together. For example, Music and Graphic Art were combined to establish a new course, titled Arts. Geography and History were also combined into a new course called History and Society. Another new comprehensive course at the middle-high level is Science, which is a combination of the previously existing Chemistry, Physics, and Biology courses (Li, 2005).

Third, traditional schooling implied use of a conventional pedagogical philosophy emphasizing the teacher's authoritative status while neglecting students' roles in the teaching and learning process. It further emphasized cognitive achievement while neglecting students' all-round development. The new curriculum reforms focus on amending the aforementioned conditions to make Chinese education more adaptive to the demands of a new information era.

As indicated earlier, Zhong and Yang (2002) identified three components of China's new curriculum reforms as textbook reforms, course reforms, and incorporation of SCP in classrooms. The first component—textbook reforms—signifies a transformation from a previous highly centralized national textbook system to a relatively decentralized one as both the city and county governments and schools are granted freedom to take actions towards meeting local needs (Zhong & Yang, 2002).

The course reforms were expected to reconfigure the structure of courses of study across all educational levels. Specifically, the Guidelines (The State Council of People's Republic of China, 2001) recommended that at the elementary level (from the 1<sup>st</sup> grade to the 6<sup>th</sup> grade), comprehensive courses should constitute the core of the curriculum; at the middle-high school level (from the 7<sup>th</sup> grade to the 9<sup>th</sup> grade), the curriculum should consist of both comprehensive and discipline-oriented courses; and at the high school level (from the 10<sup>th</sup> grade to the 12<sup>th</sup> grade), all courses are to be discipline-specific. In addition, throughout basic education, courses

aimed at improving students' competency and practical skills are required to be part of the core curriculum. Among the three components, however, SCP is considered to be most critical to implementing the new curriculum reforms (Zhu, 2008).

The third component of China's new reforms—SCP—attempts to reform classroom practices by phasing out conventional teacher-centered pedagogy and replacing that with more current approaches to instruction. SCP-related reforms emphasize that students should play a greater role in teaching and learning processes in the classroom. With regard to classroom instruction, the Guidelines highlight that “teachers should respect the dignity of students, pay attention to student differences, and meet their differentiated needs...so that every student can develop adequately” (The State Council of People's Republic of China, 2001, p. 2). With regard to the design of learning opportunities, the Guidelines underscore that “student should, under the guidance of the teacher, learn in an active and individualized fashion” (The State Council of People's Republic of China, 2001, p.3). The policy also accentuates fostering the student's capacity in identifying and solving problems.

According to Zhu (2008), the success of the entire curriculum reforms hinges on whether SCP is successfully implemented and sustained in classroom level activities. Overall, the new curriculum reforms set forth a “3-D (imensional)” (Zhong, 2005b, p. 18) goal with respect to the teaching and learning processes. According to Zhong (2005b), the 3-D goal suggests that teaching and learning should equally emphasize a) knowledge and skills, b) learning processes and learning methodology, and c) attitudes and values of life and the world. SCP-related reforms are the crucial means to achieving the 3-D goal (Zhu, 2008). In 2001, the new curriculum policies, including SCP, were tried out in thirty-eight special experimental districts scattered over ten provinces (Ma & Tang, 2002). In 2005, it was scaled up for nationwide implementation.



## **Test-related Teacher Evaluation Policies of the Chinese Educational System**

The Chinese educational system is characterized by a highly centralized structure (Cheng, 2001; Li & Xiao, 2001; Zheng, 2010). Centralization in this study refers to the vertical power transactions among different layers of authority, such as, the central government (federal), provincial governments (states), local governments (school districts), and schools. This is called “spatial centralization”(Bray, 2003, p. 22).

Although China has undertaken certain reforms that suggest decentralization in the past few decades, scholars assert that the Chinese educational system remains highly centralized (Hawkins, 2000; Zheng, 2010). The power transaction between the central and local governments in China has never reached the “devolution” level (Bray, 2003, p. 22). In other words, the central government reserves the ultimate authority and can terminate any powers delegated to lower units, as it sees fit (Cheng, 2001; Hawkins, 2000; Li & Xiao, 2001; Zheng, 2010).

Under this centralized structure, the Chinese educational system tightly controls schools and individual teachers by relying heavily on measures of performance control. In particular, the use of high-stakes teacher performance evaluation policies that tie to student outcomes on standardized tests, is a key mechanism employed to achieve teacher compliance (Chen & Li, 2007; Han & Yang, 2008; Jiang, 2008; Zhang, 2007; Zhao, 2007). Based on organizational theory on structure and control mechanisms (Ouchi, 1977, 1978, 1979), this form of control is more aligned with the notion of *output* (results-based outcome) rather than *behavior* (process) control.

China's reliance on educational outputs as a dominating mechanism of organizational control is echoed by observations of a number of Chinese researchers. For example, Jiang (2008) stated that teacher evaluation policies in China, in general, emphasize four aspects: professionalism, classroom competency, attendance regularity, and teaching outcomes based on student test results (e.g., scores of the *Gaokao* exams, admission rates to higher education). However, among the four, teaching performance outcomes based on test results by far outweigh the rest. Jiang (2008) and others (Chen & Li, 2007; Han & Yang, 2008; Zhang, 2007; Zhao, 2007) pointed out that the majority of schools in China treat the test scores as the most direct and objective indicators for evaluating job performance of teachers. Direct observations of teaching behaviors (e.g. planned or improvised class visits by school principal or administrators) are also included as part of the evaluation, but they remain nominal in weight (Chen & Li, 2007; Han & Yang, 2008; Zhang, 2007; Zhao, 2007).

Jiang (2008) also observed that the main reasons why schools reject methods of direct classroom observations is that they require too much time and resources, and the standards for good teaching have not been determined. This has led teachers to resist assessments based on direct observation. In teachers' view, Jiang (2008) observed, principals and administrators are laymen when it comes to instruction in specific subject areas. Therefore, there continues to be a reliance on *Zhongkao* and *Gaokao* assessments for 9<sup>th</sup> and 12<sup>th</sup> graders, respectively, to conduct teacher evaluations. Both are standardized tests. *Gaokao* is operated nation-wide whereas *Zhongkao* is standardized within each province.

## A Regional Case: Teacher Evaluation Policies in Jingyang District

How are the national teacher evaluation policies filtered down to the district level, in the region that is the focus of the present study? The policy document, entitled *Guideline of Personnel Policies*, was issued by the Human Resource Office in the Jingyang District Department of Education, Deyang, Sichuan. This document, written in Chinese, contains eight sections: *General Rules, Personnel Hiring, Working Code, Salary, Annual Evaluation, Rewards, Sanctions, and Appendix*. Since not all sections are relevant, only Sections 5 (*Annual Evaluation*) and 6 (*Rewards*) were translated and are reviewed here. For further details, please see *Appendix A*.

Excerpt 1 of *Appendix A* is a translation of Section 5. Although not exact, paraphrased contents of Excerpt 1 are consistent with general observations by a number of Chinese scholars that teacher evaluation policies in China generally emphasize four aspects: professionalism, classroom competency, attendance regularity, and teaching outcomes equated with student outputs on tests (Jiang, 2008). Section 5 also shows that the district's teacher evaluation policy encompasses three aspects: professionalism, teaching practice, and achievements in teaching and research. Section 5 provides rubrics for the annual evaluation of high school teachers in Jingyang District.

Excerpt 2 of *Appendix A* highlights that students' performance in *Gaokao* is directly tied to teachers' job performance evaluations and high stakes merit recognitions and punishments. Excerpt 2 is a translation of Section 6 and contains formulas used to calculate specific amounts of monetary rewards granted to teachers of the 12<sup>th</sup> grade graduating classes. The baseline for rewards is established based on the previous year's average college admission rate of the top

three high schools in the region. For each additional student admitted into higher education, the 12<sup>th</sup> grade teachers as a group get an additional ¥1000. For each additional student *below* the baseline, ¥1000 is deducted.

The formulas also stipulate that the teachers will be awarded additional ¥30000 if a top-ranking student at the provincial level is in their classes, ¥20000 for a student ranked at the second place, and ¥10000 for each student ranked from the third to the fifteenth places. In addition, for each student admitted into the top tier universities, the 12<sup>th</sup> grade teachers get another ¥5000.

Based on the formulas, a 12<sup>th</sup> grade teacher could receive approximately ¥5000 in total rewards if requirements were met. The amount would be much higher if her/his students ranked regionally or got admitted into prestigious universities. Compared to the average teacher salary in Jingyang District, which is ¥1170 monthly (official number retrieved from <http://health.scjg.com.cn/article.aspx?id=60647> ), ¥5000 is a significant and sufficient incentive for a teacher to employ a “teach-to-test” pedagogy or other strategies deemed helpful in raising student test scores.

The level of detail in formulas implies how much effort the district dedicated to create a system, and how much *Gaokao* is emphasized in the overall outcomes. Because *Gaokao* (and *Zhongkao*) are so tightly connected to student admissions and teacher evaluation policies in China, they are not only high-stakes tests for students, but also hold high-stakes for teachers.

## Definitions of Student-Centered Pedagogy: Consistency in China and the West

### How China Defines Student-Centered Pedagogy

The discussions regarding SCP in the *Guidelines* are consistent with the understanding and interpretation of SCP by other scholars who support the new curriculum reforms, led by Zhong (2005a, 2005b, 2007, 2008). Zhong and his team have substantially influenced the theoretical discussions and operational interpretations of SCP in China (Zhang, 2008). One of Zhong's 2007 publications, entitled *A Debate on Teaching and Learning Knowledge*, summarized his view regarding SCP as follows:

- First, students should be the center of the teaching and learning process. For teachers, the ultimate purpose of teaching is to stop teaching because teachers' responsibilities are to help student learn via discovery methods, and to assist them in developing a clear sense of what, why, and how to learn.
- Second, learning is not an isolated process; instead, it should happen in a collaborative environment so that social interaction, a key component of Constructivism, can be realized to facilitate learning.
- Third, learning is not about passively receiving facts and fixed knowledge from teachers; rather, learning occurs through the inquisitive seeking of knowledge.
- And finally, assessment is mainly a diagnostic means to better inform teaching and learning. Therefore, current practices in assessment should be changed from relying solely on pencil-paper testing to utilizing multiple assessment forms including direct

observations, record keepings, interviews, discussions, home assignments, projects, and portfolios, just to name a few.

### **Other Definitions of Student-Centered Pedagogy**

What are some other conceptual and operational definitions of student-centered pedagogy (SCP) in the existing educational literature? And what does SCP encompass with regard to instructional strategies when compared to teacher-centered pedagogy?

Stemming from democratic underpinnings (Dewey, 1938; Friere, 1970) and psychological bases of self-motivated learning (Deci & Ryan, 1987; Lepper & Green, 1978), teaching and learning using SCP draws largely on a social constructivist perspective, which holds that students develop personal meaning regarding the physical world through direct experience and dialogue with others about those experiences (Deboer, 2002; Piaget, 1963; Zhong, 2007, 2008). The new curriculum reforms in China reflect a perspective consistent with that of social constructivists (Zhang, 2008; Zhong, 2007, 2008).

Based on how teachers treat the subject matter, people involved, and educational processes at the classroom level, Knowlton (2000) characterized the pedagogical orientation to be either teacher-centered or student-centered. Table 1 is adapted from Knowlton (2000), and presents her contrasting views of these two types of pedagogy with respect to philosophical orientations and how topics are treated, roles people play, and processes of teaching and learning.

Table 1

*A Contrast between the Student-Centered and Teacher-Centered Pedagogy: Classroom**Characteristics under the Two Conditions.*

|  | Teacher-Centered Classroom   | Student-Centered Classroom  |
|--|--|---|
| Philosophy and Pedagogical Orientation | Positivistic (Belief that humans study, understand, and harness knowledge through objective inquiry.)                      | Constructivistic (Belief that knowledge is constructed through interactive, subjective human experience.)   |
| Treatment of topics or “Things”        | Teacher introduces “things” and suggests the implications of those things  | Both teacher and students introduce “things,” and both offer interpretations and implications   |
| Roles of People                        | Roles of teacher and student are regimented: the teacher disseminates knowledge, and the student receives that information | Role of teacher and student are dynamic: the teacher and students are a community of learners. The teacher serves as coach and mentor; the students become active participants in learning. |
| Processes                              | Teacher lectures while students take notes   | Teacher serves as facilitator while students collaborate with each other and the teacher to develop personal understanding of content.  |

*Note.* Adapted from “A theoretical framework for the online classroom: A defense and delineation of a student-centered pedagogy,” by D. Knowlton, 2000, *New Directions for Teaching and Learning*, 84, 5-14.

According to Donnell (1999), positivism views reality as independent of human consciousness, or as external, material, and objective. Because it is external to the observer, “reality” can be studied independently of the inquirer, whether they are teachers or students. An assumption is that different observers would arrive at the same conclusions, and that knowledge is defined by general and immutable laws which operate independently of individual observers

and observations. In contrast to the positivist view, constructivists deem reality as essentially subjective. There could be as many realities as there are people (Donnell, 1999).

As evident, Knowlton's model (2000) treats the two approaches as discrete and binary. This may be an over-simplified conceptualization. In real classroom settings, student-centered pedagogy and teacher-centered pedagogy are more likely to be the two ends of a continuum (Passman, 2000).

A number of other researchers offer alternative and more concrete interpretations with respect to the transformation processes needed for a shift from teacher-centered pedagogy to student-centered pedagogy (Grant & Hill, 2006; Passman, 2000; Pedersen & Liu, 2003; Zhong, 2007, 2008). Passman (2000) laid out an instructional SCP model that is better operationalized for the classroom in the U.S. public education system. Her model suggests a continuum from less teacher-directed practice to more student-centered practice. Passman's model is delineated below:

- Less whole class instruction including lecturing and teacher-led discussions, and more time spent in group and individual inquiry discussions.
- Less seatwork such as worksheets, dittos, workbook, and other "make work", and more reliance on student focused inquiry within an integrated curriculum approach.
- Less time spent by students reading text books and basal readers, and more time spent reading authentic literature from trade books.



- Less emphasis on content coverage where large quantities of material is introduced and memorized for later evaluation, and more time spent in learning to understand the content being learned.
- Less time spent in enforced silence, and more time spent in active learning, which may be noisy.
- Less emphasis on ability grouping and pull-out programs that tend to separate students from their peers, and more emphasis on heterogeneous grouping and inclusion programs.
- Less reliance on standardized testing and published assessment programs, and more reliance on portfolio assessment that is developmentally appropriate including teacher assessment (Passman, 2000, pp.5-6)

Resonating the above model, Grant and Hill (2006) identified additional differences in other aspects of teacher-centered and student-centered practice. One is associated with new roles and responsibilities. Compared to the teacher-centered approach, a teacher's role is de-centered in student-centered classrooms. As with Knowlton (2000), Grant and Hill (2006) also state that SCP-oriented teachers usually serve as facilitators of learning and partners in learning with students. New instructional strategies call on teachers being skillful in group-, time-, and project-management, instead of competency in lecturing alone. Also, assessments in SCP settings are of multiple forms, a majority of which could be performance-based and collaborative.

The present study used Zhong's theoretical discussion of SCP (2007, 2008) along with Passman's (2000) and Grant and Hill's (2006) operationalized models as the basis for understanding the technical core of SCP and constructing the survey-based construct measures.

## **Organizational Theory, Centralized Structures and Output Control Mechanisms in Large Bureaucracies: Interpreting China's Teacher Evaluation Policy**

### **Control Mechanisms from Organizational Theory**

Starting from the 1970s, organizational theory scholars began to differentiate organizational control from structure, and contended that organizational structure influences production activities in a system through control mechanisms that are devised by policy-makers and leaders (Evans, 1975; Ouchi, 1977; Williamson, 1971). This school of thought conceptualized *control* as a process of monitoring the work of members in a system, comparing it with some pre-set standards, and then providing selective rewards or adjustments to their performance so that individual production would move toward the direction of collective goals set by the organization's leaders (Ouchi, 1977). From this theoretical perspective, as indicated earlier, organizational control is realized primarily through the means of formal personnel evaluation measures and policies.

Control mechanisms utilized by organizations can be of two types: behavior-control mechanisms, which involve evaluating production *process behaviors* of personnel, and output-control mechanisms, which involve evaluating the *outcomes* of production process behaviors (Ouchi, 1977). Although most organizations can be expected to use a mix of these two forms of control, the particular mechanism of control, according to Ouchi's research (1977, 1978, 1979), is dependent on organizational structure. As organizations grow larger, the number of levels in

the hierarchy increase, and more horizontal differentiation occurs, compounding problems with regard to control. This complexity often results in a loss of control (Williamson, 1971; Evans, 1975). To deal with possible loss of control, large and complex organizations tend to employ output-control mechanisms.

Two preconditions accompanying structure determine which control mechanism an organization would employ to assess the work of its personnel: a) whether the organization is clear about how inputs get transformed into desired outcomes, and b) whether a measure of the desired outputs is available (Ouchi, 1977). For large and complex organizations such as educational systems where the educational transformation process for students is usually unclear (Hess, 1999), if a valid and reliable measure for outputs is not available, some other form of ritualized control occurs giving an illusion of rational analysis and evaluation (Meyer & Rowan, 1978, 1983; Ouchi, 1977). However, when tools for measurement for outputs are readily available, output-control mechanisms tend to take precedence with use of the existing tools (Ouchi, 1977).

### **Degree of Centralization**

The extent of centralization in an organization depends on the extent of decentralization. According to Bray (2003), spatial decentralization can be of three types based on the degree of decentralization: de-concentration, delegation, and devolution. *De-concentration* is the process whereby the central authority establishes branches or functional departments with its own staff, which can either operate out of headquarters or be dispatched to local areas. *Delegation* allows lower units to enjoy greater decision making power; however, the central government is only *lending* the power to local authorities in these cases. The delegated power can be retracted

whenever the local unit is deemed to be in violation of central authority's trust. *Devolution* is the highest level of decentralization where the local units execute power with far greater levels of autonomy. The central authority functions mostly as a hub for information exchange.

Although China has undertaken certain reforms that suggest decentralization in the past few decades, scholars assert that the Chinese educational system remains highly centralized (Hawkins, 2000; Zheng, 2010). The power transaction between the central and local governments in China has never reached the "devolution" level (Bray, 2003, p. 22). In other words, the central government reserves the ultimate authority and can terminate any powers delegated to lower units, as it sees fit (Cheng, 2001; Hawkins, 2000; Li & Xiao, 2001; Zheng, 2010). Zheng (2010) pointed out, regardless of administrative authority delegated to local governments, the central government never let go of the powers of personnel and school evaluation, including hiring and placement of staff cadres in the system to achieve their goals. According to Zheng (2010), evaluation of staff cadres is the main leverage of the central government, intended to ensure central control over local governments, while delegating only the administrative and resource distribution responsibilities to the lower levels.

Because of the highly centralized structure, the authority flow within the Chinese educational system is thus unilateral: from the top to the bottom. In other words, provincial governments are held accountable based on the performance of the pertaining local governments; local governments are held accountable based on the performance of the schools under their jurisdiction; and schools are held accountable based on the performance of their teachers.

## Output Controls Evidenced in China's Secondary Education Bureaucracy

Based on the relationships discussed thus far between organizational structure, ready availability of tools for exercising control, and the large size and hierarchy in the Chinese education system (Ouchi, 1977, 1978, 1979; Ouchi & Maguire, 1975), the predominant control mechanism at the secondary level appears to be one of *output-control* by means of the *Gaokao*. This characterization is justified on three grounds:

- a) the ambiguity (and immeasurability) of the transformation processes by which educational inputs from students and schools yield the desired outcomes of schooling (Hess, 1999);
- b) the easy availability of *output* measurements via the standardized testing systems at both the middle and high school levels that are accepted by schools and the public (Chen & Li, 2007; Han & Yang, 2008); and
- c) the *output-control* mechanism is viewed as more helpful by organizational authorities, compared to *behavior-control* mechanisms, to cope with the problem of control loss (Jiang, 2008; Ouchi, 1977).

In systems characterized by higher degrees of centralization, large size, high levels of differentiation, and multiple hierarchies, high stakes output-control mechanisms are common elsewhere, as well. Chapter I referred to similar evidence from the U.S. public education context, where the systems have shifted progressively from decentralized to more centralized structures, with high stakes testing, teacher and school evaluation policies governed by national legislation (see the No Child Left Behind Act of 2001, 2003). The next section identifies research-

supported variables known to affect large scale reform implementation in classrooms of big systems, and counterforces of high stakes testing policies from the U.S. context.

### **Factors Affecting Classroom Reform Implementation Levels in Large Education Systems: Empirical Research on Reforms in China**

As indicated in Chapter I, research on educational reforms and policy implementation in China are mostly conceptual discussions with little empirical evidence available offered. The few empirical studies conducted so far suffer from both internal and external validity problems.

For example, a survey research was conducted by Hu, Han, Wen, and Li in 2005, sponsored by the Department of Education in Shanxi Province, investigating the status of SCP implementation in secondary schools within the province. The study utilized the stratified random sampling procedure. However, only one school was chosen to represent all schools in each stratum. No rationale was provided in regard to why one school is sufficient to represent the stratum and why it is selected. In addition, causal or relational inferences were drawn based on mere univariate descriptive statistics such as means, percentages and so on. No advanced statistical methods were involved to discern random error from real relationship between variables.

Other Chinese studies were found to suffer similar methodological problems (see Li & Wang, 2008; Wang, Zhao, Duan, & Wang, 2007). Due to the lack of valid empirical findings among Chinese literature, the next section will focus primarily on educational reforms and policy implementation literature generated in the U.S.

## **Factors Affecting Classroom Reform Implementation Levels in Large Education Systems: Lessons from the U.S.**

### **Factors at the School Level**

Researchers agree that educational reform implementation comprises “co-constructed processes” (Datnow, Hubbard, & Mehan, 1998, p. 7). Accordingly to Datnow, Hubbard, and Mehan (1998), the idea of co-construction indicates a conditional matrix of activities necessary for reform implementation. Forces situated in classrooms, schools, districts, layers of the governmental hierarchy, and the surrounding environment all interact. Schools influence classroom activities in the sense that they provide the infrastructure, resources and immediate conditions for teaching and learning.

As part of the larger system (Starbuck, 1976), schools also reflect the demands experienced from the upper levels of the hierarchy and from the surrounding socio-economic, socio-cultural community and socio-political environment. For large-scale reforms initiated from the top, differences in school leadership style and interpretation of the new reform policies have been shown to be very important factors shaping reform implementation and outcomes in classrooms.

Studies that explored the principal’s leadership style and a school’s implementation of reform policies initiated at the top, adopted a variety of methods. Some used in-depth case study methods while others employed cross-sectional surveys. Although the specific type of reform varied from study to study, major findings appeared to be consistent. Researchers concluded that the strategies adopted by a school in instituting changes desired by reforms and the resulting variation of pedagogical practices are strongly dependent upon the school leaders’ vision,

understanding, and interpretation of the role and impact of the reform in the curriculum, the school's goals, as well as its history, culture and its general vision and mission (Datnow & Castellano, 2001; Leithwood & Jantzi, 2006; Stalling & Mohlman, 1981; Yuen, Law, & Wong, 2003). School-level factors such as, school climate and surrounding socio-cultural environment, also matter (Bulach & Malone, 1994). Other findings included the following (Stalling & Mohlman, 1981):

- a) in schools where reform policies were clear and more consistently enforced, more teachers changed their classroom behavior toward the reform;
- b) in schools where the principal was more collaborative and respectful, teachers had higher morale for reforms;
- c) in schools with more supportive principals, more teacher implemented the reforms.

### **Factors at the Classroom Level: Teacher Background Characteristics**

Teacher characteristics (e.g. educational degree obtained, gender, and experience) can also influence compliance with a new policy, especially when the new policy calls for substantive, innovative changes (Afshari, Baker, Luan, samash, & Fooi, 2009; Rogers, 1995; Schiller, 2003). Some specific teacher characteristics highlighted in the literature follow.

**Teacher experience.** Relevant experience in the new area can determine the extent of adoption of a new policy/program. A report by the National Center for Education Statistics in 2000 on technology reforms indicated that teachers with fewer years of experience were more likely to use computers in their classes than teachers with more years of experience. This may be



due to the fact that new teachers have been exposed to computers during their training and therefore, have more experience using this tool.

Hargreaves (2005), found a similar relationship between the level of teachers' behavioral changes with respect to a new policy implementation. Drawing on an analysis of interviews with 50 Canadian elementary, middle and high school teachers, Hargreaves (2005) found that teachers with more experience (measured by total years in teaching) are prone to being indifferent in learning new things. To mid-career teachers, older colleagues do not have the energy levels needed to deal with change that they find as "just too much work" (p. 979).

**Gender.** Gender is another teacher characteristic that may affect policy and program implementation. Venkatesh and Morris (2000) investigated gender differences in the context of individual adoption and sustained usage of technology reforms in the workplace. They studied on user reactions and technology usage behavior among 355 workers who were introduced to a new software technology application over a 5-month period. The results showed that men and women employ very different decision processes in evaluating new technologies. Particularly, women were more strongly influenced by subjective norms and perceived behavioral control by others.

**Educational degree.** Educational degree obtained by teachers, along with teacher certification and level of experience, has been treated as proxy of teacher quality in the sphere of education (Smith & Desimone, 2005). The assumption behind this connection is that teachers with higher educational attainment can provide more scholarly instruction and presumably, possess more insights about what is good or right for the student. Although this assumption was questioned by some researchers, the study does show that preparedness in certain subjects (measured by educational degree) such as mathematics and participation in content-related

professional development activities are associated with increased use of reform-oriented teaching strategies (Smith & Desimone, 2005).

### **Factors at the Classroom Level: Specific SCP-Relevant Variables**

Pressman and Wildasvasky (1973) showed that implementation is not mindless compliance to a mandate or policy, but that success of implementation ultimately depends on many semiautonomous agencies. Change ultimately is a problem of the smallest unit or teacher (McLaughlin, 1991; Odden, 1991; Weatherley & Lipsky, 1977).

**Self-perceived knowledge and capacity for reforms.** In her comprehensive literature review, entitled *Implementation Research in Education*, McLaughlin (2006) pointed out that implementation research generally focused on two themes: the technical properties of policy and individuals' ability to carry it out. In other words, teachers must know what is to be accomplished and by what means.

Implementers' knowledge of the reforms and skills to execute the reforms are the two most prominent factors that affect the implementation process. For example, Fuhrman, Clune, and Elmore (1988), in their study of educational reforms in Arizona, California, Florida, Georgia, Minnesota, and Pennsylvania between 1986 and 1987, concluded that compliance depends heavily on the extent to which relevant technical knowledge exists at the school and state levels. School personnel and teachers must *feel* competent enough to make the change.

By comparing reforms on curriculum standards vs. reforms on teacher policies (mostly teacher career ladder-related policies) across the six states, the authors observed a consistent pattern. They found that "student curriculum standards mandates were notably more straightforward and understood". Thus, they were "more easily implemented than teacher

policies...” (p. 216). The possible reason for such a phenomenon, according to Fuhrman, Clune, and Elmore (1988), is that reforms focusing on student curriculum standards were policies with which educators felt more comfortable. Creating more academically oriented high schools was a task for which teachers already had training and experience. In contrast, the weaker knowledge bases of teacher career-ladder policies showed implementation problems related to teacher performance and progress.

Fuhrman, Clune, and Elmore’s conclusions (1988) were also echoed by Sabatier (1986). Sabatier’s research (1986) focused primarily on investigating the ability of statutes in structuring implementation, and the effect of statutes on the status of the implementation process. Based on twenty empirical reform implementation studies, Sabatier (1986) concluded that implementing officials’ *commitment and skills* were the most consistently critical determinants for successful reforms across cases.

**Professional development.** McLaughlin (1988, 1991) gave highest priority to professional development factors in successful reform implementation efforts . From her point of view, professional development is the main avenue through which school systems can improve both teachers’ knowledge and skills relevant to reforms. She contended that professional growth opportunities were major incentives for teachers if the new policy entailed major shifts in instructional practices (1991). According to McLaughlin (1991), a sound and effective professional development program must be sustained over time, be directly applicable to classroom practice, and include opportunities for teachers to observe one another.

**Teacher attitudes, values and self-efficacy.** McLaughlin (1987) also stated that one of the important lessons learned from past research is that success of reforms depends on two critical factors: local capacity and will. *Will* refers to implementers' intrinsic value for reforms, including motivation or commitment, which reflects the implementer's assessment of the value of a policy or the appropriateness of a strategy.

A number of researchers studied the effects of teachers' perceived self-efficacy on classroom instruction. Teachers who believe strongly in their ability to promote learning create mastery experiences for their students, but those beset by self-doubts about their instructional efficacy construct classroom environments that are likely to undermine students' judgments of their abilities and their cognitive development (Bandura, 1997).

The evidence indicated that teachers' beliefs in their instructional efficacy partly determine how they structure academic activities in their classrooms and shape students' evaluations of their intellectual capabilities (Bandura, 1997). Gibson and Dembo (1984) conducted an observational study of how teachers of high and low perceived teaching efficacy managed their classroom activities. Teachers who had a high sense of instructional efficacy devoted more classroom time to academic activities, provided students who encounter difficulties with the guidance they needed to succeed, and praised students' academic accomplishments. In contrast, teachers of low perceived teaching efficacy spent more time on nonacademic pastimes, readily gave up on students if they did not get quick results, and criticized them for their failures.

Teachers' beliefs in their instructional efficacy also affect their receptivity to, and adaptation of, educational changes. For example, Olivier (1985) found that teachers of low

perceived mathematical efficacy distrusted their capacity to make good instructional use of computers. Similarly, school administrators who had a low sense of computer efficacy resisted adopting computers for instructional purposes (Jorde-Bloom & Ford, 1988).

Most relevant to the purpose of the present study, researchers found that teachers' beliefs in their teaching efficacy affected their general orientation toward the educational process as well as their specific instructional activities (Bandura, 1997). Those who had a low sense of instructional efficacy favored a custodial orientation that took a pessimistic view of students' motivation, emphasized control of classroom behavior through strict regulations, and relied on extrinsic inducements and negative sanctions to get students to study (Woolfolk & Hoy, 1990). Melby (1995) found that teachers with a low sense of efficacy distrusted their ability to manage their classrooms and focused more on the subject matter than on students' development. In comparison, teachers who believe strongly in their instructional efficacy tended to rely on persuasive means rather than authoritarian control and to support development of their students' intrinsic interest and academic self-directedness.

**Resources.** McLaughlin (1987) defined school capacity broadly. Capacity refers to the implementation officials and teachers having: (a) knowledge of the policy and skills to enact the corresponding changes; and (b) conditions which facilitate reforms, for example, availability of financial resources or the additional assistance from consultants or teaching aids. Ferguson and Ladd (1996) concluded that funding to support teacher development increased student achievement more than any other kind of resource, with teacher expertise and experience accounting for a larger proportion of the variance in students' achievement gains in reading and mathematics. Miles and Darling-Hammond (1998) reported case studies of five high-performance elementary and secondary schools that were redesigned to allocate teaching

resources in the classrooms in ways that better met student needs. In all five cases, resources were used to increase student-teacher contact time for instruction, reduce teacher pupil ratios, and provide more teaching aides, planning time, and materials aligned to externally mandated standards and tests.

The links between resources allocated towards reforms, teacher training/professional development, teachers' knowledge of subject area standards emphasized in reforms, and reform implementation have also been verified through cross-sectional survey research conducted by Chatterji, Sentovich, Ferron and Rendina-Gobioff (2002). These authors confirmed, using structural equation modeling with teacher survey measures, that associations between these variables were statistically significant ( $p < .05$ ) with a substantial proportion of the variance in reform implementation explained. Chatterji et al (2002) developed and validated an instrument, titled the *Teacher Readiness for Educational Reforms* (TRFR) survey, which served as a data-gathering tool in a larger study examining the influences of state-initiated standards-based assessment reforms in the state of Florida. Nine school districts and 780 teachers located in southwest Florida participated in the larger study.

Chatterji et al (2002) found that resources available for reform, teachers' content knowledge in Mathematics and Language Arts standards together accounted for 40% of the variability in reported levels of reform implementation by teachers. The standardized path coefficient between Knowledge of Language Arts standards and reform implementation was estimated to be .38, which indicated a positive influence of reform-relevant teacher knowledge on reform implementation levels. Specifically, this positive path coefficient suggested that for every standard deviation unit increase in Language Arts knowledge in teachers, there was a .38 standard deviation unit increase in reform implementation levels. The standardized path

coefficient between perceived levels of resources and reform implementation levels was estimated to be .49, which suggested that for every standard deviation unit increase in resources, there was a .49 standard deviation unit increase in reform implementation levels.

**Social supports and networks.** Many reform policies and early implementation research focused on removing or buffering constraints to effective practice—inadequate materials, lack of appropriate teacher preparation, insufficient skill to implement reforms, and so on. However, as contended by McLaughlin (1987), removing constraints or obstacles does not by itself ensure more effective practice.

A few researchers frame micro-level implementation issues through the social interaction lens (Fullan, 1991; McLaughlin, 2006; Spillane, Beiser, & Gomez, 2006). This perspective highlights that for social agents such as teachers, new mandates and policies only comprise part of their daily life; they encounter policy in a complex web of social and institutional contexts. Therefore, implementation is not about mindless compliance to a mandate or policy directive and implementation shortfalls are not just cases of individual resistance, incompetence or capability. Rather, implementation involves a process of situated sense-making (McLaughlin, 2006; Spillane, Beiser, & Gomez, 2006) that implicates an implementer's knowledge base, prior understanding, and beliefs about the best course of action. Recent research effort in policy implementation has started to pay attention to how normative factors may trump technical components of a policy (McLaughlin, 2006).

For example, Spillane, Reiser, and Gomez (2006) concluded that implementers' cognition to a new educational policy should no longer be isolated to respective individuals; rather, it should be studied as a distributed practice, which emphasizes the influence of social

interactions among implementing agents and their situation on individual understanding of and action to the policy. Using qualitative research methods such as field notes, interviews and videotapes to collect data, these researchers' work in the Distributed Leadership Study in K-8 Schools in Chicago showed that social interactions were most likely to be found in grade-level meetings, faculty meetings, and professional development workshops as well as informal interactions in the lunch room or between classes.

The influences of social networks in the school were noticed by other researchers as well. Berman and McLaughlin (1977) stated that a school's principal strongly influenced the likelihood of change because, according to their observations, "projects having the active support of the principal were the most likely to fare well" (p. 124). In comparison, Fullan (1991) addressed the importance of peer relationship in the school. The quality of working relationships among teachers is closely linked to implementation (Fullan, 1991). "Collegiality, open communication, trust, support and help, learning on the job, getting results, and job satisfaction and morale are closely interrelated" (p. 77). Therefore, in Fullan's opinion (1991), for individual teachers, support from their colleagues in school was also a critical factor in determining their practices of the new reforms in classrooms. The influences of social interaction within schools on the implementation of reforms documented by the studies reviewed above were mostly based on qualitative observations made by external researchers.

**High stakes testing, accountability system and SCP practices.** Exploratory evidence from the U.S. shows that teachers generally perceive the student-centered instructional approaches are counterproductive in raising student performance on current standardized tests and that teacher use of SCP has been very likely adversely affected by high-stakes accountability



reforms (Deboer, 2002; Meyer & Rowan, 2006; Passman, 2000; Pedersen & Liu, 2003; Spillane & Burch, 2006).

For example, Passman (2000), using the case-study method, documented how a language arts teacher, who was passionate about SCP, was compelled to abandon student-centered teaching and turn to the teach-to-test methodology because student-centered teaching was perceived by the principal as ineffective in improving students' test scores.

As documented by Passman (2000), the instructor was teaching a unit on the Age of Exploration, a long-term inquiry project about explorers. The students were asked to choose one question, do research and discover the answer by groups. They were then asked to take several weeks to prepare a report, both written and visual, to present to the school community. The students researched at the school library, connected to the Internet, and looked at the classroom resources on their topics. The teacher's role became more of a coach in which the shift of responsibility for learning was on the student. The project was a success: two of the instructor's groups gave very impressive and sophisticated presentations on the topics of "navigation" and "supplies". The teacher was amazed by the progress students made as a result of the student-centered approach. However, soon after the completion of the exploration project, documented by Passman, the principal called a faculty meeting, directly ordered teachers to abandon teaching material that is not on the Iowa test. He then reminded the teachers about probation, testing success, and job security. After the meeting, the fifth grade teacher felt compelled to return to a traditional classroom setting and abandon her efforts toward a student-centered pedagogy.

Pedersen and Liu (2003) documented similar reactions of teachers toward high-stakes testing and accountability measures. Through interviews and records of class observations of 25

school teachers, these two researchers found that teachers were most concerned about whether a new student-centered program is helpful in preparing students for standardized tests.

Pedersen and Liu noticed that this concern was most often expressed by those whose school districts were facing a dropping rating in standardized testing. One teacher commented:

My school is totally [standardized test] –driven. We dropped a rating this year and you wouldn't believe what's going on about it; the things that are required about it...We are motivated by scores. The teachers aren't necessarily, but the school district's motto is your [standardized test] scores are everything. And that's not just [out district], that's the whole state (Pedersen & Liu, 2003, p.69).

The majority of the interviewed teachers, according to Pedersen and Liu (2003), believed that for the amount of concept learning that occurs, student-centered activities are more time-consuming than teacher-directed ones; therefore, they would use only a limited number of student-centered activities in a year and be less likely to use these activities during periods when they are preparing students for a standardized test.

Watanabe (2007), based on ethnographic case studies of two teachers' classroom and interviews with 13 teachers at five middle schools, illuminated how high-stakes testing narrows the curriculum and displaces teachers' priorities for their students. Watanabe also noted that these findings are noteworthy given that many of the teachers' instructional priorities intersect with state accountability measures.

In the study, Watanabe (2007) documented teachers' three teaching priorities in teaching Language Arts based on self-report measures. The three prioritized goals were: a) personal appreciation and enjoyment of literature, b) communication and collaboration skills, and c)

writing like a real writer writes. From the view of Watanabe (2007), these goals have progressive and constructivist underpinnings. However, realization of these teaching priorities has been greatly compromised by the North Carolina's high-stakes accountability program. All 13 teachers indicated that testing has had a demonstrable effect on instruction. Specifically, teachers speak of how testing and test preparation take instructional time away from their curriculum, squash students' desire to read and appreciate literature, decrease collaborative activities between students, and make writing instruction less like that of a real writer.

### **Systems-based and Multivariate Influences**

Literature reviewed in policy implementation and educational reforms show that reform implementation involves a large number of factors. But these factors are all interrelated (Afshari, Baker, Luan, Samah & Fooi, 2009). The success of the implementation of a new policy is not dependent on the availability or absence of one individual factor, but is determined through a dynamic process involving a set of interrelated factors (Afshari, Baker, Luan, Samah & Fooi, 2009).

## Summary of the Review of Literature:

### Constructs, Variables and A Conceptual Framework

Factors identified by the preceding literature review served as the theoretical foundation for construct domains tapped by the survey instrument developed for the present study, and helped conceptualize the hypothetical relationships to be tested between the school variables, teacher background variables, specific SCP-relevant teacher variables and teachers' implementation of SCP. This section synthesizes all the variables into a cohesive conceptual framework and path model from which specific hypotheses were sequentially tested to answer the research questions.

#### Constructs and Variables

The literature suggested that differences in school backgrounds would matter when it comes to reform implementation levels. School was thus identified as a key context variable in the larger system within which teachers work. Potential factors at the school-level were treated as one categorical variable, named *School*, representing combinations of school influences on the SCP reforms in classrooms.

At the teacher level, teaching *experience* was also viewed as a potential predictor of teachers' implementation of SCP. Given the literature on technology and other reforms, teachers with more years of teaching may be more reluctant to change their instruction in accordance with new policies. *Gender* was another relevant teacher background characteristic. If women are more susceptible to external influences, it is not unreasonable to assume that there is a relationship between gender of the teacher and SCP implementation levels, with female teachers likely to have a higher tendency to practice SCP-relevant instruction. Similarly, whether

teachers' education levels lead to more substantive instructional changes towards SCP in the context of China's new curriculum reforms was also a pertinent question.

The review of literature on educational reforms and policy implementation in the U.S. revealed that teachers' implementation behaviors are not only affected by their understanding regarding technical properties of the new reforms and their ability in carrying out the new reforms, but also affected by their beliefs and values related to specific reforms, and availability of necessary support from the policy environment, including norms of the system shared by principals and colleagues. Collectively, the main constructs derived from literature review were: *Perceived Support for SCP Reforms*, *Beliefs Regarding SCP*, *Self-Efficacy in Practicing SCP*, and *Perceived Control by the Teacher Evaluation Policies*. Table 2 summarizes all these constructs in detail. Specifically, under each construct/variable, it presents name of the variable, indicators used to develop pertinent survey items, and literature sources that support the construct measure.

The policy implementation and educational reforms literature suggested that multiple, interrelated factors function dynamically in affecting teacher implementation of a new policy/program. At the school level, the influences may come from variations in school leadership style, policy interpretation, school climate and so on. At the teacher level, teacher characteristics such as gender, experience, and highest degree obtained could all influence teachers' choice of action responding to the reforms. The three teacher belief and perception factors were selected as SCP-relevant and expected to apply cross-culturally in China's setting because of the overlaps found in the literature on how reforms work in the centralized, large-scale, and accountability-driven systems. Note that these factors were derived from the angle of reform implementers at the bottom of the hierarchy.

Table 2

*Constructs and Variables for Study*

| Construct /Variable               | Indicator(s)   | Supporting Literature  |
|-----------------------------------|--|--|
| <i>School Context Factors</i>     |  |  |
| School                            | Omnibus factor representing differences in leadership style, interpretation of reform policies, student composition, and so on | Bulach & Malone, 1994; Datnow & Castellano, 2001; Datnow, Hubbard, & Mehan, 1998; Leithwood & Jantzi, 2006; Stalling & Mohlman, 1981; Yuen, Law, & Wong, 2003  |
| <i>Teacher Characteristics</i>    |  |  |
| Gender                            |  | Afshari et al., 2009; Hargreaves, 2005; Rogers, 1995; Schiller, 2003; Smith & Desimone, 2005; Venkatesh and Morris, 2000   |
| Experience                        |  |  |
| Highest Degree Obtained           |  |  |
| <i>SCP-Relevant Variables</i>     |  |  |
| Perceived Support for SCP Reforms | 1. Professional Development Opportunities  | Berman & McLaughlin, 1977; Chatterji et al., 2002; Ferguson, 1991; Ferguson & Ladd, 1996; Fuhrman, Clune, & Elmore, 1988; Fullan, 1991; McLaughlin, 1987, 1988, 1991, & 2006; Miles & Hammond, 1998; Spillane, Reiser, & Gomez, 2006 |
|                                   | 2. Necessary resources   |  |
|                                   | 3. Support from the principal  |  |
|                                   | 4. Support from colleagues   |  |
| Beliefs Regarding SCP             | 1. Beliefs in teacher's new role and responsibilities in using SCP   | Chatterji et al., 2002; Fuhrman, Clune, & Elmore, 1988; McLaughlin, 1987, 1988, 1991, & 2006; Sabatier 1986  |
|                                   | 2. Beliefs in the merits of SCP instructional strategies   |  |

Table 2 continued.

*Constructs and Variables for Study*

| Construct /Variable  | Indicator(s)   | Supporting Literature   |
|--|--|---|
| Self-Efficacy in Practicing SCP                            | Belief in self-capacity to: <ol style="list-style-type: none"> <li>1. Organize group activities</li> <li>2. Facilitate inquiry-oriented class discussions</li> <li>3. Use probing questions</li> <li>4. Employing multiple forms of assessment and SCP strategies</li> </ol> | Bandura, 1997; Gibson & Dembo, 1984; Melby, 1995; Olivier, 1985; Sabatier, 1986; Woolfolk & Hoy, 1990   |
| <i>Mediating Factor: Output Control Variable</i>           |  |   |
| Perceived Control by the Teacher Evaluation Policies       | <ol style="list-style-type: none"> <li>1. Perceived control by the student test score-related performance review policies</li> <li>2. Perceived control by the process-related performance review policies</li> </ol>  | Chen & Li, 2007; Deboer, 2002; Evans, 1975; Han & Yang, 2008; Jiang, 2008; Ouchi, 1977, 1978, & 1979; Passman, 2000; Pedersen & Liu, 2003; Watanabe, 2007; Williamson, 1971 Eisenhardt, 1985; Zhang, 2007; Zhao, 2007 |
| <i>SCP Implementation: Dependent Variable (or Outcome)</i> |  |   |
| SCP Implementation Levels                                  | <ol style="list-style-type: none"> <li>1. Frequency with which teachers use specific SCP-related practices - self-reported</li> </ol>  | Knowlton, 200; Passman, 2000; Pedersen & Liu, 2003; Zhong, 2007, 2008   |
| <i>Moderating Variables</i>                                |  |   |
| Grade Level  |  | Amrein & Berliner, 2002; Nichols, Glass, & Berliner, 2006   |
| Class Size   |  | Molnar, Smith, Zahorik, Palmer, Halbach, & Ehrle, 1999  |

## Conceptual Framework

Figure 1.0, presents a hypothesized path diagram reflecting the dynamics among the theoretically-derived variable in Table 2. Arrows indicate directions of the hypothesized relations among variables, suggested by literature.

In Figure 1.0, *SCP Implementation* is the *Dependent Variable*, with the arrows from multiple factors at the school and teacher/classroom levels directly or indirectly influencing that outcome. All other factors are *Independent Variables (IV)* that are exogenous (School) or mediating factors (teacher characteristics, beliefs, and perceived control) in the model.

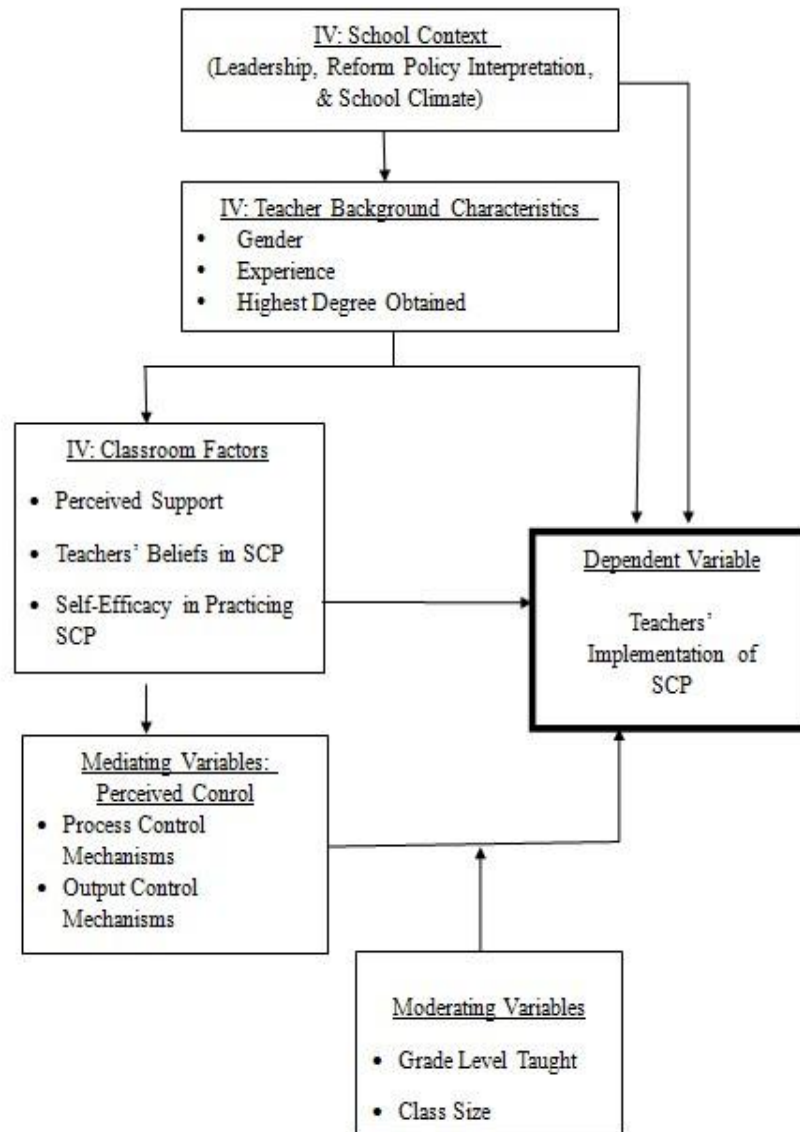
This study tested the assumption that the directional influences of factors at the school level and teacher level on SCP implementation, would be mediated by teachers' perceived levels of control by the teacher evaluation policy. China's reliance on output-control mechanisms is consistent with predictions from organizational theory on the relationship between organizational structure and control mechanisms. In this study, the influence of the output control mechanisms was operationally defined by teacher perceptions of the control enforced by performance evaluation policies. Literature in high-stakes testing and accountability system from the U.S. suggested a negative relationship between the implementation of the student-centered approaches and the high-stakes accountability system.

The relationship between teachers' perceived levels of control by the evaluation policy and teachers' implementation of SCP was expected to be *moderated* by grade level and class size. Interaction effects were expected to be significant. Because the *Gaokao* is administered as a high stakes test in grade 12, it was reasonable to expect that SCP implementation would vary by grade. Further, as smaller classes are known to be better for student oriented instruction, SCP



implementation was also expected to vary by class size. Correlations among all variables were first expected before regression models were run to test hypotheses.

Figure 1.0 Conceptual Framework of the Study



NOTE: IV=Independent Variables

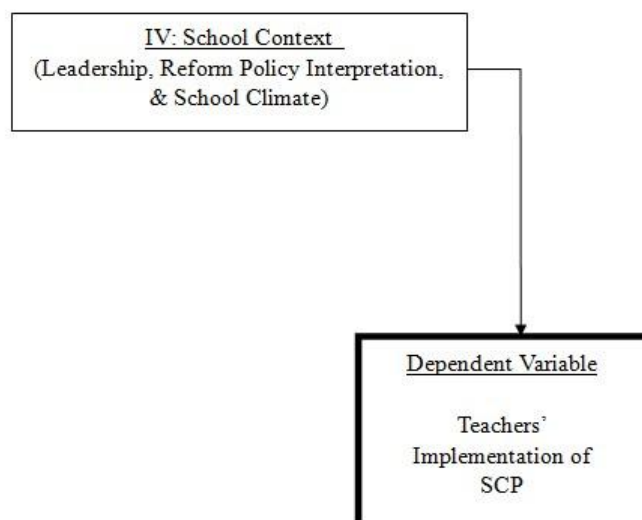
## Research Questions and Hypotheses

Based on the preceding conceptual framework, six relational hypotheses were tested. Each was designed to answer one research question. The hypotheses and research questions are now presented in sequence, with segments of the larger conceptual model extracted. ‘IV’ refers to the Independent Variables in the analytic models (Figures 1.1-1.5).

**Research question 1.0:** To what extent do *school* characteristics, as predicted by educational reform and policy implementation literature, affect SCP implementation at the classroom level reported by teachers?

**Hypothesis 1.0:** Differences in *school* context (e.g., leadership style, policy interpretation, and school climate) will significantly and substantially predict levels of *classroom implementation of SCP*.

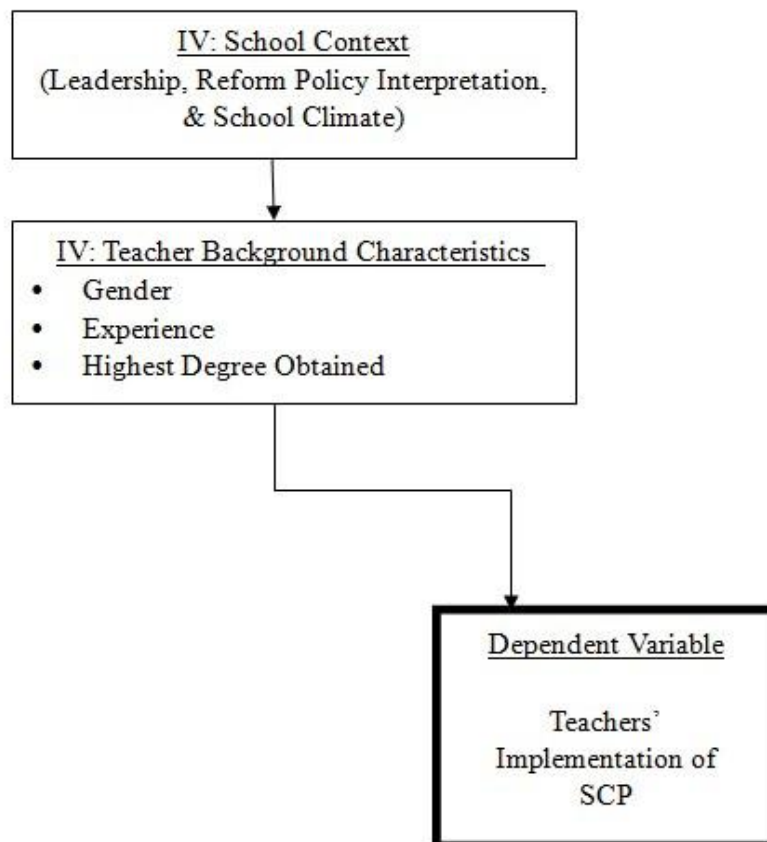
Figure 1.1 Excerpt of Figure 1.0 Relevant to Hypothesis 1.0



**Research question 2.0:** To what extent do selected *teacher characteristics*, as predicted by educational reform and policy implementation literature, affect SCP implementation in the classroom as reported by teachers, after taking school-level differences into account?

**Hypothesis 2.0:** Controlling for school level variability, *teacher background characteristics* such as Gender, Experience, and Highest Degree Obtained will significantly and substantially predict levels of *classroom implementation of SCP*.

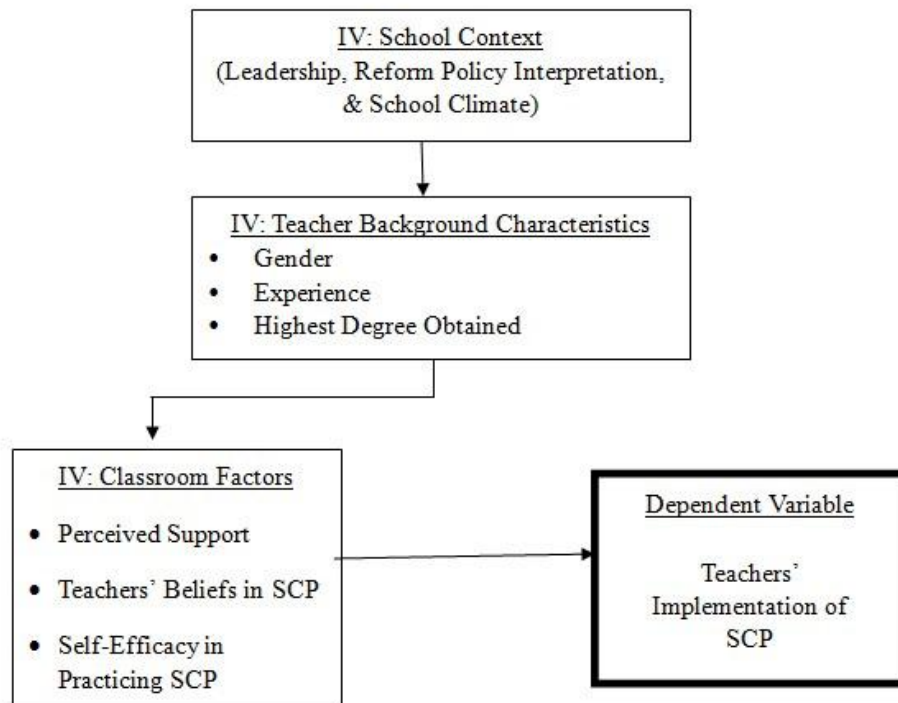
Figure 1.2 Excerpt of Figure 1.0 Relevant to Hypothesis 2.0



**Research question 3.0:** To what extent do selected SCP-relevant constructs (*Teacher Perceived Support, Teacher Beliefs in SCP, and Teacher Self-efficacy in Practicing SCP*), as predicted by the educational reform and policy implementation literature, affect *SCP implementation* in the classroom as reported by teachers, accounting for school and teacher characteristics?

**Hypothesis 3.0:** Controlling for school level variability and teacher background characteristics, *Teachers' Perceived Support for SCP Reforms, Beliefs in SCP, and Teachers' Self-efficacy in Practicing SCP* will significantly and substantially predict levels of classroom implementation of SCP.

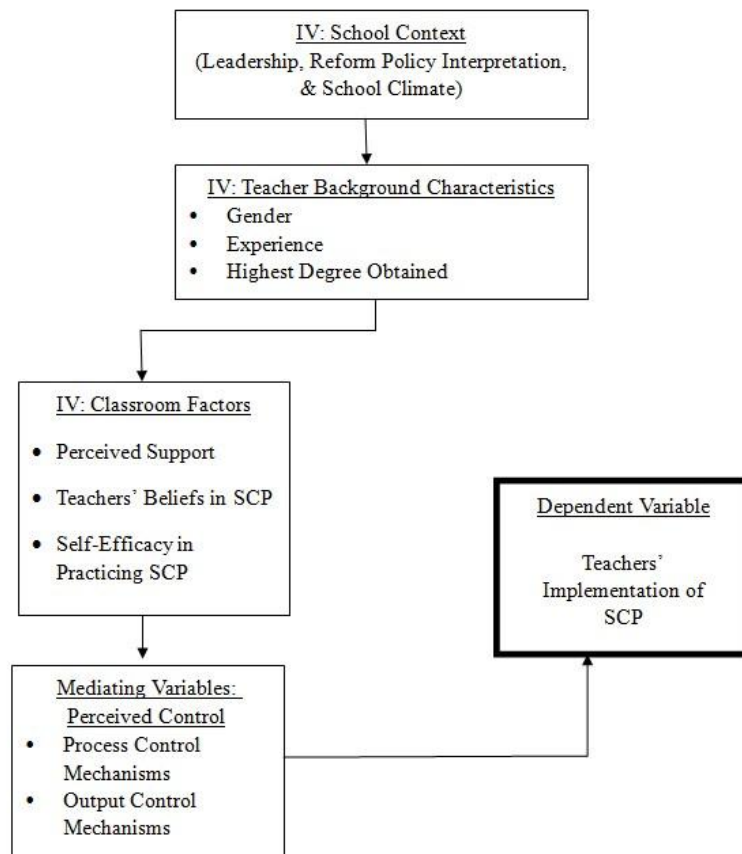
Figure 1.3 Excerpt of Figure 1.0 Relevant to Hypothesis 3.0



**Research question 4.0:** To what extent is *control* enforced by the organization's *Gaokao*-related teacher evaluation policy for secondary teachers, a negative predictor and a significant mediating variable for SCP implementation in classrooms reported by teachers?

**Hypothesis 4.0:** Controlling for school level variability, teacher background characteristics, and SCP-relevant factors, *Teachers' Perceived Levels of Control by the Teacher Evaluation Policy* will significantly and negatively predict levels of classroom *implementation of SCP*.

**Figure 1.4 Excerpt of Figure 1.0 Relevant to Hypothesis 4.0**

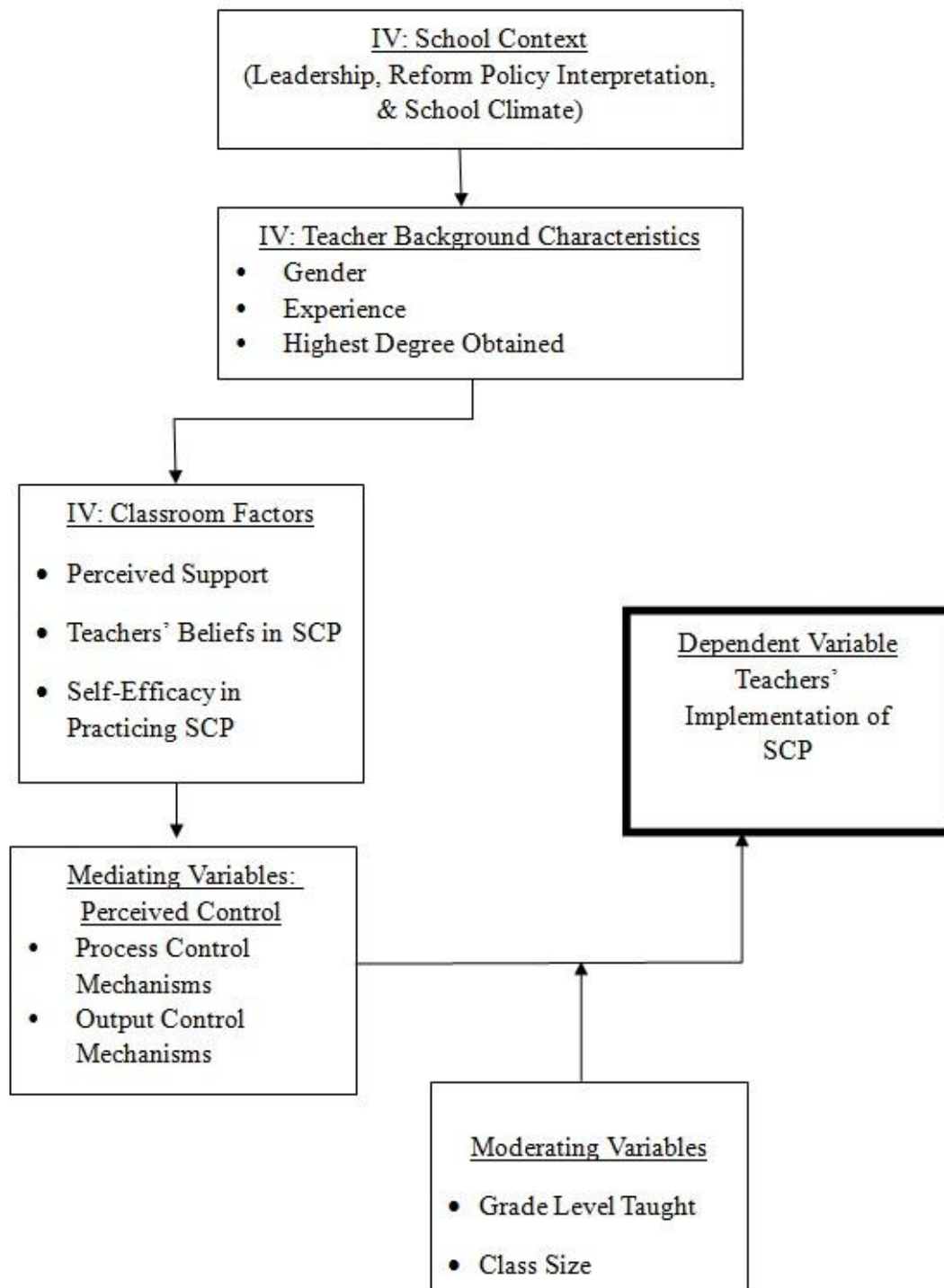


**Research question 5.0:** To what extent is the relationship between *control* enforced by the organization's *Gaokao*-related teacher evaluation policy for secondary teachers, with reported SCP implementation levels in the classroom, moderated by *Grade Level* and *Class Size*?

**Hypothesis 5.0:** The relationship between teachers' perceived levels of *control* by the teacher evaluation policy and levels of classroom *implementation of SCP* will be moderated by *Grade Level* taught by teachers. [Grade 12 teachers will show lower levels of implementation than Grade 10-11 teachers.]

**Hypothesis 6.0:** The relationship between teachers' perceived levels of control by the teacher evaluation policy and teachers' levels of implementation of SCP will be moderated by *Class Size*. [Teachers with smaller classes will show higher levels of SCP implementation than those with larger classes.]

**Figure 1.5 Excerpt of Figure 1.0 Relevant to Hypothesis 5.0 and Hypothesis 6.0**



### **CHAPTER III METHODOLOGY**

This chapter provides a description of the research methods employed to answer the research questions and investigate the hypotheses formulated for the study. First, the research site, units of analysis, the target population and the sampling procedure are described. Next, the characteristics of the teacher sample are displayed with attention given to population representativeness on two variables, school and gender. This is followed by a detailed report of the instrument design and validation procedures. Details are provided on the survey domain specifications, a pilot test on survey measures, and an exploratory factor analysis and reliability investigations using the larger study sample. The final section describes the analytical methods and equations employed to test individual hypotheses stemming from the conceptual framework and path diagram for the study, with all pertinent statistics identified.

#### **Research Site**

The study was conducted in public high schools in Jingyang District of Deyang, a mid-sized city located in Sichuan Province of China. The district has 7 high schools, 29 middle schools, and 49 elementary schools. Since 2003, the district has required district-wide adoption and implementation of the new curriculum reforms (Deyang City Bureau of Education, 2005). All seven high schools are located in the urban area of the City of Deyang, covering one district (Jingyang District) and three counties.



## **Sampling Design**

### **Unit of Analysis**

Teachers responding to the survey served as the unit of analysis for the study. Since all seven high schools in this study fell under the teacher performance evaluation policy enacted at the district level, teachers' perceptions of this policy were expected to vary across schools at the individual level. The target population consisted of 526 high school teachers from Grades 10 -12, excluding music and physical education teachers. In the population, approximately 48% were females, with 52% males.

### **Sample Size and Statistical Power Estimation**

Published power tables (Judd & McClelland, 1989) were used to estimate the best sample size so as to obtain optimal levels of statistical power for the present study. A power estimation was conducted with the following parameters:

- a) for hypotheses tests, a significance level was set at the .05 level (or 5% error), consistent with the usual practice in social science;
- b) the size of the effect desired for the hypotheses tests, adopting a conservative approach as recommended by Cohen (1988), was set at .03; and
- c) statistical power was estimated for bi-directional hypotheses.

According to the power tables, with 200 as the number of observations, the expected statistical power was .90. This indicated that the study would have a 90% chance of detecting a true and statistically significant relationship between variables at a .05 level with two-tailed

hypothesis tests. To accommodate contingencies during data collection, such as non-responders and missing data, the present study increased the targeted sample size from 200 to 300.

### **Random Sampling Procedure**

The sampling frame for the present study was the employee list provided by Office of Human Resources of Jingyang District Department of Education. This list contained names and basic information of all teachers employed by the district (Grades 10 through 12). Substitute teachers and administrative staff were not included in the employee list.

Consistency between the target population and the sampling frame lessens non-sampling bias and in turn lessens total error (Henry, 1990). Despite minor inconsistencies due to two female teachers being on maternity leave, and one male teacher out on business trips, the above list provided a complete sampling frame for the present study.

Simple random sampling was then conducted using procedures in SPSS (Version 16.0) to yield 300 teachers for the study. The sampling rate was .57.

### **Data Collection Procedure**

The survey instrument was handed out to teachers at each high school during weekly staff meetings. After obtaining permission from the school administration, at the end of the weekly staff meeting, selected teachers were asked to stay. A PowerPoint presentation was used to introduce the study to the teachers. Confidentiality was assured. The teachers were told that participation was voluntary. Teachers who were willing to take the survey then filled out the questionnaire. The same procedure was followed at all seven high schools in this study. The entire data collection took two months to complete, extending from March -May, 2010.

### **Sample Composition and Representativeness**

Of the 300 randomly selected teachers, 250 returned the questionnaires. Upon preliminary screening of the responses, 18 responses were deemed invalid because they either showed an abnormal answering pattern (e.g., respondent chose “agree” for all of the items), or left more than half of the items unanswered. After the screening process, the final data set had a total of 232 cases with complete data. Sample friction showed no particular pattern indicating selection bias.

Tables 3 and 4 provide details of the sample on demographic variables. The sample is composed of 108 (46.5%) females and 124 (53.5%) males. The vast majority of participants in the sample held 4-year college degrees (188, 81%). A small number held Master’s degrees (6, 2.6%), with the remaining holding 2-year Associate degrees (13, 5.6%). With regard to grade level distributions, there is a relatively even spread, with 61 teaching 10<sup>th</sup> grade (26.3%), 74 teaching 11<sup>th</sup> grade (31.9%), and 87 teaching 12<sup>th</sup> grade (37.5%). In terms of subject matter taught, there are a total of 9 subjects taught by this sample, with Literature (43, 18.5%), Math (42, 18.1%) and English (40, 17.2%) being taught by the largest share of teachers, and Geography (9, 3.9%) and Biology (7, 3%) on the other end of the continuum.

Table 3

*Background Characteristics of Sample on Education, Grade and Subject Taught*

|                            | Frequency | Percent |
|----------------------------|-----------|---------|
| <b>Highest Degree</b>      |           |         |
| 2 Year Associate           | 13        | 5.6     |
| 4 year college             | 188       | 81.0    |
| Master's Degree            | 6         | 2.6     |
| Missing                    | 25        | 10.8    |
| <b>Grade Level Taught</b>  |           |         |
| 10                         | 61        | 26.3    |
| 11                         | 74        | 31.9    |
| 12                         | 87        | 37.5    |
| Missing                    | 10        | 4.3     |
| <b>Main Subject Taught</b> |           |         |
| Missing                    | 17        | 7.3     |
| Biology                    | 7         | 3.0     |
| Chemistry                  | 17        | 7.3     |
| English                    | 40        | 17.2    |
| Geography                  | 9         | 3.9     |
| History                    | 19        | 8.2     |
| Literature                 | 43        | 18.5    |
| Math                       | 42        | 18.1    |
| Physics                    | 19        | 8.2     |
| Politics                   | 19        | 8.2     |

*Note.* N=232

Table 4 shows that with regard to years of teaching experience, the sample *Mean* is 14.71 years with the *Standard Deviation* being 7.97. The sample shows a large *Range*, with a minimum value of 1 year and a maximum of 38 years of experience. With respect to class size, the minimum class size reported by respondents is 30, with the largest class reported as 75, with a *Mean* of 58.48 and a *Standard Deviation* of 9.10.

Table 4

*Background Characteristics of Sample on Teaching Experience and Class Size*

|                            | Minimum | Maximum | Mean  | Standard<br>Deviation |
|----------------------------|---------|---------|-------|-----------------------|
| Total Years of<br>Teaching | 1       | 38      | 14.71 | 7.97                  |
| Class Size                 | 30      | 75      | 58.48 | 9.10                  |

*Note.* N=232

**Sample Representativeness on Gender and School Membership**

Table 5 shows that the gender distribution of the sample as the following: 108 (46.5%) females, and 124 (53.5%) males. This is fairly consistent with that of the population, among which 254 (48.3%) were females, and 272 (51.7%) were males.

Table 5

*A Comparison between the Population and the Sample: Gender*

|        | Population |             | Sample    |             |
|--------|------------|-------------|-----------|-------------|
|        | Frequency  | Percent (%) | Frequency | Percent (%) |
| Female | 254        | 48.3        | 108       | 46.5        |
| Male   | 272        | 51.7        | 124       | 53.5        |
| Total  | 526        | 100         | 232       | 100         |

Table 6 shows that among the 232 sampled teachers, 57 (24.6%) were from The 1<sup>st</sup> High School; 38 (16.4%) from The 2<sup>nd</sup> High School; 41 (17.7%) from The 7<sup>th</sup> High School; 27 (11.6%) from The 8<sup>th</sup> High School; 15 (6.5%) from The 9<sup>th</sup> High School; 39 (16.8%) from Xiaoquan High School; and 15 (6.5%) from Yangjia High School. With respect to representativeness, teachers from The 7<sup>th</sup> High School and Xiaoquan High School were at a slightly higher percentage in the sample than that in the population (17.7% vs. 15.4%, 16.8% vs. 13.1%, respectively). Teachers from The 2<sup>nd</sup> High School and The 9<sup>th</sup> High Schools had a lower percentage in the sample than that in the population (21.0% vs. 16.4%, 6.5% vs. 8.6%, respectively).

Table 6

*A Comparison between the Population and the Sample: School of Teaching*

|                        | Population |             | Sample    |             |
|------------------------|------------|-------------|-----------|-------------|
|                        | Frequency  | Percent (%) | Frequency | Percent (%) |
| The 1 <sup>st</sup> HS | 123        | 23.4        | 57        | 24.6        |
| The 2 <sup>nd</sup> HS | 110        | 21.0        | 38        | 16.4        |
| The 7 <sup>th</sup> HS | 81         | 15.4        | 41        | 17.7        |
| The 8 <sup>th</sup> HS | 58         | 11.0        | 27        | 11.6        |
| The 9 <sup>th</sup> HS | 45         | 8.6         | 15        | 6.5         |
| Xiaoquan HS            | 69         | 13.1        | 39        | 16.8        |
| Yangjia HS             | 40         | 7.6         | 15        | 6.5         |
| Total                  | 526        | 100         | 232       | 100         |

### **Instrument Design and Validation**

The main instrument used for the present study was a multiple-domain teacher questionnaire designed to collect information through teachers' self-reports on the following variables: school, background characteristics, several SCP-relevant support and affective variables, perceived control exercised by output and process aspects of the teacher performance evaluation system. All these served as independent or mediating variables in the conceptual framework of the study, shown in Figure 1 of Chapter II. The instrument also tapped self-reported levels of SCP implementation in the classroom, the dependent variable in the study's conceptual framework.

The present study followed the steps specified in Chatterji's (2003) *Process Model* for instrument design and validation, which involves a four-phase, iterative "design, check, revise, confirm" (Chatterji et al., 2002, p. 448) approach to develop measures. The instrumentation methodology here was guided by another study involving the design and validation of a teacher survey to evaluate reforms in Florida (Chatterji et al., 2002).

Use of the iterative approach helped identify and control for measurement errors in the survey-based measures prior to use of the survey instrument in the larger research investigation. There were three levels of iteration to refine the measures. A content validation, small-scale pilot-testing, and empirical validation with a large sample.

Initially, the instrument was content-validated and pilot-tested to obtain preliminary validity and reliability evidence. After data collection was completed with the full sample ( $N=232$ ), an exploratory factor analysis was conducted and reliability estimates for the final survey measures were compared with those from the pilot for the factor-defined scales.

### **Specification of Survey Domains**

As indicated in Chapter II, the teacher questionnaire contained multiple parts measuring five domains, derived from the literature review: a) perceived support for implementing SCP, b) beliefs in SCP, c) self-efficacy in practicing SCP, d) perceived controls by the teacher performance evaluation policy, and e) implementation of SCP. School context and background information on teachers were collected in a demographic section at the beginning of the instrument based on teacher reports. Appendix B gives a summary of the indicators in the five domains.



Items were written with a positive and negative orientation to match indicators in the domain,. The negatively oriented items were distributed randomly in the questionnaire to control for socially desirable responses, faking or fixed response sets from teachers. The five domains are elaborated below with positively- and negatively-oriented item examples.

**Domain 1.0: Perceived support for implementing SCP.** The first domain focused on assessing teachers' perceptions of exogenous school conditions that may affect their SCP reform behaviors, in particular the support received from the surrounding policy environment. The post-pilot version of this domain had 13 items.

One example of the positively-oriented item in this domain was: "SCP-relevant professional development programs are directly applicable to my classroom practice." One example of the negatively-oriented item for this domain was "I have not received any SCP-relevant in-service training." Domain 1.0 employed a 5-point endorsement scale. Coding for this scale was: Strongly Disagree (1 point), Disagree (2 points), Uncertain (3 points), Agree (4 points), and Strongly Agree (5 points).

**Domain 2.0: Teachers' beliefs in SCP.** The second domain focused on measuring teachers' reports of their beliefs in SCP. The post-pilot version of this domain had 9 items.

One example of the positively-oriented item in this domain was "A student should be assessed in a variety of ways, such as, with projects, essays, multiple choices, portfolios, and so on." One example of the negatively-oriented item was "The teacher's work should be mainly to transmit knowledge to students." Domain 2.0 employed a 5-point endorsement scale. Coding for this scale was: Strongly Disagree (1 point), Disagree (2 points), Uncertain (3 points), Agree (4 points), and Strongly Agree (5 points).

**Domain 3.0: Teacher reports of self-efficacy in practicing SCP.** The post-pilot version of this domain had 11 items. Teachers' competency in implementing SCP was measured on four instructional aspects: group activity, inquiry discussion, questioning, and assessment. One example of the positively-oriented item was "I can effectively ask questions that make students think in depth." One example of the negatively-oriented item was "I find myself having difficulties in designing projects that are appropriate for students of different ages and developmental stages." Identical to Domain 1.0 and Domain 2.0, Domain 3.0 also employed a 5-point endorsement scale. Coding for this scale was: Strongly Disagree (1 point), Disagree (2 points), Uncertain (3 points), Agree (4 points), and Strongly Agree (5 points).

**Domain 4.0: Perceived control by the teacher evaluation policy.** The post-pilot version of this domain had 13 items. Domain 4.0 contained two components: a) teacher perceptions of control that is output-based in their performance evaluation system, and b) teacher perceptions of control that is process-based in their performance evaluation system.

Domain 4.0 employed a 5-point ordered scale to measure teachers' perceptions on the degree to which their school emphasizes both output-based and process-based components when conducting teaching evaluations. Coding for this scale was: Very Low/Not at All (1 point), Low (2 points), Moderate (3 points), High (4 points), and Very High (5 points).

**Domain 5.0: Teacher reports of SCP implementation.** The post-pilot version of this domain had 12 items. This domain tried to measure the frequency with which teachers practiced specific strategies of SCP. These strategies included interactive learning, flexible grouping, asking questions that are more probing, assigning inquiry-based homework (such as projects), using multiple assessment methods, and involving students in designing activities and lessons.

Domain 5.0 employed a 4-point frequency scale. Coding for this scale was: Never/Rarely (1 point), Sometimes (2 points), Often (3 points), and Very Often (4 points).

### **Pre-Pilot Content Validation of Items**

As a part of the content validation process of the pre-pilot version of the tool, a structured review of items was conducted by two professors in Teachers College (TC), Columbia University. One reviewer was from the Department of Organization and Leadership (DEOL) and the other from the Department of Curriculum and Teaching (DCT). The professor from DEOL was asked to check if the factors deemed important in affecting SCP reform behaviors were adequately included in the instrument. The professor from DCT was asked to check if the items were correct in the sense of being consistent with generally accepted understandings of what the SCP is and what instructional strategies it manifests. A fellow graduate student from TC who is also familiar with the *Process Model* was invited to conduct the content validation as well. Her main goal was to check whether the writing of the items followed the rules and established guidelines for designing behavioral and affective assessments (Chatterji, 2003). The instrument items were then modified based on their feedback.

### **Back-Translation Method for Designing a Bilingual Survey**

The items were constructed in English first. Since all the respondents were Chinese nationals, a special back-translation process (Brislin, 1986) was used to prevent the essence of the item meanings from getting lost in translation. The questionnaire was first translated into Chinese by a doctoral student from College of Education, Beijing Normal University, China, and then translated back to English by another doctoral student from the same institute. These two doctoral students from Beijing Normal University were visiting scholars at Teachers College,

Columbia University. They are not only fluent in Chinese and English, but also familiar with the SCP terminology in both languages. The original English version was then compared to the version translated back from the Chinese version. Modifications were made to places where inconsistency occurred.

### **Pilot Study**

The pilot of the instrument was mainly focused on preliminary item analysis and reliability investigations, with 35 teachers participating. At the time, two domains (Teachers' Implementation of SCP and Perceived Control by the Teacher Evaluation Policy) were not included as they were still in the early developmental stages.

The first aim of the pilot was to determine whether an item was consistently measuring the same characteristic as the other items in the same domain with correlational procedures, using corrected item-total score correlations. When items showed a corrected item-to-total correlation of  $+0.30$  or better, they were considered a good addition to the domain score. Negative or low values (less than  $.30$ ) suggested item problems requiring revision or potentially removal. Cronbach's *Alpha* was calculated as well for each domain to determine the degree to which items from the same domain generate consistent patterns of responses for individual respondent. To be acceptable, Cronbach's *Alpha* should be at least  $.70$  (Chatterji, 2003; Crocker & Algina, 2006).

The results of the pilot testing were mixed, with two of the three domains showing acceptable to high internal consistency estimates (see Table 7 below). One domain appeared problematic. More specifically, Cronbach's *Alpha* for Perceived Support for SCP Reforms, and Self-Efficacy in Practicing SCP were  $.718$  and  $.936$ , respectively. But, Beliefs in SCP had an

internal consistency estimate of .596, which is lower than the acceptable .70 threshold. A review process was then initiated to revise or remove some of the problematic items.

Table 7

*Descriptive Statistics and Reliability Coefficients of the Three Domains in the First Iteration*

| Domains                           | # of Items | <i>M</i> | <i>SD</i> | Minimum | Maximum | Cronbach's <i>Alpha</i> |
|-----------------------------------|------------|----------|-----------|---------|---------|-------------------------|
| Beliefs Regarding SCP             | 16         | 57.5187  | 4.2187    | 47      | 62      | .596                    |
| Perceived Support for SCP Reforms | 14         | 47.463   | 4.879     | 37      | 57      | .718                    |
| Self-Efficacy in Practicing SCP   | 19         | 70.178   | 7.813     | 56      | 89      | .936                    |

*Note:* *M* = scale mean; *SD* = scale standard deviation.

Item statistics for the three domains are presented in Appendix C. As a result of the qualitative reviews, Items 5 and 11 in Beliefs in SCP were deleted due to the fact that the content of Items 5 and 11 were quite new to the respondents and the items performed particularly poorly with significantly negative item-to-total correlations. On the other hand, Item 3 in Beliefs Regarding SCP and Items 5 and 9 in Perceived Support for SCP Reforms were retained and subject to revision for the problem they faced was mostly wording. All the items in Self-Efficacy in Practicing SCP were retained for the second iteration.

### **Instrument Refinements Following Pilot**

Following the pilot, a number of items were either deleted or re-drafted. After consulting a professor in Sichuan Normal University who is an expert in curriculum development, items containing words such as “like” in English were deleted since unlike what it indicates in English, in Chinese, literal translation of “like” would be most likely construed as an indication of frequency instead of preference. As a result of refinement based on both pilot results and post-pilot content validation with two professors, items for domain Perceived Support for SCP Reforms were reduced from 14 to 13; for domain Beliefs Regarding SCP, reduced from 16 to 9; items for domain Self-Efficacy in Practicing SCP reduced from 19 to 11.

Per request by Jingyang District, the revised questionnaire was translated into Chinese for the actual distribution by two staff members from the Office of Policy Studies of the district. These two staff members are experienced researchers in the field of educational reforms and fluent in both Chinese and English. The final post-pilot instrument (in English) is attached as Appendix D with all five domains. The Chinese version is in Appendix E.

### **Empirical Validation of Final Instrument**

An Exploratory Factor Analysis (EFA) was conducted with 208 clean cases in the final sample, using a principal axis factor extraction procedure. This was followed by promax rotation of factors. The EFA procedure was performed with items in all domains to see if items salient to the underlying domains loaded on factors extracted.

The number of factors was decided based on evaluation of the scree plot, the size of the eigenvalues, cumulative percentage of variance explained, as well as consistency and meaningfulness of factors relative to the theoretically proposed domain structures (see Appendix

B). A promax rotation method was used based on the assumption supported by the theoretical framework that the factors would be inter-correlated. Item to factor loadings equal to or greater than .30 on a factor were considered as the cut-point for identifying items relevant to a factor/scale.

**Exploratory Factor Analyses (All Items).** A listwise deletion procedure was employed to remove cases with partially missing data. After this procedure, 208 cases remained for the analysis with similar gender distribution and grade distribution compared to the original 232 cases. The KMO (Kaiser-Meyer-Olkin) value of the dataset was .784, above the minimum requirement (.5) recommended by Kaiser (1974), indicating an adequate sample for factor analysis.

The EFA on all items resulted in the negatively-oriented items clustering to form new factors due to a semantic effect (Miller, 1987). As indicated, when designing the questionnaire, several negatively-oriented items were inserted into each domain to check as to whether respondents were faking or giving patterned answers (e.g., all agree responses). However, the literature showed that on occasion, negatively-oriented items may contain certain stimuli that would most likely invoke respondents to process survey items semantically (Miller, 1987). Although semantic processing is unintended and unconscious, it results in negatively-oriented items lumping together as a separate factor, confounding theoretical interpretations of the factor analysis results. When more than one such item is built into domains, they complicate understanding of the empirically-derived factors. Therefore, a new round of factor analysis was conducted excluding the negatively-oriented items. The second-round EFA results are reported below.

**Exploratory Factor Analyses (Positively-oriented Items).** With the negatively-oriented items excluded, the EFA results showed that the first thirteen factors had eigenvalues greater than 1. The scree plot suggested one break after five factors, another after seven, and another after nine. Cumulative percentage variance explained by first five factors was around 41%, with the first seven factors explaining approximately 46% and the first nine factors approximately 50% of the total variance in items (See Table 8).

Table 8

*Percentage of Variance Explained for Extracted Factors in Exploratory Factor Analysis*

| Factor | Initial Eigenvalues |               |              | Extraction Sums of Squared Loadings |               |              | Rotation |
|--------|---------------------|---------------|--------------|-------------------------------------|---------------|--------------|----------|
|        | Total               | % of Variance | Cumulative % | Total                               | % of Variance | Cumulative % | Total    |
| 1      | 8.22                | 17.12         | 17.11        | 7.81                                | 16.27         | 16.27        | 5.48     |
| 2      | 4.67                | 9.71          | 26.83        | 4.24                                | 8.83          | 25.10        | 4.58     |
| 3      | 4.01                | 8.37          | 35.2         | 3.63                                | 7.56          | 32.67        | 5.21     |
| 4      | 2.64                | 5.51          | 40.71        | 2.23                                | 4.64          | 37.30        | 4.64     |
| 5      | 2.29                | 4.77          | 45.48        | 1.83                                | 3.81          | 41.12        | 3.27     |
| 6      | 1.67                | 3.49          | 48.97        | 1.24                                | 2.58          | 43.69        | 2.74     |
| 7      | 1.64                | 3.42          | 52.39        | 1.16                                | 2.41          | 46.10        | 3.41     |
| 8      | 1.41                | 2.94          | 55.33        | .97                                 | 2.02          | 48.12        | 2.78     |
| 9      | 1.31                | 2.74          | 58.07        | .86                                 | 1.79          | 49.91        | 2.89     |

The above results led to a preliminary decision to retain the nine-factor structure. However, further investigation revealed that four items loaded substantially on to more than one factor, and thus were deleted. Further, two factors appeared to contain only two items, below the



minimum-3 criterion. As a result, the instrument eventually yielded a *seven-factor structure*, corresponding largely with the originally-specified scales (domains), but with a few modifications. Table 9 displays results from the rotated (promax) seven-factor solution with item saliency for each factor. Due to limited space, item numbers are indicated. The item content can be found in Appendix D, the English version of the survey questionnaire.

Generally speaking, as seen in Table 9, salient items had pattern coefficients well above the .30 criterion. The structure coefficients are generally consistent with the pattern coefficients, with minor exceptions. Field (2005) explained that if several variables loaded highly onto more than one factors in the structure matrix, this is due to the relationship between factors. Field's explanation is reflected in Table 10, in which, for example, Perceived Control by the Process-based Components of Teacher Evaluation Policy had relatively high correlations with Perceived Support from Colleagues and Professional Development Programs and Self-efficacy in Practicing SCP respectively, and vice versa. Self-efficacy in Practicing SCP had relatively high correlations with Perceived Support from Colleagues and Professional Development Programs and General Beliefs Regarding SCP. Beliefs Regarding Teacher's Role in SCP had relatively high correlations with Perceived Control by the Output-based Components of Teacher Evaluation Policy. All of these inter-factor correlations were either explicitly or implicitly expected by the literature surveyed in Chapter II, as shown in the conceptual framework in Figure 1.

Table 9

## Results of Promax-Rotated Factor Pattern Showing Standardized Loading and Structure Coefficients (in Parenthesis)

| Theoretical Domain/Factor                          | Empirical Domain/Factor   | Item                              | Non-Gaokao  | Support         | Self-Efficacy | SCP             | Teacher's Role | General Belief | Gaokao    |
|--|---|-----------------------------------|---|-----------------|---------------|-----------------|----------------|----------------|-----------|
| Perceived Control by the Teacher Evaluation Policy | Perceived Control by the Process-based Components of Teacher Evaluation Policy (Non-Gaokao-related) | Q53                               | <b>85*</b> (81)   | -19 (20)        | -02 (26)      | 00 (00)         | -19 (-17)      | 09 (20)        | -02 (-03) |
|  |   | Q51                               | <b>82*</b> (76)   | 02 (18)         | -14 (14)      | -02 (03)        | 02 (08)        | 13 (19)        | 14 (18)   |
|  |   | Q55                               | <b>72*</b> (74)   | -05 (28)        | 09 (30)       | -07 (-06)       | -20 (-12)      | 07 (14)        | -04 (-04) |
|  |   | Q49                               | <b>70*</b> (73)   | 08 (29)         | 08 (24)       | -07 (-11)       | 09 (-11)       | -10 (10)       | 12 (08)   |
|  |   | Q52                               | <b>70*</b> (66)   | 02 (23)         | -20 (08)      | 15 (10)         | 02 (03)        | 02 (12)        | -01 (-01) |
|  |   | Q48                               | <b>68*</b> (75)   | 07 (40)         | -08 (28)      | 21 (-21)        | 21 (12)        | -11 (-01)      | -03 (-07) |
|  |   | Q56                               | <b>53*</b> (65)   | 07 (48)         | 12 (47)       | 00 (-04)        | 00 (-11)       | 03 (19)        | -13 (-28) |
|  |   | Q54                               | <b>50*</b> (63)   | 08 (36)         | 07 (20)       | -10 (-07)       | -04 (-06)      | -16 (-04)      | 18 (08)   |
|  |   | Q45                               | <b>43*</b> (40)   | 05 (07)         | -06 (14)      | 13 (20)         | 11 (29)        | 04 (16)        | 09 (33)   |
|  |   | Perceived Support for SCP Reforms | Perceived Support from Colleagues & Professional Development Programs | Q16             | -11 (17)      | <b>80*</b> (64) | -13 (19)       | 02 (02)        | 11 (10)   |
| Q15  | 16 (37)   |                                   |   | <b>73*</b> (72) | 01 (29)       | 19 (12)         | -18 (-20)      | 00 (00)        | -09 (-21) |
| Q18  | -04 (24)  |                                   |   | <b>70*</b> (80) | -02 (32)      | -04 (-10)       | 04 (-16)       | 00 (00)        | -11 (-36) |
| Q17  | 00 (24)   |                                   |   | <b>45*</b> (51) | 07 (31)       | 03 (00)         | 05 (00)        | -07 (01)       | 11 (-03)  |

| Theoretical Domain/Factor       | Empirical Domain/Factor                 | Item | Non-Gaokao | Support   | Self-Efficacy   | SCP             | Teacher's Role  | General Belief | Gaokao    |
|---------------------------------|---|------|------------|-----------|-----------------|-----------------|-----------------|----------------|-----------|
| Self-efficacy in Practicing SCP | Self-efficacy in Practicing SCP         | Q33  | -04 (17)   | 03 (27)   | <b>72*</b> (65) | 00 (18)         | 04 (05)         | -06 (24)       | -06 (-10) |
|                                 |   | Q34  | -11 (18)   | 09 (29)   | <b>67*</b> (71) | -06 (09)        | 03 (09)         | 06 (25)        | -07 (-07) |
|                                 |   | Q38  | -07 (23)   | -18 (13)  | <b>61*</b> (64) | 08 (30)         | 00 (06)         | 05 (40)        | 03 (-03)  |
|                                 |   | Q37  | 06 (12)    | -02 (22)  | <b>61*</b> (64) | 14 (27)         | -02 (09)        | -12 (24)       | 21 (14)   |
|                                 |   | Q36  | 18 (25)    | -13 (10)  | <b>55*</b> (56) | 14 (30)         | 02 (12)         | 04 (38)        | -06 (-02) |
|                                 |   | Q39  | -03 (23)   | 27 (37)   | <b>45*</b> (47) | -05 (12)        | -03 (-02)       | 12 (24)        | 05 (-11)  |
| Teachers' Implementation of SCP | Teachers' Implementation of SCP         | Q42  | -02 (08)   | -15 (02)  | <b>35*</b> (46) | 15 (39)         | -05 (09)        | 22 (48)        | 00 (10)   |
|                                 |   | Q57  | -17 (-05)  | 05 (09)   | 09 (38)         | <b>65*</b> (66) | -03 (04)        | 00 (30)        | 15 (18)   |
|                                 |   | Q60  | -09 (00)   | 00 (04)   | -02 (28)        | <b>56*</b> (63) | -08 (-06)       | 20 (42)        | 12 (08)   |
|                                 |   | Q64  | 00 (10)    | 14 (21)   | -03 (21)        | <b>44*</b> (54) | -01 (00)        | 08 (28)        | -12 (-02) |
|                                 |   | Q63  | 17 (19)    | 09 (19)   | 00 (22)         | <b>37*</b> (39) | 12 (11)         | -03 (17)       | -08 (01)  |
|                                 |   | Q25  | -05 (02)   | 13 (02)   | -12 (12)        | 01 (12)         | <b>81*</b> (82) | 06 (20)        | 08 (31)   |
| Beliefs Regarding SCP           | Beliefs Regarding Teacher's Role in SCP | Q26  | -01 (-02)  | -05 (-13) | 02 (07)         | 02 (17)         | <b>74*</b> (71) | 05 (19)        | -10 (17)  |
|                                 |   | Q28  | -05 (00)   | -03 (-10) | 16 (21)         | -02 (19)        | <b>61*</b> (66) | 08 (25)        | 10 (31)   |

| Theoretical Domain/Factor                          | Empirical Domain/Factor  | Item | Non-Gaokao | Support   | Self-Efficacy | SCP      | Teacher's Role | General Belief  | Gaokao          |
|--|--|------|------------|-----------|---------------|----------|----------------|-----------------|-----------------|
| Beliefs Regarding SCP                              | General Beliefs Regarding SCP  | Q31  | 01 (13)    | 09 (-02)  | 02 (30)       | 00 (30)  | 11 (28)        | <b>79*</b> (76) | 06 (13)         |
|  |  | Q29  | 07 (12)    | -04 (-07) | -12 (24)      | 08 (28)  | 03 (18)        | <b>62*</b> (63) | -04 (14)        |
|  |  | Q30  | -06 (09)   | 04 (02)   | 03 (29)       | 04 (28)  | -03 (09)       | <b>55*</b> (59) | 09 (12)         |
|  |  | Q32  | 17 (29)    | -11 (06)  | 08 (34)       | -19 (00) | 23 (21)        | <b>45*</b> (49) | -14 (-08)       |
| Perceived Control by the Teacher Evaluation Policy | Perceived Control by the Output-based Components of Teacher Evaluation Policy (Gaokao-Related) | Q50  | 04 (07)    | 00 (-16)  | -04 (-05)     | 03 (17)  | -04 (17)       | 05 (09)         | <b>74*</b> (68) |
|  |  | Q44  | 05 (10)    | 06 (-14)  | 14 (10)       | 08 (25)  | 11 (40)        | -02 (11)        | <b>70*</b> (76) |
|  |  | Q47  | 15 (20)    | 03 (-06)  | 00 (00)       | -03 (19) | 01 (23)        | 00 (11)         | <b>59*</b> (65) |

*Note.* Factor loadings > .3 are in boldface with asterisks. SCP = Student-Centered Pedagogy.

Table 10

*Inter-factor Correlation Matrix*

| Factor  | 1.  | 2    | 3    | 4   | 5   | 6   | 7 |
|---|-----|------|------|-----|-----|-----|---|
| 1. Perceived Control by the Output-based Components of Teacher Evaluation Policy  | —   |      |      |     |     |     |   |
| 2. Perceived Support from Colleagues and Professional Development Programs        | .38 | —    |      |     |     |     |   |
| 3. Self-efficacy in Practicing SCP  | .32 | .38  | —    |     |     |     |   |
| 4. Teachers' Implementation of SCP  | .00 | -.04 | .28  | —   |     |     |   |
| 5. Beliefs Regarding Teacher's Role in SCP  | .02 | -.13 | .09  | .15 | —   |     |   |
| 6. General Beliefs Regarding SCP  | .15 | -.01 | .43  | .38 | .17 | —   |   |
| 7. Perceived Control by the Process-based Components of Teacher Evaluation Policy | .01 | -.26 | -.06 | .21 | .33 | .08 | — |

*Note.* SCP = Student-Centered Pedagogy.

Table 9 also revealed that the theoretical conceptualization of some of the domains needed revision based on the empirical factor structures. Domain 3.0 (Self-efficacy in Practicing SCP) was validated as designed. Domain 1.0 (Perceived Support for SCP Reforms) had fewer items. Items from Domain 2.0 (Beliefs Regarding SCP) comprised two new factors (Beliefs Regarding Teacher's Role in SCP and General Beliefs Regarding SCP). Items from Domain 4.0 (Perceived Control by Teacher Evaluation Policy) comprised two new factors: Perceived Control by the Output-based Components of Teacher Evaluation Policy (*Gaokao*-related) and Perceived Control by the Process-based Components of Teacher Evaluation Policy (Non-*Gaokao*-related). Items from Domain 5.0 (Teachers' Implementation of SCP) comprised fewer items than originally specified.

Of direct interest to the present study, the most interesting EFA results were validation of Domain 4.0 (Perceived Control by the Teacher Evaluation Policy), which now contained two different scales consistent with the organizational theory literature on *output* versus *behavior* control mechanisms (Ouchi, 1977). The two scales were thus renamed as: a) Perceived Control by the Output-based Components of Teacher Evaluation Policy (*Gaokao*-related), and b) Perceived Control by the Process-based Components of Teacher Evaluation Policy (Non-*Gaokao*-related). The first scale included three items: students' college admission rates (Q50), Students' test scores on standardized tests such as *Gaokao* (Q44), and students' scores from other locally-administered standardized tests (Q47). The rest of the items were lumped into the other scale. The three items were strictly and directly related to *Gaokao* because standardized tests employed by local educational agencies are commonly viewed as the "simulations of *Gaokao*" by schools.

Students' graduation rates (Q52) did not load onto the *Gaokao*-related factor probably because in China, students' graduation is determined by a special exam called "graduation test," not *Gaokao*. The rest of the items comprised the factor on teaching processes and behaviors (non-*Gaokao*-related). Table 11 sums up the modified factor structure, including new titles and short labels of the empirically-supported factors and items retained. The present study used the short labels of the empirically-supported factors in the rest of the document and chapters.

Table 11

*Modified Survey Domains based on Empirical Factor Structure: Items Defining Scales*

| Theoretically-Supported Domain/Factor              | Empirically-Supported Domain/Factor  | Short Labels of the Empirically-Supported Domain/Factor | Items Included                              |
|--|--|---|---|
| Perceived Support for SCP Reforms                  | Perceived Support from Colleagues and Professional Development Programs                                      | Support   | Q15, Q16, Q17, Q18                          |
| Beliefs in SCP                                     | Beliefs Regarding Teacher's Role in SCP  | Teacher's Role  | Q25, Q26, Q28                               |
|  | General Beliefs Regarding SCP  | General Beliefs   | Q29, Q30, Q31, Q32                          |
| Self-Efficacy in Practicing SCP                    | Self-Efficacy in Practicing SCP  | Self-efficacy   | Q33, Q34, Q37, Q38, Q39, Q42                |
| Perceived Control by the Teacher Evaluation Policy | Perceived Control by the Process-based Components of Teacher Evaluation Policy (Non- <i>Gaokao</i> -Related) | Control by the Process-based Components                 | Q45, Q48, Q49, Q51, Q52, Q53, Q54, Q55, Q56 |
|  | Perceived Control by the Output-based Components of the Teacher Evaluation Policy ( <i>Gaokao</i> -Related)  | Control by the Output-based Components                  | Q44, Q47, Q50                               |
| Teachers' Implementation of SCP                    | Teachers' Implementation of SCP  | SCP Implementation                                      | Q57, Q60, Q63, Q64                          |

**Final Scales: Descriptive Statistics and Reliability Estimates**

Following the EFA, descriptive analyses were performed on each of the factor-defined scales. For descriptive statistics, the study looked at: mean, maximum, minimum, and standard deviation. For distribution statistics, the study looked at: skewness and kurtosis. Cronbach's

*Alpha* (Chatterji, 2003; Crocker & Algina, 2006) was calculated as well for each factor to help further determine homogeneity of items. The summary statistics are shown in Table 12.

Table 12

*Descriptive Statistics and Cronbach's Alpha Estimates for Seven Survey Scales*

| Domains                                 | # of Items | Minimum | Maximum | M     | SD   | Skewness | Kurtosis | Cronbach's $\alpha$ |
|---|------------|---------|---------|-------|------|----------|----------|---------------------|
| Support                                 | 4          | 6       | 19      | 13.49 | 2.65 | -.531    | .023     | .77                 |
| Self-Efficacy                           | 6          | 12      | 30      | 22.48 | 2.92 | -.238    | .526     | .72                 |
| General Beliefs                         | 4          | 11      | 20      | 16.03 | 2.00 | -.195    | .223     | .65                 |
| Teacher's Role                          | 3          | 5       | 15      | 12.26 | 1.68 | -.959    | 2.862    | .65                 |
| Control by the Process-based Components | 3          | 6       | 15      | 11.72 | 2.19 | -.181    | -.704    | .70                 |
| Control by the Output-based Components  | 9          | 10      | 44      | 33.23 | 5.88 | -.985    | 1.634    | .87                 |
| SCP Implementation                      | 4          | 7       | 16      | 11.14 | 2.08 | .453     | -.179    | .68                 |

*Note.* M = scale mean; SD = scale standard deviation.

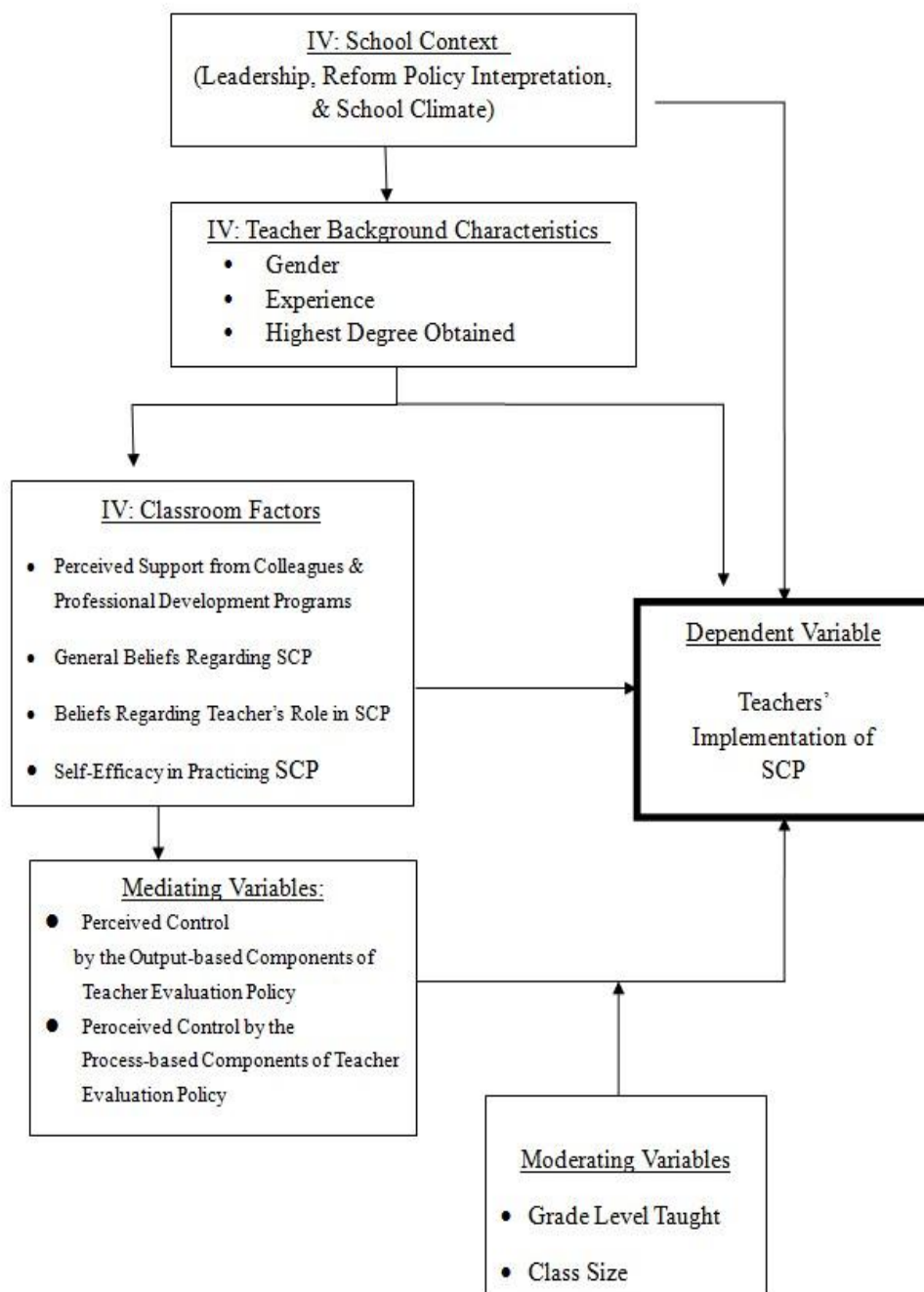
Most of the Cronbach's alpha values were above the .70 criterion except for three scales, indicating homogenous items under the same domain. Although three factors had an alpha estimates that was just below the .70 criterion, the rounded values approached .70.

### **Revised Conceptual Framework Based on EFA Results**

Based on EFA results, the revised conceptual framework is presented in Figure 2 with labels of the validated and reliable construct measures representing variables. The scale titles have been revised per Table 11.



Figure 2.0 Revised Conceptual Framework with Variables Tapped by Scales



## Data Analysis Plan

### Differences by Grade and Class Size

The study proposed to investigate two moderating variables in the key analyses: grade and class size (Hypothesis 6.0). To examine whether there were differences among teachers on the seven survey measures, the study began with a series of exploratory analyses. Means ( $M$ ), Standard Deviations ( $SD$ ), and independent-samples  $t$ -tests were performed, followed by multiple group comparisons. For differences by Grade, since the comparison was conducted three times with each survey construct measure serving as the dependent variable in these analyses, a Bonferroni adjustment procedure was applied, with the  $p$  value of .05 divided by 3. For differences by Class Size, the comparison was conducted only once for each construct measure. No adjustment procedure was applied. The  $p$  value for statistical significance remained at the .05 level.

### Bivariate Correlations among Survey Measures

Initially, Pearson correlations were obtained to examine relationships between pairs of the seven survey construct measures. This analysis was also exploratory and descriptive in nature, and used the entire sample. The purpose was to examine whether the direction and magnitude of the relationships were consistent with the literature review.

### Coding Procedure for the School Variable

The school context factor, in this study, is a combined contrast-coded variable encompassing differences on a range of school-level variables relevant for SCP implementation in the classroom, such as leadership style, policy interpretation, school climate and so on.

Socioeconomic data on students and other school-related variables were not readily available in Jingyang District, and could not be formally measured. The differences at the school level are thus collectively represented as a categorical variable with 7 levels. Since it is a multi-level categorical variable, Helmert coding was employed (Wendorf, 2004) to test if SCP implementation of any school is significantly different from other schools due to variability at schools.

### **Hierarchical Multiple Regression Models**

To test the series of hypotheses formulated based on the revised conceptual framework given in Figure 2, the selected school and teacher factors and SCP-relevant independent variables were incrementally added in regression models to explain the variance in the criterion (dependent) variable, SCP Implementation. The order of variables entered was: School, teacher background characteristics (Gender, Experience and Highest Degree Obtained), Perceived Support, General Beliefs, Teacher's Role, Self-efficacy, Control by the Output-based Components, and Control by the Process-based Components.

For regression models, the statistical significance of the overall model  $F$  was examined at the .05 alpha level.  $R$ -Squared and  $R$ -Squared changes, and individual standardized  $\beta$  values, were examined to interpret the magnitude and direction of the variable relationships against expectations from the literature review. Variables that were statistically non-significant in a given regression analysis were dropped from following models.

Stemming from the conceptual framework, the regression equations used to test the series of hypotheses were as follows. Each research question is now followed by the hypothesis statement, and corresponding equation.

**Research question 1.0:** To what extent do *school* characteristics, as predicted by educational reform and policy implementation literature, affect SCP implementation at the classroom level as reported by teachers?

**Hypothesis 1.0:** Differences in *school* context (e.g., leadership style, policy interpretation, and school climate) will significantly and substantially predict levels of classroom implementation of SCP.

Model 1:

$$Y_{scp} = \beta_0 + \beta_1 X_{sch1} + \beta_2 X_{sch2} + \beta_3 X_{sch3} + \beta_4 X_{sch4} + \beta_5 X_{sch5} + \beta_6 X_{sch6} + \varepsilon$$

**Research question 2.0:** To what extent do selected *teacher characteristics*, as predicted by educational reform and policy implementation literature, affect SCP implementation in the classroom as reported by teachers, after taking school-level differences into account?

**Hypothesis 2.0:** Controlling for school level variability, *teacher background characteristics* such as Gender, Experience, and Highest Degree Obtained will significantly and substantially predict levels of classroom implementation of SCP.

Model 2:

$$Y_{scp} = \beta_0 + \beta_1 X_{sch1} + \beta_2 X_{sch2} + \beta_3 X_{sch3} + \beta_4 X_{sch4} + \beta_5 X_{sch5} + \beta_6 X_{sch6} + \beta_7 X_{exp} + \beta_8 X_{degree} + \beta_9 X_{gender} + \varepsilon$$

**Research question 3.0:** To what extent do selected SCP-relevant constructs (*Teacher Perceived Support, Teacher Beliefs, and Teacher Self-efficacy in Practicing SCP*), as predicted by the educational reform and policy implementation literature, affect SCP implementation in the classroom reported by teachers, accounting for school and teacher characteristics?

**Hypothesis 3.0:** Controlling for school level variability and teacher background characteristics, *Teachers' Perceived Support for SCP Reforms, Beliefs in SCP, and Teachers' Self-efficacy in Practicing SCP* will significantly and substantially predict levels of classroom implementation of SCP.

Model 3:

$$Y_{scp} = \beta_0 + \beta_1 X_{sch1} + \beta_2 X_{sch2} + \beta_3 X_{sch3} + \beta_4 X_{sch4} + \beta_5 X_{sch5} + \beta_6 X_{sch6} + \beta_7 X_{exp} + \beta_8 X_{degree} + \beta_9 X_{gender} + \beta_{10} X_{supt} + \varepsilon$$

Model 4:

$$Y_{scp} = \beta_0 + \beta_1 X_{sch1} + \beta_2 X_{sch2} + \beta_3 X_{sch3} + \beta_4 X_{sch4} + \beta_5 X_{sch5} + \beta_6 X_{sch6} + \beta_7 X_{exp} + \beta_8 X_{degree} + \beta_9 X_{gender} + \beta_{10} X_{supt} + \beta_{11} X_{role} + \beta_{12} X_{belief} + \varepsilon$$

Model 5:

$$Y_{scp} = \beta_0 + \beta_1 X_{sch1} + \beta_2 X_{sch2} + \beta_3 X_{sch3} + \beta_4 X_{sch4} + \beta_5 X_{sch5} + \beta_6 X_{sch6} + \beta_7 X_{exp} + \beta_8 X_{degree} + \beta_9 X_{gender} + \beta_{10} X_{supt} + \beta_{11} X_{role} + \beta_{12} X_{belief} + \beta_{13} X_{effi} + \varepsilon$$

**Research question 4.0:** To what extent is *control* enforced by the organization's *Gaokao*-related teacher evaluation policy for secondary teachers, a negative predictor and a significant mediating variable for SCP implementation in classrooms, as reported by teachers?

**Hypothesis 4.0:** Controlling for school level variability, teacher background characteristics, and SCP-relevant factors, *Teachers' Perceived Levels of Control by the Teacher Evaluation Policy* will significantly and substantially predict levels of classroom implementation of SCP.

Model 6:

$$\begin{aligned}
 Y_{scp} = & \beta_0 + \beta_1 X_{sch1} + \beta_2 X_{sch2} + \beta_3 X_{sch3} + \beta_4 X_{sch4} + \beta_5 X_{sch5} + \beta_6 X_{sch6} + \beta_7 X_{exp} \\
 & + \beta_8 X_{degree} + \beta_9 X_{gender} + \beta_{10} X_{supt} + \beta_{11} X_{role} + \beta_{12} X_{belief} + \beta_{13} X_{effi} \\
 & + \beta_{14} X_{gk} + \beta_{15} X_{nongk} + \varepsilon
 \end{aligned}$$

**Research question 5.0:** To what extent is the relationship between *control* enforced by the organization's *Gaokao*-related teacher evaluation policy for secondary teachers, with reported SCP implementation levels in the classroom, moderated by *Grade Level* and *Class Size*?

**Hypothesis 5.0:** The relationship between teachers' perceived levels of control by the teacher evaluation policy and levels of classroom implementation of SCP will be moderated by *Grade Level Taught* by teachers.

Model 7:

$$Y_{scp} = \beta_0 + \beta_1 X_{sch1} + \beta_2 X_{sch2} + \beta_3 X_{sch3} + \beta_4 X_{sch4} + \beta_5 X_{sch5} + \beta_6 X_{sch6} + \beta_7 X_{exp} \\ + \beta_8 X_{degrees} + \beta_9 X_{gender} + \beta_{10} X_{supt} + \beta_{11} X_{role} + \beta_{12} X_{belief} + \beta_{13} X_{effi} \\ + \beta_{14} X_{gk} + \beta_{15} X_{nongk} + \beta_{16} X_{grd} + \beta_{17} X_{gk*grd} + \varepsilon$$

**Hypothesis 6.0:** The relationship between teachers' perceived levels of control by the teacher evaluation policy and teachers' levels of implementation of SCP will be moderated by *Class Size*.

Model 8:

$$Y_{scp} = \beta_0 + \beta_1 X_{sch1} + \beta_2 X_{sch2} + \beta_3 X_{sch3} + \beta_4 X_{sch4} + \beta_5 X_{sch5} + \beta_6 X_{sch6} + \beta_7 X_{exp} \\ + \beta_8 X_{degree} + \beta_9 X_{gender} + \beta_{10} X_{supt} + \beta_{11} X_{role} + \beta_{12} X_{belief} + \beta_{13} X_{effi} \\ + \beta_{14} X_{gk} + \beta_{15} X_{nongk} + \beta_{16} X_{cls} + \beta_{17} X_{gk*cls} + \varepsilon$$



### Testing Statistical Assumptions in the Data Set

The study used standardized residuals to identify outliers. If outliers exceeded the acceptable limits, the cases were excluded. Using criteria derived from the rule of normality (Field, 2005), the study expected to have approximately 12 cases (5%) of standardized residuals outside the  $\pm 2$  limits and 3 cases (1%) outside of the  $\pm 2.5$  limits. Corresponding SPSS outputs (Appendix F) show that 14 cases in the sample of the study lie outside of the  $\pm 2$  limits and 1 (case 122) lies outside of the  $\pm 2.5$  limits. Therefore, the sample of the study conformed with assumptions for regression analyses in terms of outliers.

Other assumptions examined were: (a) independent errors, (b) multicollinearity, (c) independence of the observed outcome variable, (d) linearity, and (e) normally distributed errors. All the five basic assumptions were met.

The Durbin-Watson statistic was used to determine whether the assumption of independent errors is tenable. According to Field (2005), the Durbin-Watson test statistics can vary between 0 and 4 with a value of 2 meaning that the residuals are uncorrelated. The Durbin-Watson value of the present study was 1.94, which is very close to 2, indicating that the residuals (error) in the model are independent.

With respect to multicollinearity issues, diagnosis results revealed that although some of the factors are inter-correlated, there was no substantial evidence for multicollinearity. The tolerance value of each factor was above the .2 threshold (Menard, 1995), indicating that collinearity is not a problem for this study.

The \*ZPRED/\*ZRESID graph reported by SPSS helped to check both the assumption of the normality of residuals and the assumption of linearity (see Appendix G). As shown in the figure, the residual dots are randomly and evenly dispersed throughout the plot.

## CHAPTER IV RESULTS

This chapter now reports the study's results in accordance with research hypotheses and overall conceptual framework. To start, the chapter presents results of the exploratory analyses with respect to differences by Grade and Class Size and the matrix of bivariate correlations among variables. Following that, the results from the series of multiple-regression models are presented with reference to specific research questions and hypotheses.

### **Grade Level Differences on Survey Construct Measures**

Tables 13-19 display results of differences by Grade on the seven survey measures. Overall, descriptive statistics reveal a consistent pattern for each variable, with the group means very close in value by grade, as are the standard deviations. Independent samples *t*-tests confirmed that there were no statistically significant differences. None of the *p* values in the parenthesis were below .016, the *p* value adjusted for multiple comparisons. The exception was a difference between Grade 11 and 12 teachers on Beliefs Regarding Teachers' Role in SCP. The difference was statistically significant at the .01 level.

Table 13

*Grade Level Differences: Perceived Support from Colleagues and Professional Development Programs*

|          | N  | Mean  | SD   | Mean Difference in Perceived Support |                             |                             |
|----------|----|-------|------|--------------------------------------|-----------------------------|-----------------------------|
|          |    |       |      | Grade 10                             | Grade 11                    | Grade 12                    |
|          |    |       |      | <i>t</i> value ( <i>p</i> )          | <i>t</i> value ( <i>p</i> ) | <i>t</i> value ( <i>p</i> ) |
| Grade 10 | 58 | 14.83 | 2.21 | —                                    |                             |                             |
| Grade 11 | 71 | 12.77 | 2.30 | 2.05 (.82)                           | —                           |                             |
| Grade 12 | 87 | 13.21 | 2.72 | 1.62(.09)                            | .43 (.12)                   | —                           |

*Note.* *p* value required for statistical significance=.016

Table 14

*Grade Level Differences: General Beliefs Regarding SCP*

|          | N  | Mean  | SD   | Mean Difference in General Beliefs Regarding SCP |                             |                             |
|----------|----|-------|------|--|-----------------------------|-----------------------------|
|          |    |       |      | Grade 10   | Grade 11                    | Grade 12                    |
|          |    |       |      | <i>t</i> value ( <i>p</i> )                      | <i>t</i> value ( <i>p</i> ) | <i>t</i> value ( <i>p</i> ) |
| Grade 10 | 61 | 16.20 | 2.02 | —  |                             |                             |
| Grade 11 | 73 | 15.95 | 1.95 | .25 (.75)  | —                           |                             |
| Grade 12 | 87 | 16.00 | 2.05 | .20 (.73)  | .06 (.48)                   | —                           |

*Note.* *p* value required for statistical significance=.016

Table 15

*Grade Level Differences: Beliefs Regarding Teachers' Role in SCP*

|          | N  | Mean  | SD   | Mean Difference in Beliefs Regarding Teachers' Role in SCP |                             |                             |
|----------|----|-------|------|--|-----------------------------|-----------------------------|
|          |    |       |      | Grade 10   | Grade 11                    | Grade 12                    |
|          |    |       |      | <i>t</i> value ( <i>p</i> )                                | <i>t</i> value ( <i>p</i> ) | <i>t</i> value ( <i>p</i> ) |
| Grade 10 | 61 | 12.18 | 2.00 | —  |                             |                             |
| Grade 11 | 74 | 12.59 | 1.10 | .41 (.04)  | —                           |                             |
| Grade 12 | 87 | 12.00 | 1.77 | .18 (.97)  | .59 (.01**)                 | —                           |

Note. *p* value required for statistical significance=.016

\*\*  $p < .016$

Table 16

*Grade Level Differences: Self-Efficacy in Practicing SCP*

|          | N  | Mean  | SD   | Mean Difference in Self-Efficacy |                             |                             |
|----------|----|-------|------|----------------------------------|-----------------------------|-----------------------------|
|          |    |       |      | Grade 10                         | Grade 11                    | Grade 12                    |
|          |    |       |      | <i>t</i> value ( <i>p</i> )      | <i>t</i> value ( <i>p</i> ) | <i>t</i> value ( <i>p</i> ) |
| Grade 10 | 61 | 22.77 | 2.99 | —                                |                             |                             |
| Grade 11 | 74 | 22.20 | 2.67 | .57 (.55)                        | —                           |                             |
| Grade 12 | 86 | 22.44 | 2.96 | .33 (.95)                        | .24 (.46)                   | —                           |

Note. *p* value required for statistical significance=.016

Table 17

*Grade Level Differences: Perceived Control by the Output-based Components of Teacher Evaluation Policy (Gaokao-related)*

|          | N  | Mean  | SD   | Mean Difference in Perceived Control by the Output-based Components |   |   |
|----------|----|-------|------|---|---|---|
|          |    |       |      | Grade 10<br><i>t</i> value ( <i>p</i> )                             | Grade 11<br><i>t</i> value ( <i>p</i> ) | Grade 12<br><i>t</i> value ( <i>p</i> ) |
| Grade 10 | 61 | 11.77 | 2.16 | —   |   |   |
| Grade 11 | 73 | 12.09 | 2.17 | .32 (.83)   | —                                       |   |
| Grade 12 | 87 | 11.35 | 2.24 | -.42 (.40)  | -.75(.51)                               | —                                       |

*Note.* *p* value required for statistical significance=.016

Table 18

*Grade Level Differences: Perceived Control by the Process-based Components of Teacher Evaluation Policy (Non-Gaokao-Related)*

|          | N  | Mean  | SD   | Mean Difference in Perceived Control by the Process-based Components |   |   |
|----------|----|-------|------|--|---|---|
|          |    |       |      | Grade 10<br><i>t</i> value ( <i>p</i> )                              | Grade 11<br><i>t</i> value ( <i>p</i> ) | Grade 12<br><i>t</i> value ( <i>p</i> ) |
| Grade 10 | 61 | 32.93 | 6.13 | —  |   |   |
| Grade 11 | 73 | 33.77 | 5.56 | .84(.55)   | —                                       |   |
| Grade 12 | 87 | 33.12 | 5.37 | .18 (.30)  | -.65(.64)                               | —                                       |

*Note.* *p* value required for statistical significance=.016

Table 19

*Grade Level Differences: Teachers' Implementation of SCP*

|          | N  | Mean  | SD   | Mean Difference in SCP Implementation |                             |                             |
|----------|----|-------|------|---------------------------------------|-----------------------------|-----------------------------|
|          |    |       |      | Grade 10                              | Grade 11                    | Grade 12                    |
|          |    |       |      | <i>t</i> value ( <i>p</i> )           | <i>t</i> value ( <i>p</i> ) | <i>t</i> value ( <i>p</i> ) |
| Grade 10 | 61 | 11.44 | 1.93 | —                                     |                             |                             |
| Grade 11 | 73 | 11.05 | 1.95 | .39 (.82)                             | —                           |                             |
| Grade 12 | 87 | 10.78 | 2.10 | .66 (.58)                             | .27 (.73)                   | —                           |

*Note.* *p* value required for statistical significance=.016

### **Class Size Differences on Survey Construct Measures**

Tables 20-26 display results of differences by Class Size on the seven survey measures. Overall, descriptive statistics reveal a consistent pattern for each variable, with the group means very close in value between larger ( $\geq 50$ ) and smaller ( $<50$ ) classes, as are the standard deviations. Independent samples *t*-tests confirmed that there were no statistically significant differences. None of the *p* values in the parenthesis were below .05. The exception was a difference on Perceived Control by the Output-based Components of Teacher Evaluation Policy. The mean difference between larger and smaller classes was 1.04 and it was statistically significant at the .01 level.

Table 20

*Class Size Differences: Perceived Support from Colleagues and Professional Development Programs*

|                         | N   | Mean  | SD   | Mean Difference in Perceived Support<br><i>t</i> value ( <i>p</i> ) |
|-------------------------|-----|-------|------|---|
| Big Class ( $\geq 50$ ) | 175 | 13.50 | 2.60 | —   |
| Small Class ( $< 50$ )  | 32  | 13.91 | 2.18 | -.41 (.40)  |

*Note.* \*  $p < .05$ ; \*\*  $p < .01$ .

Table 21

*Class Size Differences: General Beliefs Regarding SCP*

|                         | N   | Mean  | SD   | Mean Difference in General Beliefs<br><i>t</i> value ( <i>p</i> ) |
|-------------------------|-----|-------|------|---|
| Big Class ( $\geq 50$ ) | 177 | 16.10 | 1.85 | —   |
| Small Class ( $< 50$ )  | 33  | 16.33 | 2.38 | -.23 (.52)  |

*Note.* \*  $p < .05$ ; \*\*  $p < .01$ .

Table 22

*Class Size Differences: Beliefs Regarding Teacher's Role in SCP*

|                         | N   | Mean  | SD   | Mean Difference in Teacher's Role<br><i>t</i> value ( <i>p</i> ) |
|-------------------------|-----|-------|------|--|
| Big Class ( $\geq 50$ ) | 178 | 12.25 | 1.66 | —  |
| Small Class ( $< 50$ )  | 33  | 12.42 | 1.37 | -.17 (.53)   |

Note. \*  $p < .05$ ; \*\*  $p < .01$ .

Table 23

*Class Size Differences: Self-efficacy in Practicing SCP*

|                         | N   | Mean  | SD   | Mean Difference in Self-efficacy<br><i>t</i> value ( <i>p</i> ) |
|-------------------------|-----|-------|------|---|
| Big Class ( $\geq 50$ ) | 177 | 22.54 | 2.77 | .03 (.96)   |
| Small Class ( $< 50$ )  | 33  | 22.52 | 3.03 | —   |

Note. \*  $p < .05$ ; \*\*  $p < .01$ .



Table 24

*Class Size Differences: Perceived Control by the Output-based Components of Teacher*

*Evaluation Policy (Gaokao-related)*

|                         | N   | Mean  | SD   | Mean Difference in Control by<br>the Output-based Components<br><br><i>t</i> value ( <i>p</i> ) |
|-------------------------|-----|-------|------|---|
| Big Class ( $\geq 50$ ) | 177 | 11.92 | 2.19 | 1.04 (.01**)  |
| Small Class ( $< 50$ )  | 33  | 10.88 | 1.71 | —   |

*Note.* \*  $p < .05$ ; \*\*  $p < .01$ .

Table 25

*Class Size Differences: Perceived Control by the Process-based Components of Teacher*

*Evaluation Policy (Non-Gaokao-related)*

|                         | N   | Mean  | SD   | Mean Difference in Control by<br>the Process-based Components<br><br><i>t</i> value ( <i>p</i> ) |
|-------------------------|-----|-------|------|--|
| Big Class ( $\geq 50$ ) | 177 | 33.59 | 5.52 | .38 (.71)  |
| Small Class ( $< 50$ )  | 33  | 33.21 | 5.42 | —  |

*Note.* \*  $p < .05$ ; \*\*  $p < .01$ .

Table 26

*Class Size Differences: Teachers' Implementation of SCP*

|                         | N   | Mean  | SD   | Mean Difference in SCP Implementation<br><i>t</i> value ( <i>p</i> ) |
|-------------------------|-----|-------|------|--|
| Big Class ( $\geq 50$ ) | 177 | 11.10 | 2.04 | .31 (.42)  |
| Small Class ( $< 50$ )  | 33  | 10.79 | 2.04 | —  |

Note. \*  $p < .05$ ; \*\*  $p < .01$ .

### Pearson Correlations

The Pearson correlations are displayed in Table 27. Generally speaking, the bivariate Pearson correlations among composite survey measures were consistent with the study's theoretical expectations and the inter-factor correlations based on EFA results per the revised conceptual framework in Figure 2. Based on literature review, the study expected factors to be inter-correlated. The direction of the relationships was also consistent with the study's assumptions. A positive and significant correlation exists between Self-efficacy and Support ( $r [220] = .310, p < .05$ ), and between General Beliefs and Beliefs Regarding Teacher's Role ( $r [228] = .188, p < .01$ ). Teachers' Perceived Control by the Process-based Components has a positive and significant relationships with Support ( $r [220] = .354, p < .05$ ), Self-efficacy ( $r [229] = .209, p < .05$ ), and General beliefs ( $r [227] = .171, p < .01$ ), respectively.

The most important information in Table 21 is contained in the last row, which shows the bivariate relationships between the study's dependent variable, SCP Implementation, and the other independent variables in the conceptual framework in Figure 2. Consistent with the study's

expectations, SCP Implementation has positive, significant relationships with three variables: Support ( $r[220] = .204, p < .05$ ), Self-efficacy ( $r [228] = .335, p < .05$ ), and General Beliefs ( $r [227] = .302, p < .01$ ). It has a negative, significant relationship with Perceived Control by the Output-based Components ( $r [228] = -.175, p < .05$ ).

Table 27

*Pearson-Correlation Matrix*

|   | 1      | 2      | 3     | 4      | 5      | 6      | 7     | 8      | 9    | 10     | 11   | 12 |
|---|--------|--------|-------|--------|--------|--------|-------|--------|------|--------|------|----|
| 1. Grade                                    | —      |        |       |        |        |        |       |        |      |        |      |    |
| 2. Degree                                   | .055   | —      |       |        |        |        |       |        |      |        |      |    |
| 3. Gender                                   | -.039  | -.084  | —     |        |        |        |       |        |      |        |      |    |
| 4. Experience                               | -.001  | -.179* | .241* | —      |        |        |       |        |      |        |      |    |
| 5. Class size                               | -.160* | .045   | .070  | .009   | —      |        |       |        |      |        |      |    |
| 6. Support                                  | -.093  | .007   | -.124 | -.043  | -.042  | —      |       |        |      |        |      |    |
| 7. Self-Efficacy                            | -.003  | .059   | .010  | .186** | .008   | .310*  | —     |        |      |        |      |    |
| 8. General Belief                           | -.015  | -.044  | -.036 | .051   | -.078  | .094   | .301* | —      |      |        |      |    |
| 9. Teachers' Role                           | -.120  | .034   | .049  | .015   | -.009  | -.034  | .052  | .188** | —    |        |      |    |
| 10. Control by the Output-based Components  | -.132* | .183*  | -.014 | -.008  | .253** | -.153* | -.107 | -.014  | .042 | —      |      |    |
| 11. Control by the Process-based Components | -.024  | .038   | -.087 | .037   | .002   | .354*  | .209* | .171** | .028 | .118   | —    |    |
| 12. SCP Implementation                      | -.109  | -.018  | -.090 | .081   | .088   | .204*  | .335* | .302** | .035 | -.175* | .057 | —  |

Note. \*  $p < .05$  (2-tailed).

\*\*  $p < .01$  (2-tailed)

## Influence of School Characteristics on SCP Implementation by Teachers

**Research question 1.0:** To what extent do *school* characteristics, as predicted by educational reform and policy implementation literature, affect SCP implementation at the classroom level as reported by teachers?

**Hypothesis 1.0:** Differences in *school* context (e.g., leadership style, policy interpretation, and school climate) will significantly and substantially predict levels of classroom implementation of SCP.

Table 28 displays results of the regression analysis for Hypothesis 1.0. Table 28 shows that none of the coded school context variables is statistically significant. Although Table 28 shows that the overall model explained nearly 5% of the variance in SCP Implementation ( $R^2 = .047$ ), the influence of school factors coded categorically was mostly likely obtained by chance,  $F(6,225)=1.832, p=.094$ .

Table 28

*Teachers' Reports of SCP Implementation Regressed on School Context*

| SCP Implementation      |                    |              |
|-------------------------|--------------------|--------------|
| Independent Variable    | $\beta$            | Std. $\beta$ |
| Coded School Variable 1 | .131               | .054         |
| Coded School Variable 2 | .107               | .039         |
| Coded School Variable 3 | .432               | .152         |
| Coded School Variable 4 | -.206              | -.057        |
| Coded School Variable 5 | .120               | .032         |
| Coded School Variable 6 | -.426              | -.109        |
| Constant                | 11.343             |              |
| <i>F</i>                | 1.832              |              |
|                         | ( <i>p</i> = .094) |              |
| <i>R</i> <sup>2</sup>   | .047               |              |

*Note.* Coded school variables should be interpreted as follows: School Variable 1 versus others influenced SCP implementation with a  $\beta$  of .131; School Variable 2 versus others influenced SCP implementation with a  $\beta$  of .107; and so on.

\*  $p < .05$ ; \*\*  $p < .01$ .

## Influences of Teacher Background Characteristics on SCP Implementation

**Research question 2.0:** To what extent do selected *teacher characteristics*, as predicted by educational reform and policy implementation literature, affect SCP implementation in the classroom as reported by teachers, after taking school-level differences into account?

**Hypothesis 2.0:** Controlling for school level variability, *teacher background characteristics* such as Gender, Experience, and Highest Degree Obtained will significantly and substantially predict levels of classroom implementation of SCP.

Table 29 displays results of the regression analyses of Hypothesis 2.0. Table 29 shows that controlling for school context characteristics, none of the teacher background characteristic variables are statistically significant predictors. As a matter of fact, SCP Implementation seems to be neither depending on school context factors nor on teacher background characteristics. The overall model is not significant  $F(9,222)=1.649, p = .103$ .

Table 29

*Teachers' Reports of SCP Implementation Regressed on Teacher Background Characteristics*

| Variable                | SCP Implementation |              |                    |              |
|-------------------------|--------------------|--------------|--------------------|--------------|
|                         | Model 1            |              | Model 2            |              |
|                         | $\beta$            | Std. $\beta$ | $\beta$            | Std. $\beta$ |
| School Variable 1       | .131               | .054         | .142               | .059         |
| School Variable 2       | .107               | .039         | .160               | .058         |
| School Variable 3       | .432               | .152         | .378               | .133         |
| School Variable 4       | -.206              | -.057        | -.105              | -.029        |
| School Variable 5       | .120               | .032         | .027               | .007         |
| School Variable 6       | -.426              | -.109        | -.412              | -.105        |
| Gender                  |                    |              | -.241              | -.113        |
| Experience              |                    |              | .026               | .096         |
| Highest Degree Obtained |                    |              | -.064              | -.017        |
| Constant                | 11.343             |              | 11.065             |              |
| <i>F</i>                | 1.832              |              | 1.649              |              |
|                         | ( <i>p</i> = .094) |              | ( <i>p</i> = .103) |              |
| <i>R</i> <sup>2</sup>   | .047               |              | .063               |              |
| $\Delta R^2$            |                    |              | .016               |              |

*Note.* Coded school variables should be interpreted as follows: School Variable 1 versus others influenced SCP implementation with a  $\beta$  of .131; School Variable 2 versus others influenced SCP implementation with a  $\beta$  of .107; and so on.

\* *p* < .05; \*\* *p* < .01.



### **Influences of Classroom Level SCP-Relevant Variables on SCP Implementation**

**Research question 3.0:** To what extent do selected SCP-relevant constructs (*Teacher Perceived Support, Teacher Beliefs, and Teacher Self-efficacy in Practicing SCP*), as predicted by the educational reform and policy implementation literature, affect SCP implementation in the classroom reported by teachers, accounting for school and teacher characteristics?

**Hypothesis 3.0:** Controlling for school level variability and teacher background characteristics, *Teachers' Perceived Support for SCP Reforms, Beliefs in SCP, and Teachers' Self-efficacy in Practicing SCP* will significantly and substantially predict levels of classroom implementation of SCP.

Hypothesis 3.0 was intended to test the influences of the selected SCP-relevant factors on teachers' implementation of SCP, controlling for both schools and teacher background factors. These factors were incrementally entered in the following order: 1) Support, 2) General Beliefs and Teacher's Role, and 3) Self-efficacy. Since none of the school context and teacher characteristic variables was statistically significant in the previous regression analyses, they were excluded from the hypothesis testing hereafter.

Table 30 displays results of the regression analyses of Hypothesis 3.0.

Table 30

*Teachers' Reports of SCP Implementation Regressed on SCP-Relevant Variables*

| Variable              | SCP Implementation |              |                   |              |                   |              |
|-----------------------|--------------------|--------------|-------------------|--------------|-------------------|--------------|
|                       | Model 3            |              | Model 4           |              | Model 5           |              |
|                       | $\beta$            | Std. $\beta$ | $\beta$           | Std. $\beta$ | $\beta$           | Std. $\beta$ |
| Support               | .161**             | .201         | .140**            | .174         | .090              | .112         |
| General Beliefs       |                    |              | .301**            | .287         | .231**            | .221         |
| Teacher's Role        |                    |              | -.016             | -.013        | -.016             | -.011        |
| Self-Efficacy         |                    |              |                   |              | .169**            | .237         |
| Constant              | 8.967              |              | 4.622             |              | 2.434             |              |
| <i>F</i>              | 9.667**            |              | 10.457**          |              | 15.335**          |              |
|                       | ( <i>p</i> =.002)  |              | ( <i>p</i> =.000) |              | ( <i>p</i> =.000) |              |
| <i>R</i> <sup>2</sup> | .040               |              | .121              |              | .168              |              |
| $\Delta R^2$          |                    |              | .081              |              | .047              |              |

Note. \*  $p < .05$ ; \*\*  $p < .01$ .

Table 30 shows that as the four factors were added into the regression model in the aforementioned order, the model's explanatory power on the variance of the criterion variable (SCP Implementation) increased. This is reflected by the change of both the *F* and the *R squared* value across the four models. In Model 3 where only the factor Support was entered as the predictor, *F* value is significant at the .01 level,  $F(1,230)=9.667$ ,  $p = .002$ . The variable itself is a significant predictor of teachers' implementation of SCP,  $\beta = .161$ ,  $t(230)=3.109$ ,  $p=.002$ , and explained about 4% of the variance in SCP Implementation ( $R^2 = .040$ ).

In Model 4, where two more SCP-relevant factors were entered (General Beliefs and Teacher's Role), the model is again significant,  $F(3,228) = 10.457, p=.000$ . *R squared* value increased from .040 to .121, indicating that 12% variance in teachers' implementation of SCP can now be explained by the predictors. Except for Teacher's Role, the other two SCP-relevant factors were statistically significant at the .01 level. General Beliefs explained an added approximately 8% variance in SCP implementation ( $\Delta R^2 = .081$ ).

As the fourth SCP-relevant factor Self-efficacy was entered into the model, results are again significant,  $F(3,228) = 15.335, p=.000$ . The explanatory capacity increased from .121 to .168, indicating that an additional 5% variance in teachers' Implementation of SCP was now explained by Self-efficacy. However, Support became insignificant in Model 5. Taken individually, two variables turned out to be significant predictors of teachers' implementation of SCP: General Beliefs,  $\beta = .231, t(228)=3.488, p=.001$  and Self-efficacy,  $\beta = .169, t(228)=3.593, p=.000$ .

Overall, in Model 5, Self-efficacy had a slightly higher standardized  $\beta$  (.237) compared to General Beliefs (.221), which means that in comparison, Self-efficacy exerts slightly bigger influence on teachers' implementation of SCP. One unexpected result was the negative correlation between Teacher's Role and SCP Implementation, but the relationship was not significant,  $t(228) = -.200, p=.841$ , and marginal in magnitude ( $\beta = -.016$ ). As suggested by literature reviewed, implementation of SCP can be expected to rise when teachers have more positive views of their own role in SCP. Yet, that conceptualization of this scale showed no relationship in the sample surveyed. There may also be statistical suppressor effects of similar variables entered in the model.

## **Influences of Perceived Control by Teacher Evaluation Policy on SCP Implementation**

**Research question 4.0:** To what extent is *control* enforced by the organization's *Gaokao*-related teacher evaluation policy for secondary teachers, a negative predictor and a significant mediating variable for SCP implementation in classrooms, as reported by teachers?

**Hypothesis 4.0:** Controlling for school level variability, teacher background characteristics, and SCP-relevant factors, *Teachers' Perceived Levels of Control by the Teacher Evaluation Policy* will significantly and substantially predict levels of classroom implementation of SCP.

Hypothesis 4.0 was intended to test the influences of the two organizational control factors (Perceived control by the Output-based Components of Teacher Evaluation Policy and Perceived control by the Process-based Components of Teacher Evaluation Policy) on SCP Implementation, controlling for SCP-relevant factors.

Table 31 displays results of the regression analyses of Hypothesis 4.0.

Table 31

*Teachers' Reports of SCP Implementation Regressed on Perceived Control by Teacher**Evaluation Policy*

| Variable                                | SCP Implementation |              |                   |              |
|---|--------------------|--------------|-------------------|--------------|
|   | Model 5            |              | Model 6           |              |
|   | $\beta$            | Std. $\beta$ | $\beta$           | Std. $\beta$ |
| Support                                 | .090               | .112         | .091              | .113         |
| General Beliefs                         | .231**             | .221         | .240**            | .229         |
| Self-Efficacy                           | .169**             | .237         | .165**            | .232         |
| Control by the Output-based Components  |                    |              | -.117*            | -.123        |
| Control by the Process-based Components |                    |              | -.020             | -.055        |
| Constant                                | 2.434              |              | 4.361             |              |
| <i>F</i>                                | 15.335**           |              | 10.431**          |              |
|   | ( <i>p</i> =.000)  |              | ( <i>p</i> =.000) |              |
| <i>R</i> <sup>2</sup>                   | .168               |              | .188              |              |
| $\Delta R^2$                            |                    |              | .020              |              |

Note. \*  $p < .05$ ; \*\*  $p < .01$ .

As seen in Table 31, the overall model is significant,  $F(5,226) = 10.431, p=.000$ , after the two new “control” variables were added. Compared to Model 5, Model 6’s explanatory power on variance in SCP Implementation increased from .168 to .188, meaning that 19% of the variance in teachers’ classroom implementation of SCP is now attributable to the predictors in Model 6.

At the individual factor level, the variable that is of the direct interest to the present study, Teachers' Perceived Control of the Output-based Components of Teacher Evaluation Policy (*Gaokao*-related), does have a significant influence on SCP Implementation,  $\beta = -.117$ ,  $t(226) = -1.982$ ,  $p = .049$ . An added 2% variance in teachers' implementation of SCP was uniquely explained by Perceived control by the Output-based Components ( $\Delta R^2 = .02$ ). Importantly, the direction of the linear relationship is negative. In contrast, the Perceived Control by the Process-based Components factor was not a significant predictor,  $\beta = -.020$ ,  $t(226) = -.833$ ,  $p = .406$ .

To sum up, in Model 6, there are three significant predictors of SCP Implementation: General Beliefs, Self-efficacy, and Perceived Control by the Output-based Components. With respect to the standardized regression coefficients, Control by the Output-based Components had a standardized  $\beta$ ,  $-.123$ , indicating that for every one standard deviation increase in the degree to which teachers perceive their school emphasizes *Gaokao* and *Gaokao*-related components when conducting job performance evaluations, teachers' implementation of SCP decreases by .12 standard deviation units.

### The Moderating Influences of Grade Level and Class Size

**Research question 5.0:** To what extent is the relationship between *control* enforced by the organization's *Gaokao*-related teacher evaluation policy for secondary teachers, with reported SCP implementation levels in the classroom, moderated by *Grade Level* and *Class Size*?

**Hypothesis 5.0:** The relationship between teachers' perceived levels of control by the teacher evaluation policy and levels of classroom implementation of SCP will be moderated by *Grade Level* taught by teachers.

**Hypothesis 6.0:** The relationship between teachers' perceived levels of control by the teacher evaluation policy and teachers' levels of implementation of SCP will be moderated by *Class Size*.

Hypothesis 5.0 expected that the relationship between Perceived Control by the Output-based Components (Gaokao-related) of Teacher Evaluation Policy and SCP Implementation is moderated by Grade Level since reviews of literature showed that 12<sup>th</sup> grade teachers face higher pressures generated by *Gaokao* and other related high-stakes, standardized tests.

Table 32 displays results of the regression analyses regarding Hypothesis 5.0. For the purpose of comparison, information for analytical Model 6 was included in the table as well.

Table 32

*Moderating Effects of Grade Level*

| Variable                                | SCP Implementation |              |              |              |
|---|--------------------|--------------|--------------|--------------|
|   | Model 6            |              | Model 7      |              |
|   | $\beta$            | Std. $\beta$ | $\beta$      | Std. $\beta$ |
| Support                                 | .091               | .113         | .080         | .099         |
| General Beliefs                         | .240**             | .229         | .240**       | .229         |
| Self-Efficacy                           | .165**             | .232         | .168**       | .235         |
| Control by the Output-based Components  | -.117*             | -.123        | -.139*       | -.146        |
| Control by the Process-based Components | -.020              | -.055        | -.018        | -.051        |
| Grade Level                             |                    |              | .180         | .083         |
| Grade Level * Output-based Components   |                    |              | -.036        | -.198        |
| Constant                                | 4.361              |              | 4.609        |              |
| <i>F</i>                                | 10.431**           |              | 8.041**      |              |
|   | ( $p=.000$ )       |              | ( $p=.000$ ) |              |
| $R^2$                                   | .188               |              | .201         |              |
| $\Delta R^2$                            |                    |              | 0.013        |              |

Note. \*  $p < .05$ ; \*\*  $p < .01$ .



Table 32 shows that after adding the interaction variable, the overall model is significant,  $F(7,224)= 8.041, p=.000$ . However, the model's explanatory power was not attributed to the interaction variable. The individual  $p$  value of the interaction variable was not significant at the .05 level,  $\beta= -.036, t(224)= -.670, p=.503$ . The null hypothesis is therefore accepted that there is no interaction effect. In other words, the relationship between SCP Implementation and the system's centralized structure and the associated output control mechanism is not moderated by grade of teaching.

Hypothesis 6.0 expected that the relationship between teachers' perceived control of the Output-driven evaluation policy and teachers' implementation of SCP is moderated by class size.

Table 33 displays results of the regression analyses. For comparison, information for analytical Model 6 was included in the table as well.

Table 33

*Moderating Effects of Class Size*

| Variable                                | SCP Implementation |              |              |              |
|---|--------------------|--------------|--------------|--------------|
|   | Model 6            |              | Model 8      |              |
|   | $\beta$            | Std. $\beta$ | $\beta$      | Std. $\beta$ |
| Support                                 | .091               | .113         | .096         | .120         |
| General Beliefs                         | .240**             | .229         | .248**       | .237         |
| Self-Efficacy                           | .165**             | .232         | .159**       | .223         |
| Control by the Output-based Components  | -.117*             | -.123        | -.071        | -.075        |
| Control by the Process-based Components | -.020              | -.055        | -.020        | -.058        |
| Class Size                              |                    |              | 1.216        | .406         |
| Class Size * Output-based Components    |                    |              | -.085        | -.335        |
| Constant                                | 4.361              |              | 3.682        |              |
| <i>F</i>                                | 10.431**           |              | 7.935**      |              |
|   | ( $p=.000$ )       |              | ( $p=.000$ ) |              |
| $R^2$                                   | .188               |              | .199         |              |
| $\Delta R^2$                            |                    |              | .011         |              |

Note. \*  $p < .05$ ; \*\*  $p < .01$ .

Table 33 displays a very similar pattern in comparison with Table 32. Although the overall model is significant,  $F(7,224)=7.935$ ,  $p=.000$ , its explanatory power was not due to addition of the interaction variable. The individual  $p$  value of the interaction variable was not significant at the .05 level,  $\beta = -.085$ ,  $t(226) = -.927$ ,  $p = .355$ . The null hypothesis is therefore

accepted that there is no interaction effect. In other words, the relationship between teachers' implementation of SCP and teachers' perceived control by the system's output control mechanism is not moderated by class size in the sample surveyed.

### **Summary of Significant Variable Relationships**

This investigation set out to examine the direct and mediating influences of schools, selected teacher background characteristics, and a number of classroom and teacher evaluation policy variables on SCP implementation, using survey-based measures completed by 232 teachers. Results showed that, based on teacher self-reports, teachers' perceived levels of control exercised by the output-based evaluation mechanisms of the Chinese educational system significantly and negatively influenced classroom-level implementation of SCP strategies. Consistent with factors identified via the literature review, Teacher Self-efficacy in SCP (see for example, Bandura, 1997; Gibson & Dembo, 1984; Melby, 1995; Olivier, 1985; Sabatier, 1986; Woolfolk & Hoy, 1990) and Teachers' General Beliefs in SCP (see for example, Chatterji et al., 2002; Fuhrman, Clune, & Elmore, 1988; McLaughlin, 1987, 1988, 1991, & 2006; Sabatier 1986), constituted significant predictors of SCP implementation in a Chinese education reform environment, with all other variables controlled statistically in the models.

Schools, teacher characteristics (Gender, Experience, Highest Degree Obtained), and Support from schools became non-significant in comparison with Self-efficacy and General Beliefs in SCP. It should be noted, however, that Support was a significant predictor in earlier models. This was consistent with the literature reviewed (see for example, Chatterji et al., 2002; Fuhrman, Clune, & Elmore, 1988; McLaughlin, 1987, 1988, 1991, & 2006; Sabatier 1986). However, in comparison with other predictors the Support factor became a non-significant

variable in the final analytic models. A comprehensive discussion of the validated conceptual framework, with reference to the literature review in Chapter II, is provided in Chapter V.

## CHAPTER V DISCUSSION

This chapter starts with a presentation of the validated conceptual framework based on the overall results of the analyses, followed by specific interpretations of results corresponding to individual research questions and hypotheses. Implications of the results for theory on reforms, large scale educational reforms in China, and education policy at large, are discussed. The conclusion chapter also identifies the study's limitations and makes suggestions for future research.

### **The Validated Conceptual Framework**

Research on China's new curriculum reforms have documented poor implementation levels of a classroom-level curriculum strategy that forms the centerpiece of China's national education reform policies, namely, *Student-Centered Pedagogy* (SCP). This dissertation set out to investigate the influence of schools, teacher background characteristics, and a number of classroom-level and teacher evaluation policy variables on SCP implementation levels by teachers, using survey-based measures.

The aim of the present research was to study a number of related factors affecting teachers' SCP implementation levels with the help of a proposed conceptual framework. The study particularly examined the potential adverse influence of an output-driven teacher evaluation policy tied to secondary school students' performance on the national college entrance examination, *Gaokao*, on SCP implementation levels reported by high school teachers. Eight other contextual and reform-related factors derived from a review of existing literature were tied together with the above variables in the framework. Based on the framework, paths by which the

variables could affect SCP implementation levels directly and indirectly were tested in stages with a series of hierarchical regression models.

The theoretical premise of the study was that the large size and highly centralized structure of the Chinese educational system led it to adopt an output-control mechanism in the form of high-stakes teacher evaluation policies tied to student performance on *Gaokao*. The adoption of such an output-control mechanism resulted in a mismatch between the philosophy underlying the newer SCP reforms and the pre-existing teacher evaluation policies, which in turn led to poor implementation levels of SCP in classrooms by teachers. Past research in China has largely overlooked the importance of policy incompatibility issues in examining effects of reforms in classrooms.

Figure 3 shows the conceptual framework, with validated variable relationships and pathways indicated in bold double lines. With all the specified independent and mediating variables in the conceptual model, the cumulative variance explained in the dependent variable, SCP Implementation levels, was 20%,  $R^2 = .199$ . The overall model was statistically significant,  $F(7,224)=7.935$ ,  $p=.000$ .

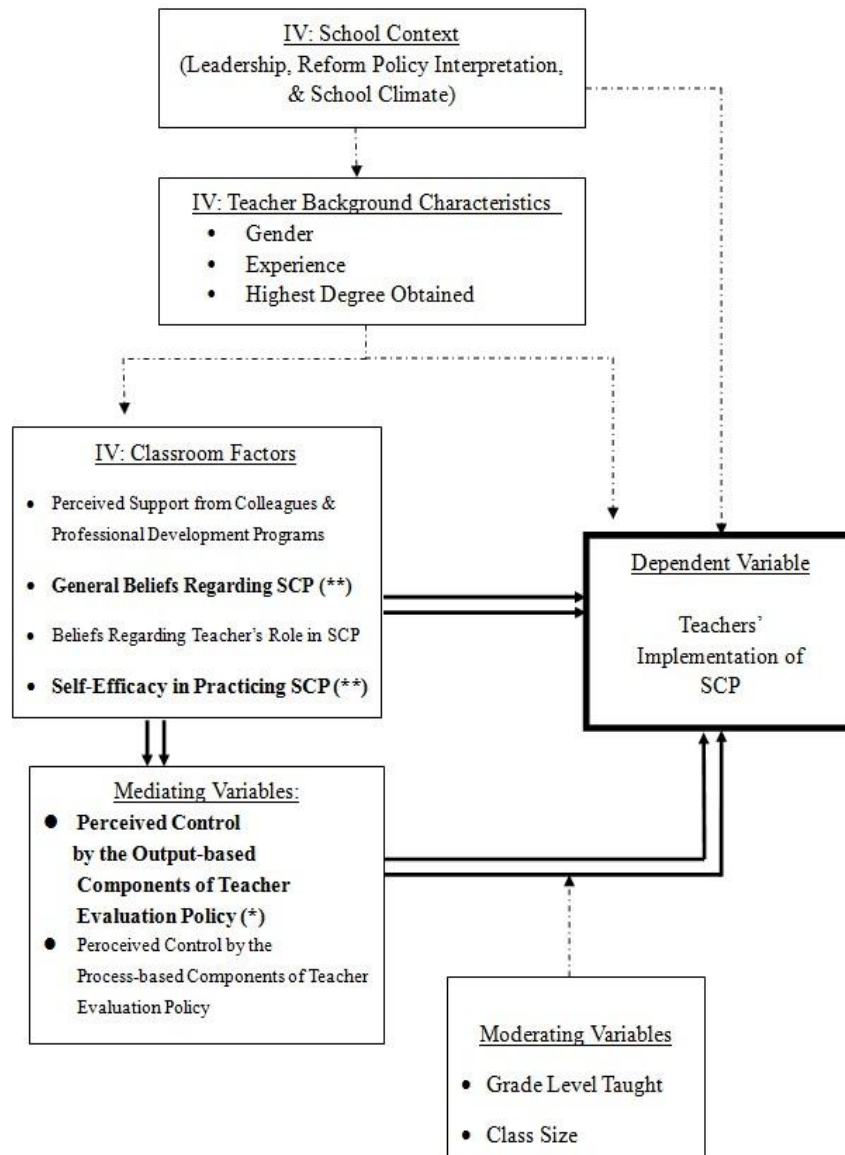
Consistent with the literature, the influences of both Beliefs in SCP and Self-efficacy in Practicing SCP on SCP implementation were significant at the .05 level (for Beliefs Regarding SCP,  $t(224)=3.745$ ,  $p=.000$ , for Self Efficacy,  $t(224)=3.387$ ,  $p=.001$ ). Other studies, mostly qualitative research from the U.S., have found similar results. For example, Fuhrman, Clune, and Elmore (1988) found that student curriculum standards mandates were implemented to a much higher degree than other types of reforms, such as teacher career-ladder-related policies in six states because teachers had better understanding of the former and felt competent to make the change. Other researchers also pointed out the importance of beliefs in implementing new

reforms, including beliefs in their ability to carry them out (see Bandura, 1997; Chatterji et al., 2002; Gibson & Dembo, 1984; McLaughlin, 1988, 2006; Melby, 1995; Olivier, 1985; Sabatier 1986; Woolfolk & Hoy, 1990).

Also consistent with the literature, the influence of teachers' perceived levels of control by high stakes teacher evaluation policy based on outputs on the *Gaokao*, was significant at the .05 level on levels of SCP implementation. The significant, yet negative, correlation confirms the main hypothesis of this study. This result was also consistent with findings from the U.S. education reform context, where exploratory studies have shown that teacher changes with regard to SCP-related reforms have been adversely affected by the output-driven, high stakes accountability measures (see Deboer, 2002; Meyer & Rowan, 2006; Passman, 2000; Pedersen & Liu, 2003; Spillane & Burch, 2006; Watanabe, 2007).

Contrary to the literature, school factors (e.g., leadership style, interpretation of reform policies, school climate and so on) and teacher background characteristics (Gender, Experience, and Highest Degree Obtained) were not found statistically significant. Perceived support, including resources, professional development programs, support from principals and colleagues, was also not found to be a statistically significant predictor in the end. Furthermore, moderators (Grade Level Taught and Class Size) were found not statistically significant either, which was contradictory to what the U.S. literature suggests. A variety of reasons, such as measurement and coding issues, or differences in research contexts could explain the non-significant results. These issues will be further explored in the discussion of corresponding individual hypothesis in the next section.

Figure 3.0 The Validated Conceptual Framework and Statistically Significant Variable Relationships and Paths



Note: IV=Independent Variables

Bold double lines indicate significant relationships between the dependent variable (SCP Implementation) and the predictors. Dotted lines indicate paths that were not significant in regression analyses. Significant predictors are in **bold font** and with statistical significance levels indicated with asterisks: \*  $p < .05$ ; \*\*  $p < .01$ .



## Discussion of Results by Research Question and Hypothesis

### Results on Research Questions 1.0-2.0

**Research question 1.0:** To what extent do *school* characteristics, as predicted by educational reform and policy implementation literature, affect SCP implementation at the classroom level as reported by teachers?

**Research question 2.0:** To what extent do selected *teacher characteristics*, as predicted by educational reform and policy implementation literature, affect SCP implementation in the classroom as reported by teachers, after taking school-level differences into account?

Hypotheses 1.0 and 2.0 could not be confirmed. School context and teachers' background characteristics were not significant predictors of teacher SCP implementation, contrary to the literature ( see for examples, Afshari, Baker, Luan, samash, & Fooi, 2009; Bulach & Malone, 1994; Datnow & Castellano, 2001; Datnow, Hubbard, & Mehan, 1998; Hargreaves, 2005; Leithwood & Jantzi, 2006; Rogers, 1995; Schiller, 2003; Smith & Desimone, 2005; Stalling & Mohlman, 1981; Venkatesh and Morris, 2000; Yuen, Law, & Wong, 2003). The answers to Research Question 1.0 and Research Question 2.0 were thus inconclusive.

The results obtained with school context factors may be explained in two ways. First, the categorically coded school variable failed to yield significance between-school variance on SCP implementation levels by teachers. The hypothesis was intended to examine whether teachers' implementation of SCP is significantly different from one school to others due to differences in collective conditions. The *F* tests found no significant difference between schools, confirming

the study's assumption that teachers' implementation of SCP varies to a much higher degree at the intra-school level, instead of the inter-school level. From this perspective, the results supported use of teachers as the unit of analysis instead of teachers grouped by school.

Alternatively, the way the school context variables were constructed and coded may have overshadowed detection of the effects of school contextual factors. The school variables were contrast-coded. This study did not separately measure school-level factors, such as leadership style and interpretation of reform policies, collapsing them instead into a 7-level categorical factor. It is possible that if measured differently, particular contextual factor(s) would have significant relationships with teacher implementation of SCP.

The decision to categorically code the school variable was made because of a lack of available instruments in Chinese tapping school level constructs such as, leadership style, climate and policies. Data could thus not be gathered to derive these measures for the present study. Future research may focus on addressing this limitation.

### Results on Research Question 3.0

**Research question 3.0:** To what extent do selected SCP-relevant constructs (*Teacher Perceived Support, Teacher Beliefs, and Teacher Self-efficacy in Practicing SCP*), as predicted by the educational reform and policy implementation literature, affect SCP implementation in the classroom as reported by teachers, accounting for school and teacher characteristics?

The evidence showed that out of the four specific SCP-relevant independent variables at the classroom level, two (General Beliefs Regarding SCP and Self-efficacy in Practicing SCP) turned out to be significant predictors of the criterion variable of *SCP Implementation*. In other words, how often teachers implemented specific SCP strategies in their classrooms was influenced by the level of their endorsement of and beliefs in the technical properties of SCP and their beliefs in their own capacity to handle the new instructional activities. This finding is consistent with theoretical predictions stemming from the previous literature review (see Bandura, 1997; Chatterji et al., 2002; Fuhrman, Clune, & Elmore, 1988; Gibson & Dembo, 1984; McLaughlin, 1987, 1988, 1991, & 2006; Melby, 1995; Olivier, 1985; Sabatier 1986; Woolfolk & Hoy, 1990).

The variable, Support, was a significant predictor in both analytical models 3 and 4. However, in Model 5, after Self-efficacy was entered, Support became statistically insignificant. This change indicated that most of the variance in the *SCP Implementation* variable initially explained by Support, was now attributable to teachers' Self-efficacy. The results from the separate regression models suggest that teachers' SCP implementation was influenced positively

by resources required for changes, professional development opportunities for teachers, and school supports for SCP. However, it was also influenced significantly by Self-efficacy, and this latter variable was more dominant. When teachers have self –efficacy or beliefs in their own capacity to engage in SCP was entered in models, school resources and support for reforms were not as relevant as predictors.

The indicators operationally defining the Support factor were: a) professional development programs, and b) support from colleagues. Based on the literature review (McLaughlin, 1988, 1991), the availability, usefulness, and sustainability of relevant professional development programs are critical for reforms. Support from colleagues and the perceived overall level of practices used by fellow teachers, also matter (Fullan, 1991; McLaughlin, 2006; Spillane, Beiser, & Gomez, 2006). The present study concluded, based on the significant change in the Support factor after Self-efficacy was entered in analytic models, that a) Support factor may not directly affect teachers' implementation of SCP as much as self-perceived capacity and attitudes towards SCP, and b) professional development programs and positive social interaction with colleagues are the main venues by which to improve teachers' competency in carrying out new reform policies. Thus, it could well be that initial levels of support provided to teachers led to higher levels of self-efficacy in SCP. Support for teachers, thus, should be continued to build optimum level of capacity and self-efficacy in SCP.

## Results on Research Question 4.0

**Research question 4.0:** To what extent is *control* enforced by the organization's *Gaokao*-related teacher evaluation policy for secondary teachers, a negative predictor and a significant mediating variable for SCP implementation in classrooms, as reported by teachers?

Hypothesis 4.0 was confirmed by the results, and the answer to the Research Question 4.0 was that that teachers' implementation of SCP was significantly and negatively mediated by Perceived Control by the Output-based Teacher Evaluation Policy tied to *Gaokao*.

These results supported the study's central premise, that the output-driven performance evaluation policy (Ouchi, 1977, 1978, 1979) would be a deterrent to reform implementation because of its incompatibility with the philosophy underlying of SCP (Nichols & Berliner, 2007; Passman, 2000; Pedersen & Liu, 2003; Watanabe, 2007).

In contrast, Perceived Control by the Process-based Components of Teacher Evaluation Policy turned out to be an insignificant predictor of SCP implementation. This finding further confirms observations made by a number of Chinese researchers (Chen & Li, 2007; Han & Yang, 2008; Jiang, 2008; Zhang, 2007; Zhao, 2007) that although teacher evaluation policies in China generally encompass four aspects (professionalism, competency, attendance regularity, and teaching outcomes), teaching outputs, particularly those related to *Gaokao*, outweigh the others in affecting SCP implementation actions.

As per organizational theory on relative influences of behavior control versus output control mechanisms (Ouchi, 1977, 1978, 1979), these results suggest that evaluations of teaching processes in Jingyang District may have been viewed as a benign ritual by teachers. That is,

teachers may have felt confident that they would be evaluated on non-*Gaokao*-related aspects, but that they would be given favorable evaluation scores. Because evaluation of direct inspection of teaching is ritualized, it fails to affect teachers' practices on new SCP policies. Therefore, the factor, Perceived Control by the Process-based Components (Non-*Gaokao*-related), turned out to be statistically insignificant compared with the Perceived Control by the Output-based Components (*Gaokao*-related).

In contrast, evaluation results on output-based components are crucial to teachers, because it differentiates them from others, and is linked to merit pay rewards/sanctions. Upholding more conventional teacher-centered instruction serves as an easy avenue to meet the output standards set by the evaluation policies. The findings, therefore, are not unexpected that teachers' perceived control of the *Gaokao*-related evaluation components has a statistically significant yet negative, linear relation with teachers' implementation of SCP.

With respect to the magnitude of the standardized regression coefficients ( $\beta$ ), the factor Perceived Control by the Output-based Components had a relatively lower absolute value (.123), compared to General Beliefs Regarding SCP (.229) and Self-efficacy (.232). However, the relatively low  $\beta$  is because it constituted a mediator in the conceptual framework of the postulated relation between teachers' implementation of SCP and the other two predictors.

## Results on Research Question 5.0

**Research question 5.0:** To what extent is the relationship between *control* enforced by the organization's *Gaokao*-related teacher evaluation policy for secondary teachers with reported SCP implementation levels in the classroom, moderated by *Grade Level* and *Class Size*?

Hypothesis 5.0 assumed that, should a negative linear relationship exist between teachers' implementation of SCP and their perceived control of the evaluation policy's *Gaokao*-related components, such a relationship would be moderated by the grade taught. This hypothesis stems from the U.S. literature, which suggests that the degree of teacher practices differ significantly depending on grade levels that are the focus of high-stakes testing (Nichols, Glass, & Berliner, 2006). More student-oriented teaching is observed in grades with no state-administered standardized tests because teachers at the grade levels at which the test is given are particularly vulnerable to the pressure of teaching to the test (Amrein & Berliner, 2002). In the Chinese context, *Gaokao* is administered only in Grade 12. It was therefore reasonable to assume that if SCP implementation was influenced by output-driven, high stakes teacher evaluation policy, teachers who are teaching the 12<sup>th</sup> grade would be more susceptible to such an influence. Such a hypothesis, however, was not supported by regression results. The interaction variable was *not* statistically significant at the 5% level.

Why was Grade Level not a significant moderator? Chinese teachers move up the grade levels following a cohort of students all the way to graduation and then cycle back to the starting grade (Y. Peng, Head of the Deyang City Bureau of Education, personal communication, April 2010). Under such an arrangement, even though the respondents were teaching the 10<sup>th</sup> or 11<sup>th</sup> grade when the survey was conducted, they clearly knew that in order to get positive results on *Gaokao*, they must work hard toward that end in Grade 12. In some instances, schools organize a special group comprised primarily of experienced teachers to strategically handle the preparation for *Gaokao*. However, these teachers would also teach either or both of the other two grades. As such, it is unlikely that such preparation would cause them to adjust their teaching practices as

grade changes since they also knew that working mindfully towards *Gaokao* at the 10<sup>th</sup> or 11<sup>th</sup> grade today would benefit them at 12<sup>th</sup> grade tomorrow.

The lack of significant interaction effects of grade level was also borne out by Tables 13-19 showing no significant differences by Grade Level on SCP Implementation and other survey measures. It indicated a situation in which the influences of the *Gaokao*-driven evaluation policies are so pervasive that school activities may revolve around it regardless of grade level.

Hypothesis 6.0 assumed that, should a negative linear relation exist between teachers' implementation of SCP and their perceived control of the evaluation policy's *Gaokao*-related components, such a relationship would also be moderated by class size. This assumption also arises from the U.S. literature, which showed that that reduced class size significantly affected teaching methodologies (Molnar, Smith, Zahorik, Palmer, Halbach, & Ehrle, 1999). More SCP-oriented teaching behaviors appeared more often by those with smaller class size (Molnar et al., 1999). However, multiple regression results here did not support this prediction.

The insignificant results can be better understood based on results in Tables 20 to 26. Descriptive analyses on the variable, Class Size, revealed that generally classes in this sample were too large for SCP implementation to fit the theoretical ideal (Knowlton, 2000; Passman, 2000). The median was 60, and 84.4% of all the classes reported a size of 50 and above (see Table 4). The minimum class size was 30, but only one teacher reported teaching a class of this size.

The study divided the classes into two groups using 50 as the cut-off point for the interaction analysis. Regardless of how Class Size is operationally defined, even the smallest class in the data set might not be small enough to expect teachers to substantively change their



instructional practices to conform with SCP. This may be a main reason as to why no significant differences surfaced in this analysis.

### **Implications of Results for Policy, Practice and Theory**

#### **Policy Levers in the Context of China's Education Reforms**

Since the inception of the new curriculum reforms, significant resources have been allocated towards SCP. Numerous SCP-oriented professional development programs have been developed (Liu, 2011). Financial resources and materials for implementation have been provided (Zhu, 2004). Despite such expenditures, teachers' implementation of SCP still appears to be no more than symbolic and without substantive assimilation of the SCP-related principles in classroom practices and behaviors (Li, 2008; Ma & Tang, 2002; Xia, 2008; Yan & Zhou, 2008; Zhong, 2005). What factors lie at the root the observed lack of substantive SCP implementation by the teachers? How can the problem be addressed?

The present empirical examination found teachers' implementation of SCP to be affected on two levels:

a) when teachers have stronger beliefs regarding the merits and technical properties of SCP and also in their own capacity (self-efficacy) to carry out the strategies, their SCP implementation behaviors increase; and

b) when there are conflicts between teacher evaluation policies enforcing the high stakes, output-control mechanisms the Chinese educational system, and new reform policies advancing SCP, SCP implementation is negatively affected in classrooms.

Both the above findings were suggested by the literature on reform implementation and organizational theory but never formally explored. The influence of the output control factor had not been raised as a potential barrier to reforms by other investigators, nor thoroughly was it treated in the existing reform implementation literature in China.

Consistent with the U.S. education reforms literature, support from the school colleagues and through professional development opportunities were also significant predictors of SCP implementation, but these diminished in influence when the three factors identified above were present in analytic models. Together, the results suggest that the availability of necessary resources, professional development opportunities and supportive working environments are effective policy levers in enhancing teachers' beliefs in SCP and their competency in practicing SCP. Thus, it is recommended that policymakers continue these policy efforts in order to improve teachers' attitudes, cognition, and beliefs in the new SCP-related reforms and also to improve their self-efficacy in implementing the new SCP-related instruction.

However, this study also showed that teachers' SCP implementation was negatively influenced by their perceptions of controls enforced via the output-driven, high stakes teacher evaluation policies. Thus, relying on the professional development and resource policy levers will not be sufficient to foster SCP-relevant reforms. Teachers' implementation of a new policy is affected by a number of interrelated factors, and contradictory policy factors must be removed from the practice environments or disengaged from monetary reward systems.

In a system characteristic of a rigid, highly centralized structure, the dominant output-control mechanisms can put severe restraints on teachers' choice of instructional methods. As long as teachers have significant concerns surrounding annual *Gaokao*-oriented student

performance, they will likely choose conventional teacher-centered practices and test preparation techniques over SCP-oriented classroom instruction, even though they may believe in SCP and their competence in SCP implementation.

Policy makers should recognize that, in the current Chinese education system, the links between the highly centralized structure, the output control mechanisms, and the conventional mode of classroom instruction are all in direct conflict with the philosophy of new SCP reforms, reducing the effectiveness of dedicated resources, professional development programs and supportive working environments. Resistance or passivity on the part of the teachers regarding SCP should not be interpreted as a lack of training or support. The tensions and policy conflicts of the old system and the new reforms should be confronted and examined.

### **Recommendations for Change**

The current situation with SCP implementation in the classrooms, or lack thereof, calls for new approaches. Policy attention must be directed towards the structural constraints of the highly centralized organization and control mechanisms, as a starting point for a new policy cycle. To solve the contradiction between SCP and the output-driven control mechanisms of the Chinese education system, matching reforms must be undertaken aiming to remove the mismatches. This may be realized by changing the output control mechanisms, making the *Gaokao* system more consistent philosophically with SCP. Another route would be to reform the system's structure towards greater decentralization, and weighting teaching process variables more heavily in the teacher evaluation policy. A move towards *behavior control mechanisms* may be a more practical option (Ouchi, 1977).

**Reforming the High-Stakes Test, *Gaokao*.** A standardized test need not be high-stakes in nature, unless it is connected to high-stakes actions and consequences for teachers (Cole & Osterlind, 2008). For instance, in China, two standardized tests are currently applied at the 12<sup>th</sup> grade, the graduation test and *Gaokao*. The graduation test is used to determine students' qualifications for a high school diploma. *Gaokao* is used for student admission to college. For high school students, both tests are high-stakes tests. For teachers, only the *Gaokao* is considered high-stakes due to its close ties with the teacher evaluation measures and policies. These policies ought to be re-evaluated along with the *Gaokao*, if SCP reforms are to succeed.

Reforming the *Gaokao* has been a debated topic among Chinese reformers since the 1990s (Gao & Deng, 2008). Chinese educators and researchers generally agree that currently *Gaokao* is severely tilted toward assessing students' knowledge and pays little attention to building holistic skills/capacities consistent with SCP philosophy (Chen, Huang, & Huang, 2009; Ling & Long, 2009). In addition, it is the sole criterion for college admissions (Chen, Huang, & Huang, 2009; Hu, 2006; Ling & Long, 2009). Accordingly, educational efforts in schools, particularly in classrooms, and conventional teacher-centered instruction are arguably the most effective way to transmit knowledge and raise student performance on the current version of the *Gaokao*.

Strategies to reform the *Gaokao* should start by looking for ways to break the philosophical stand-off between *Gaokao* and SCP. First, *Gaokao* should be revised into a tool that accurately assesses both knowledge and other skills/capacities of students that make it more consistent with SCP principles. If such a transformation were successful, more student-centered instruction to help develop students' skills at all levels may follow.

However, an emphasis on reforming *Gaokao* alone will not put an end to the “teach-to-test” phenomenon in China. Due to the traditional emphasis on education in Chinese society as a primary means to success, *Gaokao* has also come to symbolize the Chinese values of social justice and upward mobility. Given its iconic status, any change proposed to substantially alter the current structure and procedure of *Gaokao* will potentially encounter system-wide resistance as well as criticism from the general public.

In addition, no matter how consistent a reformed *Gaokao* may be made with SCP-oriented classroom instruction, the system would likely eventually regress to a “teach-to-test” work culture because of the merit pay scheme tied to test performance for teachers and schools. In a system where educational output control mechanisms prevail, the standardized test becomes a substitute for the full curriculum and the broader educational goals suffer (Nichols & Berliner, 2007; Watanabe, 2007).

**Towards decentralization.** As suggested by organizational theory on structure and control, there is a relationship between structure, control mechanisms and teacher evaluation policies in large bureaucracies like the Chinese education system (Evans, 1975; Ouchi, 1977, 1978, & 1979; Williamson, 1971). *Gaokao* represents a means of organizational control. Empirical evidence from this study confirmed the negative tensions between SCP and *Gaokao*.

One possible remedial strategy is decentralization. SCP-related reforms are designed to change the core teaching and learning process, but monetary incentives in the current Chinese educational system discourage teachers to attempt any change that might potentially undermine their teaching outputs. A reform of the organizational structure toward decentralization could dismantle the mismatch between policy intentions and implementation incentives.

If the structure is decentralized, for each autonomous or semiautonomous unit within the organization, direct oversight of the teaching processes would become a viable alternative method for personnel evaluation. Once the SCP-oriented behavior-control mechanisms become dominant, teacher incentives would shift, thus affecting their pedagogical preferences (Evans, 1975; Ouchi, 1977, 1978, & 1979; Williamson, 1971).

Reform efforts toward decentralization do not mean that the influences of the output-control mechanisms would be entirely eliminated. For either socio-cultural or political reasons, some units in the decentralized system might still consider output-control as a supplementary means of teacher evaluation.

### **Contributions to Theory and Research Base on China's New Curriculum Reforms**

The main contributions of the present research lie in filling identifiable gaps in the existing literature on factors affecting China's reforms, and in developing a validated conceptual framework and survey instrument to guide future studies on SCP implementation levels in large education systems in both China and the U.S.

As discussed in Chapter II, with the exception of a few, studies focusing on educational reforms and policy implementation were found to be mostly qualitative and observational in the U.S. context. They were largely lacking in China. Qualitative research presents two main challenges: first, findings are usually not generalizable beyond the cases studied, and, second, it is difficult to evaluate the relative importance of an array of different factors on reform-relevant variables like SCP implementation.

Survey methodology adopted in the present study was useful in addressing these critical issues. Because findings are based on a representative random sample, results reported here are

generalizable to the larger teacher population in the school district from where teachers were surveyed. This study also developed a set of validated survey measures to evaluate the relative importance of a series of reform-relevant measures on SCP implementation levels by teachers. Using hierarchical analytic models, it compared the incremental amount of variance explained by each independent variable on teacher-reported levels of SCP implementation. The analyses yielded a theoretically-validated conceptual framework and path model.

The findings about direct or indirect influences of the eight variables on the dependent variable (SCP Implementation) also fill gaps in the literature on educational reforms and policy implementation in China, and certainly in the district studied. Factors widely cited by prior research, such as professional development (McLaughlin, 1988, 1991) and social supports and networks (Fullan, 1991; McLaughlin, 2006; Spillane, Beiser, & Gomez, 2006) were also significant at first, but less important with the presence of the other two. Validation of the present study's premise that teachers' implementation of SCP is also affected by the system's output-control mechanisms (e.g. the formal personnel evaluation policies) further confirmed suspected issues of incompatibility between old and new policies. For researchers interested in exploring SCP-related implementation issues, these findings provide a starting point for new research.

Lastly, given the acceptable results of the investigations on final scales, the research effort simultaneously generated an original, scientifically validated teacher survey instrument to study educational reform implementation in future. There are two versions of the survey in both English and Chinese languages, permitting future studies in educational contexts in China and the U.S.

The teacher questionnaire developed by the present study may also serve as a foundation for new or broadened instrumentation research. Domains of this questionnaire were based on extensive literature review. Items were constructed around operational definitions of SCP shared by researchers in both China and the U.S. But future iterations could yield improved measures.

### **Implications for Reform Policy and Theory in U.S. Education**

Findings of this study may have some lessons for U.S. public education contexts where reformers are pushing for high stakes testing and teacher evaluation policies coupled with SCP reforms in classrooms. The policies are incompatible. SCP will very likely be undermined as long as NCLB-like policies are enforced with punitive sanctions for schools and teachers based on students' test scores. Tools and conceptual models offered through this research could be employed to investigate similar issues in U.S. contexts.

Second, the U.S. literature has not examined the potential negative influence of output-control mechanisms from an organizational theory perspective. This gap may also be addressed by future research in U.S. contexts.

### **Limitations of the Study**

#### **Generalizability**

The study was conducted within one local district. That scope limits the inferences that can be made from the results of the analyses. The principal findings of this study are applicable to high school teachers in Jingyang District, Deyang City. The sample was representative of the population on two variables, School and Gender. However, the generalizability of results to other districts and the nation remain limited.



Jingyang District was chosen as the research site due to its typical district profile on economic advancement, population density, and geographic features. However, one should be cautious in applying findings to other areas simply based on superficially similar characteristics. The degree of emphasis placed on high-stakes, standardized tests such as *Gaokao* might vary greatly from district to district. The monetary rewards in Jingyang District's teacher evaluation policy are quite substantial and can incentivize conventional teaching practices in this district. However, such incentives might not exist in regions with more advanced economic development. It is also possible that such incentives lie beyond the fiscal capacity of districts in poorer regions. In sum, without a nationwide comprehensive survey, this study's findings should be applied with caution to other districts or larger administrative units, such as prefects or provinces in China.

The generalizability of the study's findings is further limited when applied to educational systems with a structure different from the Chinese educational system. Although the study's main hypothesis partly stemmed from U.S. educational reforms and organizational theory literature bases, the U.S. educational system has a very dissimilar structure from the highly centralized Chinese system. Some researchers have described the U.S. educational system's structure as fragmented centralization (Meyer, Scott, & Strang, 1987). As a result, this study can only serve as a reference in conjunction with other similar research to explain or predict American teachers' SCP-related practices.

### **Measurement of Variables**

Given the researcher's limited time and resources, the study used a written, structured questionnaire as the primary data collecting tool. The large number of close-ended questions provided two advantages. First, it made the data collection process efficient in comparison with

other types of survey methods, such as interviews or open-ended questions. Second, asking respondents to apply numeric values to questionnaire items reduces subjectivity in treating the data for analysis.

However, measurement problems with self-report measures are also well-known. Responses can be easily faked or influenced by factors that are irrelevant to the construct. For example, the social desirability factor cannot be ignored. Under the influences of this factor, respondents may be inclined to, consciously or unconsciously, choose options that they believe others would want to hear without truly considering their perspective on the issue. To cope with possible errors associated with these adverse factors, following Chatterji (2003), the study adopted several strategies when designing the instrument, including assurance of privacy and confidentiality, assurance of anonymity, random mixing of negatively and positively stated items, pilot-testing to enhance the clarity in items, directions and standardization of data collection procedures.

However, the effectiveness of these strategies' may be limited, especially when respondents were asked to tell how often they implement SCP strategies in their classrooms. Since the differences between the concepts of student-centered-pedagogy and teacher-centered-pedagogy are easily distinguishable, it was easy for teachers to figure out which items were related to the former and which to the latter. Under the influences of social desirability, respondents could report a higher level of SCP implementation than was actually true. Such responses could lead to potential inaccurate conclusions on the relation between teachers' implementation of SCP and the perceived organizational control, making the estimates more conservative.

Another concern associated with the written, structured survey method relates to its capacity to measure complicated behavioral constructs such as SCP implementation. Are several items truly able to reflect the interactive, open-ended nature of the SCP-oriented classroom climate? Although the iterative process on designing and validating the questionnaire helped to address these concerns, it might not fully eliminate such shortcomings. For example, the alpha reliability estimates for the outcome variable, *teachers' implementation of SCP*, was .68, slightly below the .70 criterion.

In addition, SCP Implementation was cast in a general form rather than tailored to domains of instructional functioning. Teachers' practices of SCP are not necessarily uniform across different subjects. Teachers in foreign language or Chinese literature instruction, for instance, might employ the memorization strategy more often than teachers in mathematical or science instruction. Due to these potential pitfalls in the instrumentation, further research and development of the survey-based scales and other measures is recommended.

Thus, validation studies should continue. In particular, the English version of the instrument should be content-validated, field-tested and evaluated in U.S. contexts before use.

### **Analytic Models and Multiple Regression Procedure**

Multiple-regression is useful in examining statistical significance and magnitude of the relationships between the criterion variable and the variable of interest while statistically removing the influences of other factors that are previously entered in the models. However, since the explanatory power of multiple-regression is built on analysis of variance, it might fail to explain generic actions. In other words, should a relationship exist between teachers' implementation of SCP and their perceived organizational control by the evaluation policy,

multiple-regression can only detect such a relationship if both variables contain a sufficient level of variance. This analytical method might prove insufficient when answering questions pertaining to a perceived organizational control's relationship to universal non-compliance of the new SCP policy and to the degree of such a relationship. Future research should examine mediator effects with Sokol's test (Baron & Kenny, 1986).

Further, when multiple and similar variables are entered, some variables can suppress the influences of others. Statistical suppression may have been the cause for the findings on the Support variable in the present research. This anomaly should be further investigated.

Limitations of this kind may also be to blame when the study failed to detect possible influences of class size on teachers' implementation of SCP due to a lack of between-group variance. It is possible that class size has contributed to the uniform non-compliance of SCP implementation among teachers, to certain extent. It is unfortunate, therefore, that the present study could not satisfactorily capture this relationship.

### **Suggestions for Future Study**

The line of inquiry initiated by this study can help in furthering China's education reform goals. Future investigations are therefore suggested that broaden the scale of data collection to be national in scope. The influences of factors where non-significant results were found, like school, grade, class size might be more easily detected using a larger sample with higher levels of variance in factors.

The present study used individual teachers the unit of analysis. Future research should explore the viability of multi-level models that examine influences of the control policy on teachers' SCP implementation at the organizational or school level, with teachers nested within

schools. Top-tiered schools may be compared to lower-tiered schools, or urban schools with rural schools on SCP implementation. It would be interesting to see how centralized controls affect SCP implementation at both the inter-school and the intra-school levels. Further, addition of a school leadership and climate survey might shed greater light on correlates of reform implementation in classrooms.

It is further recommended that future research adopt the mixed-methods approaches instead of relying solely on quantitative survey research methods. As discussed earlier, written, structured questionnaire items contain potential measuring problems, especially when addressing complicated behavioral constructs, such as SCP implementation. In the future, qualitative data such as classroom observations can be employed to help address potential measurement issues and obtain convergent validity of findings.

## REFERENCES

- Afshari, Bakar, Luan, Samash, & Fooi (2009). Factors affecting teachers' use of information and communication technology. *International Journal of Instruction*, 2(1), 77-104.
- Amrein, A. L., & Berliner, D. C. (2002). High-stakes testing, uncertainty and student learning. *Education Policy Analysis Archives*, 10(18), 1-74.
- Anagnostopoulos, D. (2003). The new accountability, student failure, and teachers' work in urban high schools. *Educational Policy*, 17(3), 291-316.
- Anonymous. (2009). Salary of teachers in Sichuan province [Web post by the Network for Agengices in Sichun]. Retrieved from <http://health.scjg.com.cn/article.aspx?id=60647>
- Babbie, E. (1990). *Survey Research Methods* (2<sup>nd</sup> ed.). Belmont, CA: Wadsworth Publishing.
- Baker, B., Green, P., & Richards, C. (2008). *Financing Educational Systems*. Upper Saddle River, N.J.: Pearson/Merrill Prentice Hall.
- Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York: W. H. Freeman and Company.
- Barkdale-Ladd, M.A., & Thomas, K.F. (2000). What's at stake in high-stakes testing: Teachers and parents speak out. *Journal of Teacher Evaluation*, 51(5), 384-397.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.

- Beniger, J. R. (1986). *The Control Revolution: Technological and Economic Origins of the Information Society*. Cambridge, MA: Harvard University Press.
- Berman, P., & McLaughlin, M. (1979). *An Exploratory Study of School District Adaptation*. Santa Monica, CA: Rand Corporation
- Berman, P., & McLaughlin, M. (1977). Factors affecting implementation and continuation. In authors, *Federal Programs Supporting Educational Change*. Santa Monica, CA: The Rand Corporation.
- Bransford, J., Brown, A. L., & Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington: National Academies Press.
- Bray, M. (2003). Control of education: Issues and tensions in centralization and decentralization. In Robert F. Arnove & Carlos A. Torres (eds.), *Comparative Education: The Dialectic of the Global and the Local* (2<sup>nd</sup> ed.). Lanham: Rowman & Littlefield.
- Brislin, R.W. (1986). The wording and translation of research instrument. In Walter J. Lonner and John W. Berry (Eds.), *Field Methods in Cross-Cultural Research and Methodology Series*, Vol. 8 (pp. 137-164). Thousand Oaks, CA: Sage Publications, Inc.
- Bulach, C., & Malone, B. (1994). The relationship of school climate to the implementation of school reform. *ERS Spectrum*, 12(4), 3-8.
- Chatterji, M. (2003). *Designing and Using Tools for Educational Assessment*. Boston: Pearson Education, Inc.

- Chatterji, M., Sentovich, C., & Ferron, J. et al. (2002). Using an iterative model to conceptualize, pilot test, and validate scores from an instrument measuring teacher readiness for educational reforms. *Educational and Psychological Measurement*, 62(3), 444-465.
- Chen, D.J., Huang, C., & Huang, X.X. (2009). The comparison and contrasting of the Gaokao systems between China and Japan. *Journal of Sichuan Economic Management Institute*, 67(1), 63-65.
- Chen, Y. L., & Li, H. S. (2007). An exploration of teacher management in basic education. *College Journal of Jingmen Tech (China)*, 4, 43-46.
- Cheng, Y. (2001). In retrospect of the 50-year history of centralization and decentralization in Chinese education. *Journal of College of Adult Education, Hubei University*, 19(1), 38-40.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (2<sup>nd</sup> ed.). New York: Academic Press.
- Cole, J. S., & Osterlind, S. J. (2008). Investigating differences between low- and high-stakes test performance on a general education exam. *The Journal of General Education*, 57(2), 119-130.
- Crocker, L., & Algina, J. (2006). *Introduction to Classical and Modern Test Theory*. Pacific Grove, CA: Wadsworth.
- Datnow, A., & Castellano, M. E. (2001). Managing and guiding school reform: Leadership in Success for All Schools. *Educational Administration Quarterly*, 37(2), 219-249.



- Datnow, A., Hubbard, L., & Mehan, H. (1998). *Educational Reform Implementation: A Co-Constructed Process*. A research report sponsor by the Center for Research on Education, Diversity and Excellence, UC Berkeley.
- Debard, R., & Kubow, P. K. (2002). From compliance to commitment: The need for constituent discourse in implementing testing policy. *Educational Policy*, 16(3), 387-405.
- Deboer, G.E. (2002). Student-centered teaching in a standards-based world: Finding a sensible balance. *Science and Education*, 11, 405-417.
- Deci, E. & Ryan, R. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology*, 53(6), 1024-1037.
- Dewey, J. (1938). *Experience and Education*. New York: Collier Macmillan.
- Deyang City Bureau of Education, (2005). The Report of Progress of the New Curriculum Reform [Web documents]. Retrieved from <http://www.scjks.net/Article/kgxl/jycl/2000609/1300.html>
- Donnell, A. (1999). The philosophy of science and its implications for astrology. *The Wholistic Astrologer*, 1. Retrieved from <http://www.aplaceinspace.net/Pages/AndrePhilosophyofScience.html>
- Duffy, T.M. & Jonassen, D. (Eds.), (1992). *Constructivism and the technology of instruction: A conversation*. Hillsdale NJ: Lawrence Erlbaum Associates.
- Edmonds, R. R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37, 15-24.

- Eisenhardt, K. M. (1985). Control: organizational and economic approaches. *Management Science*, 31 (2), 134-155.
- Evans, P.B. (1975). Multiple hierarchies and organizational control. *Administrative Science Quarterly*, 20, 250-259.
- Ferguson, R.F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation*, 28, 465-498.
- Ferguson, R.F., & Ladd, H.F. (1996). How and why money matters: An analysis of Alabama schools. In H. Ladd (Ed.), *Holding Schools Accountable*. Washington, DC: Brookings Institution.
- Field, A. (2005). *Discovering Statistics Using SPSS* (2<sup>nd</sup> Ed.). Thousand Oaks, CA: SAGE Publications.
- Friere, P. (1970). *Pedagogy of the Oppressed*. New York: Herder and Herder.
- Fullan, M. (1991). *The Meaning of Educational Change*. New York: Teachers College Press.
- Fullan, M. (1985). Change processes and strategies at the local level. *The Elementary School Journal*, 85(3), 391-422.
- Fuhrman, S., Clune, W., & Elmore, R. (1988). Research on education reform: Lessons on the implementation of policy. In David H. Monk and Julie Underwood (Eds.), *Microlevel School Finance: Issues and Implications for Policy*. Copyright at 1988 by the American Education Finance Association.

- Gao, Y.H., & Deng, J.B. (2008). A discussion on the progress of the new curriculum reform under the environment of the college entrance exam. *Anhui Literature*, 9, 319-320.
- Gibson, S., & Dembo, M. H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*, 76, 569-582.
- Grant, M.M., & Hill, J. R. (2006). Weighing the risks with the rewards: Implementing student-centered pedagogy within high stakes testing. R. Lambert and C. McCarthy (Eds.), *Understanding Teacher Stress in an Age of Accountability*. The Information Age Publishing.
- Guidelines for New Curriculum Reforms in Basic Education in China (the experimental version). (2001). [Web post by China Education and Research Network]. Retrieved from [http://edu.cn/20010926/3002911\\_2.shtml](http://edu.cn/20010926/3002911_2.shtml)
- Han, B., & Yang, C. (2008). The remedy to teacher evaluation in the new curriculum reform. *Science and Technology Information*, 33, 465-466.
- Hargreaves, A. (2005). Educational change takes ages: Life, career, and generational factors in teachers' emotional responses to educational change. *Teaching and Teacher Education*, 21 (8), 967-983.
- Hawkins, J. N. (2000). Centralization, decentralization, recentralization: Educational reform in China. *Journal of Educational Administration*, 38(5), 442-455.
- Hedges, L.V., Laine, R.D., & Greewald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher*, 23(3), 5-14.

- Henry, G. T. (1990). *Practical Sampling*. Newbury Park, CA: SAGE
- Hess, F.M. (1999). *Spinning Wheels: The Politics of Urban School Reform*. Washington, D.C.:  
Brookings Institution Press.
- Holt, D. G.; Willard-Holt, C. (2000). "Let's get real: Students solving authentic corporate problems". *Phi Delta Kappan* 82 (3).
- Honig, M.I. (2006). Complexity and Policy Implementation: Challenges and Opportunities for the Field. In author (Ed.), *New Directions in Education Policy Implementation: Confronting Complexity*. Albany, NY: State University of New York Press.
- Honig, M.I. (2004a). Where's the "up" in bottom-up reform? *Education Policy*, 18(4), 527-561.
- Honig, M.I. (2004b). The new middle management: Intermediary organizations in education policy implementation. *Educational Evaluation and Policy Analysis*, 26(1), 65-87.
- Hu, W. P., Han, Q., Wen, P. N., & Li, J. B. (2005). A survey of new curriculum reform implementation status in elementary schools. *Curriculum, Teaching Material and Method (China)*, 25(2), 8-15.
- Hu, Z. P. (2006, November). Reforming the Gaokao system: A challenge to social honesty. In the *Illumination Daily (China)*. Retrieved from [http://news.xinhuanet.com/theory/2006-11/08/content\\_5303324.htm](http://news.xinhuanet.com/theory/2006-11/08/content_5303324.htm)
- Jiang, J. (2008). Teacher evaluation in rural basic education. *Modern Education Science (China)*, 3, 4-7.

- Jorde-Bloom, P., & Ford, M. (1988). Factors influencing early childhood administrators' decisions regarding the adoption of computer technology. *Journal of Educational Computing Research, 4*, 31-47.
- Judd, C.M., & McClelland, G.H. (1989). *Data Analysis: A Model-Comparison Approach*. New York: Harcourt Brace Jovanovich, Inc.
- Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika, 39*, 31-36.
- Kirst, M.W., & Jung, R. (1980). The utility of a longitudinal approach in assessing implementation. *Educational Evaluation and Policy Analysis, 2*(5), 17-34.
- Knowlton, D.S. (2000). A theoretical framework for the online classroom: A defense and delineation of a student-centered pedagogy. *New Directions for Teaching and Learning, 84*, 5-14.
- Leithwood, K., & Jantzi, D. (2006). Transformational school leadership for large-scale reform: Effects on students, teachers, and their classroom practices. *School Effectiveness and School Improvement, 17*(2), special issue on transformational school leadership.
- Lepper, M. & Green, D. (1978). *The hidden costs of reward: New perspectives on the psychology of human motivation*. Hillsdale, NJ: Laurence Erlbaum.
- Lewis-Beck, M.S. (1980). *Applied Regression: An Introduction*. Newbury Park, CA: SAGE Publications, Inc.
- Li, Ch.L., & Xiao, Y.J. (2001). A Discussion on Centralization and Centralization in Chinese Educational System [Web post by Research Center for Government by Law]. Retrieved from [http://law.china.cn/features/200906/18/content\\_2971180.htm](http://law.china.cn/features/200906/18/content_2971180.htm)

- Li, G. X., & Wang, X. H. (2008). An investigation of the implementation of the new high school Chemistry standards in Shandong Province. *Chemistry Education (China)*, 7, 48-51.
- Li, J. P. (2005, October 18). A report on the comprehensive courses for elementary and secondary education in China [Web news post]. Retrieved from <http://www.edu.cn/20051018/3156324.shtml>
- Li, K.J. (2008). Reasons of teacher resistance of the new curriculum reform and the according reaction strategies. *Contemporary Science of Education*, 6, 15-18.
- Li, X.Y., & Long, Y. (2009). The ill-logic of the current Gaokao system. *Education Research Monthly (China)*, 3, 98-102.
- Li, Y. (2009). The relationship between the central government and local governments: A constitutional perspective [Web document posted by 110 Network for Legal Consultation]. Retrieved from <http://www.110.com/ziliao/article-135060.html>
- Lipsky, M. (1971). Street-level bureaucracy and the analysis of urban reform. *Urban Affairs Review*, 6, 391-409.
- Liu, K. (2011, October 25). Why the majority of teachers are not satisfied with implementation of the new curriculum reforms [Web log post]. Retrieved from <http://blog.people.com.cn/open/articleFine.do?articleId=1319524216268>
- Lv, W., & Pang, M. C. (2002). Report of the project “An Investigation of the Public Financial Input for Basic Education,” sponsored by the Ministry of Education, China [Web report]. Retrieved from

<http://academy.dufe.edu.cn/resources/neibuwengao/images/nongcunjichujiaoyugonggongontouru.pdf>

- Ma, Y.P., & Tang, L.F. (2002). The implementation status of the new curriculum reform and the corresponding remedies: An analysis of and reflection on the evaluation results of some of the special experimental districts. *Journal of Northeast Normal University (Philosophy and Social Science)*, 5, 124-131.
- Mazmanian, D., & Sabatier, P.A. (1980). The implementation of public policy: A framework of analysis. *Policy Studies Journal*, 8, 538-560.
- McDonnell, L.M. (1991). Ideas and values in implementation analysis: The case of teacher policy. In Allan R. Odden (Ed.), *Education Policy Implementation*. Albany, NY: State University of New York Press.
- McLaughlin, M.W. (2006). Implementation research in education: Lessons learned, lingering questions and new opportunities. In Meredith I. Honig (Ed.), *New Directions in Education Policy Implementation: confronting Complexity*. Albany, NY: State University of New York Press.
- McLaughlin, M.W. (1991). The Rand Change Agent Study: Ten years later, in Allan R. Odden (Ed.), *Education Policy Implementation*. Albany, NY: State University of New York Press.
- McLaughlin, M.W., & Yee, S. (1988). School as a place to have a career, in Ann Libberman (Ed.), *Developing the Case for a Professional Culture*. New York: Teachers College Record.

Mclaughlin, M.W. (1987). Learning from experience: Lessons from policy implementation.

*Educational Evaluation and Policy Analysis*, 9(2), 171-178.

Meckler, L. (2011, March 11). Education push includes merit pay? In *The Wall Street Journal*, p.

A6. Retrieved from <http://online.wsj.com/article/SB123668036405881929.html>

Melby, L. C. (1995). Teacher efficacy and classroom management: A study of teacher cognition, emotion, and strategy usage associated with externalizing student behavior. Ph. D. diss., University of California, Los Angeles.

Meyer, H., & Rowan, B. (2006). Institutional analysis and the study of education. In authors (Eds.), *The New Institutionalism in Education*. Albany: State University of New York Press.

Meyer, J.W., & Rowan, B. (1983). The structure of educational organizations. In John W. Meyer and W. Richard Scott (Eds.), *Organizational Environments: Ritual and Rationality*. Beverly Hills, CA: Sage Publications.

Meyer, J.W., & Rowan, B. (1978). The structure of educational organizations. In W. Meyer (Ed.), *Environments and organizations*. San Francisco: Jossey-Bass.

Meyer, J., Scott, W. R., & Strang, D. (1987). Centralization, fragmentation, and school district complexity. *Administrative Science Quarterly*, 32(2), 186-201.

Miles, K. H., & Darling-Hammond, L. (1998). Rethinking the allocation of teaching resources: Some lessons from high-performing schools. *Educational Evaluation and Policy Analysis*, 20(1), 9-29.



- Miller, J (1987). Priming is not necessary for selective-attention failures: Semantic effects of unattended, unprimed letters. *Perception and Psychophysics*, 41(5), 419-434.
- Nichols, S. N., & Berliner, D. C. (2007). *Collateral Damage: The Effects of High-Stakes Testing on America's Schools*. Cambridge, MA: Harvard Education Press.
- Nichols, S. L., Glass, G. V., & Berliner, D. C. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning? *Education Policy Analysis Archives*, 14(1), 1-56.
- No Child Left Behind (NCLB) Act of 2001, 20 U.S.C.A. § 6301 et seq. (West 2003)
- Noddings, N. (1999). Renewing democracy in schools. *Phi Delta Kappan*, April, 579-583.
- Odden, A.R. (1991). New patterns of education policy implementation and challenges for the 1990s, in author (Ed.), *Education Policy Implementation*. Albany, NY: State University of New York Press.
- Oliver, T. E. (1985). The relationship of selected teacher variables with self-efficacy for utilizing the computer for programming and instruction. Ph. D. diss., University of Houston. *Dissertation Abstracts International*, 46, 1501-A.
- Ouchi, W.G. (1977). The relationship between organizational structure and organizational control. *Administrative Science Quarterly*, 22, 95-113.
- Ouchi, W. G. (1978). The transmission of control through organizational hierarchy. *Academy of Management Journal*, 21, 173-192.

- Ouchi, W. G. (1979). A conceptual framework for the design of organizational control mechanisms. *Management Science*, 25 (9), 833-848.
- Ouchi, W.G., & Maguire, M.A. (1975). Organizational control: Two functions. *Administrative Science Quarterly*, 20, 559-569.
- Passman, R. (2000). *Pressure Cooker: Experiences with Student-centered Teaching and Learning in High-Stakes Assessment Environments*. Paper presented at the annual meeting of the Southwest Educational Research Association (Dallas, TX, January 27-29, 2000).
- Pedersen, S., & Liu, M. (2003). Teachers' beliefs about issues in the implementation of a student-centered learning environment. *Education Technology Research and Development*, 51(2), 57-76.
- Piaget, J. (1963). *Origins of Intelligence in Children*. New York: Norton.
- Pressman, J., & Wildavsky, A. (1973). *Implementation*. Berkeley, CA: University of California Press.
- Raudenbush, S. W., & Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20 (4), 307-335.
- Ravitch, D. (2010). *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. New York: Basic Books, A Member of the Perseus Books Group.
- Robson, C. (2002). *Real World Research*. Malden, MA: Blackwell Publishing.

- Rogers, E. M. (1995). *Diffusion of Innovations* (4<sup>th</sup> Ed.). New York: the Free Press.
- Sabatier, P.A. (1986). Top-down and bottom-up approaches to implementation research: A critical analysis and suggested synthesis. *Journal of Public Policy*, 6(1), 21-48.
- Schiller, J. (2003). Working with ICT perceptions of Australian principals. *Journal of Educational Administration*, 19 (8), 171-185.
- Siskin, L. S. (2003). Outside the core: Accountability in tested and untested subjects. In M. Carnoy, R. Elmore, & L. S. Siskin (Eds.), *The New Accountability: High schools and High-stakes Testing* (pp. 87-98). New York: Routledge Falmer.
- Smith, T. M., & Desimone, L. M. (2005). “Highly qualified” to do what? The relationship between NCLB teacher quality mandates and the use of reform-oriented instruction in middle school mathematics. *Educational Evaluation and Policy Analysis*, 27 (1), 75-109.
- Spillane, J.P., Beiser, B.J., & Gomez, L.M. (2006). Policy implementation and cognition: The role of human, social, and distributed cognition in framing policy implementation. In Meredith I. Honig (Ed.), *New Directions in Education Policy Implementation: Confronting Complexity*. Albany, NY: State University of New York Press.
- Spillane J., & Burch, P. (2006). *The institutional environment and instructional practice: Changing patterns of guidance and control in public education*. In H. Meyer and B. Rowan (Eds.), *The New Institutionalism in Education*. Albany: State University of New York Press.

- Stallings, J. A., & Mohlman, G. G. (1981). School policy, leadership style, teacher change and student behavior in eight schools. Washington, DC: National Institute of Education.  
Retrieved from ERIC database. (ED 209759)
- Starbuck, W. H. (1976). Organizations and their environments. In *Handbook of Industrial Organizational Psychology*, M. D. Dunnette (Ed.). Chicago: Rand McNally.
- The State Council of People's Republic of China. (2001). The Guideline for New Curriculum Reforms, 2001 (21).
- Venkatesh, V., & Morris, M. G. (2000). A longitudinal field investigation of gender differences in individual technology adoption decision-making processes. *Organizational Behavior and Human Decision Processes*, 83(1), 33-60.
- Vygotskii, L.S. (1978). *Mind in society: The development of higher mental processes*. Cambridge, MA: Harvard University Press
- Wang, Z. G., Zhao, Z. J., Duan, C. S., & Wang, T. (2007). A report on the inquiry learning in high school chemistry. *Journal of Teacher College of Qingdao University (China)*, 24(4), 113-119.
- Watanabe, M. (2007). Displaced teacher and state priorities in a high-stakes accountability context. *Educational Policy*, 21(2), 311-367.
- Weatherley, R., & Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special education reform. *Harvard Educational Review*, 47(2), 171-197.
- Weber, M. (1947). *The Theory of Social and Economic Organization*. New York: Oxford University Press

- Weber, M. (1946). *Essays in Sociology*. New York: Oxford University Press.
- Wendorf, C. A. (2004). Primer on multiple regression coding: Common forms and the additional case of repeated contrasts. *Understanding Statistics*, 3(1), 47-57.
- Williamson, O.E. (1971). *Corporate Control and Business Behavior*. Englewood Cliffs, NJ: Prentice Hall.
- Woolfolk, A. E., & Hoy, W. K. (1990). Prospective teachers' sense of efficacy and belief about control. *Journal of Educational Psychology*, 82, 81-91.
- Wu, J.X. (2005). On the goal of education and the new curricular reformation in mainland China. *Hong Kong Teachers' Centre Journal*, 4, 11-18.
- Xia, X.M. (2008). Conservative or progressive: An analysis on the new-old-mixed phenomenon occurred in the implementation process of the new curriculum reform. *Contemporary Science of Education*, 18, 3-8.
- Yan, Y.J., & Zhou, G.P. (2008). Silence of teachers: Causes and remedies. *Theory and Practice of Education*, 7, 22-23.
- Yuen, A., Law, N., & Wong, K. C. (2003). ICT implementation and school leadership: Case studies of ICT integration in teaching and learning. *Journal of Educational Administration*, 41(2), 158-170.
- Zhang, L. B. (2007). Teacher evaluation: Student performance determines teacher performance? *China Education (China)*. Retrieved from <http://fzxpj.cersp.com/JSPJ/200707/3033.html>

- Zhang, R.W. (2008). *New Curriculum Reforms: What on Earth They Bring to Us?* Fuzhou, Fujian Province (China): Fujian Educational Press.
- Zhao, L. M. (2007). An analysis of the issues in teacher evaluation. *Contemporary Education: Science of the General Education (China)*, 1, 60-62.
- Zheng, Y.N. (2010). *The Chinese Model: Experience and Dilemma*. Hangzhou, Zhejiang, China: Zhejiang People's Press.
- Zhong, Q.Q. (2008). Social constructivism: Learning in conversations and collaborations. *Shanghai Education (China)*, 45-48.
- Zhong, Q.Q. (2007). A debate on knowledge. *Shanghai Education (China)*, 4, 4-8.
- Zhong, Q.Q. (2005a). An investigation in theories regarding educational activities. *Education Research (China)*, 304(5), 36-44.
- Zhong, Q.Q. (2005b). The Chinese new curriculum reforms: Challenges and Meditations. *Comparative Education Review*, 12, 18-23.
- Zhong, Q.Q., & Yang, Q.M. (2002). Background and concepts of the new curriculum reforms in basic education. *Henan Education (China)*, 282, 14-16.
- Zhu, M. J. (2008) (Ed.). *Walking into the New Curriculum: A Dialogue to the Curriculum Implementers*. Beijing: Beijing Normal University Press.
- Zhu, M. J. (2004, September 3). A read on the new curriculum reforms [Web new]. Retrieved from [http://jyx.zjc.edu.cn/shownews.asp?news\\_id=27](http://jyx.zjc.edu.cn/shownews.asp?news_id=27)

**APPENDICES***Appendix A*

Excerpts of the Teacher Evaluation Policy by Jingyang District, Deyang City, China

(Excerpts 1 and 2)

**Excerpt 1***Jingyang District High School Teacher Annual Evaluation Rubric*

| Evaluation Item                           | Points  | Self-evaluation | Evaluation by Department | Evaluation by Committee |
|---|---|-----------------|--------------------------|-------------------------|
| A,<br>Professionalism<br>(10 points)      | 1) Do not abuse student physically or mentally; do not discriminate against student; do not seek personal gain using the position of teacher; do not take money for extracurricular consultation; dress properly for class; do not make personal phone call during class. | 10              |                          |                         |
| B,<br>Teaching<br>practice<br>(70 points) | 2) workload (full)  | 4               |                          |                         |
|   | 3) attendance   | 4               |                          |                         |
|   | 4) professional development   | 4               |                          |                         |
|   | 5) class preparation (based on random inspection from the Office of Academic Affairs)   | 6               |                          |                         |
|   | 6) class instruction (based on syllabus, student feedback, and peer review)   | 6               |                          |                         |
|   | 7) timely feedback on homework  | 6               |                          |                         |
|   | 8) meet standard for inspection   | 3               |                          |                         |
|   | 9) organization and supervision of extracurricular activities   | 2               |                          |                         |
|   | 10) participation in research (based on record of ongoing research project)   | 3               |                          |                         |



---

|   |  |    |
|---|--|----|
|   | 11) Presentation or open class in Provincial, Municipal, or district level conferences   | 4  |
|   | 12) mentoring junior faculty member  | 2  |
|   | 13) student and parent evaluation  | 3  |
|   | 14) Class supervisor, department chair, grade coordinator evaluation   | 5  |
|   | 15) supportive of school policy, achieve educational goals, actively engage in teaching as well as research  | 20 |
| C, Education, instruction, and research achievements<br>(40 points) | 16) 10 points for having student ranked the first at <i>Gaokao</i> or other standardized tests, 7 points for the second, and 3 points for the third  | 10 |
|   | 17) winning award in teaching competition (national 7-5 points, provincial 5-3 points, municipal 3-1 points, county 1-0.5 point)   | 7  |
|   | 18) winning award for research (national 4 points, provincial 3 points, municipal 2 points, county 1 point)  | 4  |
|   | 19) Supervising student to win award in competition (for competition organized by educational authorities, national 5 points, provincial 4 points, municipal 3 points, county 2 points; Points | 5  |

---

---

|   |     |
|---|-----|
| deducted half for competition organized by academic associations at corresponding level)  |     |
| 20) Publication or presentation of educational papers (for publication, national 5 points, provincial 4 points, municipal 3 points, county 2 points, half the points for presentation at corresponding level, non-publication, non-presentation paper will get 1 point if submitted to educational research office) | 5   |
| 21) Awarded for outstanding/excellence (national 5 points, provincial 4 points, municipal 3 points, county 2 points, school 1 point)  | 5   |
| 22) Making significant contribution in areas of education, instruction, research or administration that can be considered historical breakthrough (major contribution to the development of the school)   | 5   |
| Total   | 120 |

---

## Excerpt 2

### *Formulas of Instruction Award for Graduating Classes*

The school shall reward teachers who excel in class instruction and student services. The reward for faculty and staff working on graduating classes will be given according to the following guideline:

A, Guideline of Rewards for Instructional Quality Related to College Admission Exam  
(*Gaokao*)

Class instruction is the core of the school's mission. The outputs of the college admission exam are directly related to the reputation of the school and its future development. This guideline is developed to encourage faculty and staff's dedication and creativity in class instruction. This guideline bases reward on both the quality and quantity of efforts.

a) College admission goal:

Graduating class college admission goal = (First tier college admission goal) + (regular four-year college admission goal)

- a. First tier college admission goal = (number of students taking *Gaokao*) \* (average first-tier college admission rate of the top three high schools last year)
- b. Four-year college admission goal = (number of students taking *Gaokao*) \* (average four-year college admission rate of the top three high schools last year)
- c. Academic affair office and Graduating Class office will determine specific college admission goals for each class based on the freshman year entry test results

b) Determination of reward amount

College admission reward amount is calculated based on the year's average admission rate of the top three high schools

a. Base amount:

Graduating class reward = ¥15000 \*number of classes

Repeating class reward = ¥10000 \*number of classes

- i. When perform below the admission goal, the amount corresponding to the percentage below the admission goal will be deducted from the base amount until deduced entirely

b. Reward for out/under-performance:

Based on the formula in A-a, ¥1000 increment for each one more student over the first tier college admission goal from the graduating class (¥700 for student from the repeating class), and ¥1000 deduction for each one less student under from graduating class (¥700 for the repeating class).

c. Reward for prestigious college admission

- i. ¥30000 for having a top rank student at the provincial level, 20000 Yuan for a second rank student, and 10000 Yuan for each student ranked from the third to the fifteenth (liberal art major and science major students will be considered separately)
- ii. ¥5000, 4000, 3000, 2000, and 1000 for having students ranked first to fifth in the municipal (student who's already ranked in the top 15 in the province will not be counted repeatedly)

- iii. Classes with QingHua University or BeiJing University admission will be rewarded on Class basis. For regular class, ¥5000 per student; for advanced class, reward will be calculated the same way after deducting the baseline goal. If within the base line goal, ¥1000 per student. Baseline goal is determined by the graduating class coordination committee.
  - d. ¥2000 for each student with single subject test score ranked top of the province, and ¥500 for each student ranked top of the municipal (this reward is for individual teacher)
- c) Total rewards for each class = (base amount) + (out-performance reward) + (Prestigious college admission reward)
- a. Class coordinator reward = (class reward) \* 8%
  - b. Team work reward = (class reward \* 10%) ÷ (number of subject teachers)
- Note: This reward is only for teachers teaching subjects included in the college admission exam (*Gaokao*)
- c. Quality Education reward
    - i. Reward due to reaching the goals in A-a:  
 Reward based on subject = [(class reward \* 40%) ÷ passing number of all subjects] \* (passing number of subject in one particular subject)
  - d. Coordination and administrative reward
    - i. Senior class coordinator reward = (total amount for all class coordinators) ÷ (total numbers of class coordinators)
    - ii. Administrative staff serving the graduating grade = (total amount for all class coordinators ÷ total numbers of class coordinators) \* 80%

*Appendix B*

## Domain Specifications and Indicators for the Teacher Survey Questionnaire

## Instrument Design and Validation

### Domain Specification

| DOMAINS   | GENERAL INDICATOR   | SPECIFIC INDICATOR   |
|---|---|--|
| <p><b>1.0 SUPPORT FOR IMPLEMENTATION RECEIVED FROM THE SURROUNDING ENVIRONMENTS</b></p> <p>Based on reform implementation literature in the U.S., necessary support from the surrounding policy environments include factors such as reform-related professional development opportunities, necessary resources for implementation, and supports from the principal and survey respondents' colleagues. Professional development programs in SCP are supposed to be not only provided to the respondent, but also helpful to their daily practice, and also sustained over time. SCP-relevant textbooks, supplementary books, SCP-related teaching materials, and financial resources are supposed to be provided</p> | <p style="text-align: center;"><i>(the Agree-Disagree scale)</i></p> <p>1.1 SCP-relevant professional development opportunities</p> | <p>1.1.1 respondent reports that SCP-relevant professional development opportunities were provided</p> <p>1.1.2 ...professional development opportunities were helpful to the SCP practice</p> <p>1.1.3 ... professional development opportunities were sustained over time</p> <p>1.2 necessary resources for implementation</p> <p>1.2.1 respondent reports that SCP-related textbooks were provided</p> |

to teachers.

1.2.2 ... SCP-relevant supplementary books were provided

1.2.3 ... other teaching materials for SCP practice were provided

1.2.4 ... financial resources for SCP practice were provided

### 1.3 support from the principal

1.3.1 respondent reports positive attitude of the principal toward the new curriculum reform

1.3.2 reports the principal's supportive leadership actions with regard to SCP

### 1.4 support from colleagues

1.4.1 respondent reports the generally positive attitude of colleagues toward the new curriculum reform

1.4.2 reports supportive actions of colleagues with regard to SCP

---

*(the Agree-Disagree scale)*

## 2.0 KNOWLEDGE REGARDING

---



## SCP

Under SCP, a teacher's role is de-centered and becomes the facilitator and partner in learning with students. A teacher should foster students' thinking skills, instead of rote memorization. With respect to instruction, SCP calls on more time spent on group activities and individual- or group-inquiry discussions; teachers are supposed to raise probing questions. Teachers are also expected to use multiple methods of assessment.

2.1 Beliefs about a teacher's new role and responsibilities in general under SCP

- 2.1.1 respondent communicates belief that a teacher's role is to be a partner of student (rather than a directive teacher)
- 2.1.2 ... a teacher should serve as the facilitator of student learning
- 2.1.3 ... a teacher should foster students' thinking skills

2.2 Beliefs about instructional strategies regard to SCP

- 2.2.1 respondent communicates belief that use of grouping help promote socially-constructed learning
- 2.2.2 ... flexible grouping strategies help foster learning
- 2.2.3 ... inquiry driven discussions facilitate learning
- 2.2.4 ... use of probing questions facilitate learning

2.2.5 ...use of multiple assessment methods capture learning better

---

**3.0 SELF-EFFICACY IN IMPLEMENTING SCP**

*(the Agree-Disagree scale)*

This domain is to assess teachers' competency and skills in implementing SCP. Teachers are expected to be able to effectively organize group activities, facilitate individual- or group-inquiry discussions, ask questions that are probing, and employ multiple forms of assessment.

3.1 respondent reports self-efficacy in organizing group activities

3.2 respondent reports self-efficacy in facilitating inquiry-oriented class discussions

3.3. respondent reports self-efficacy in using probing questions

3.4 respondent reports self-efficacy in employing multiple forms of assessment

---

4.0 PERCEPTIONS ON THE  
STUDENT-OUTCOME-  
ORIENTATION OF THE TEACHER  
PERFORMANCE EVALUATION  
MEASURES

(not at all—very low—low—high—very high)

---

A teacher performance evaluation system is highly student-outcome-oriented if teacher performance criteria are heavily weighted on student outcome indicators such as student test scores, student graduation rates, student grade detention rates, college admission rates, and student regularity/attendance; but not so heavily on teacher indicators such as: teacher absenteeism, classroom teaching performance, class management, knowledge of subject, instructional design, communication skills, organization and planning of lessons.

4.1 student outcome indicators:

Respondent reports that their performance evaluation measures place more emphasis on these student outcome indicators:

4.1.1 high-stakes, standardized test scores such as “*Zhongkao*” and “*Gaokao*”

4.1.2 other locally-organized standardized test scores

4.1.3 student results of assessments designed by respondent

4.1.4 student grade retention rates

4.1.5. college admission rates ( or high-school admission rates)

4.1.6 student graduation rates

4.1.7 student regularity/attendance

4.2 teacher-related indicators:

Respondent reports that their performance evaluation measures place less emphasis on these teacher-related indicators:

4.2.1 teacher absenteeism

4.2.2 classroom teaching performance

4.2.3 class management

4.2.4 knowledge of subject matter

4.2.5 instructional design

4.2.6 communication skills

4.2.7 organization and planning of lessons

---

*(not all all—seldom—often—very often—all the time)*

## 5.0 IMPLEMENTATION OF SCP

For a teacher is said to be practicing SCP, she or he would report high frequencies in employing SCP-related instructional strategies. More specifically, she or he would frequently use group activity as the primary instructional strategy. If no group activity is involved, she or he would use individual- or group-based class discussion very often to encourage students to share and interact with both the teacher and other students; she or he would frequently raise questions that lead students to think in depth; and she or he would report the use assessment methods other than testing.

### 5.1 group activity

5.1.1 respondent reports higher frequencies in using group activity as the main instructional strategy

### 5.2 inquiry-oriented class discussion

5.2.1 respondent reports higher frequencies in employing group- or individual-based inquiry-oriented class discussions

### 5.3 questioning

5.3.1 respondent reports higher frequencies in asking questions that are probing

### 5.4 multiple assessment methods

5.4.1 respondent reports higher frequencies in using assessment forms other than testing

---

**Background Demographics:**

- Gender/sex
- Years of teaching experience
- Degree of education
- Grade level
- Course subject
- Class size
- school

*Appendix C*

Item-Total Statistics for Three Domains (the pilot version)

## Item-Total Statistics for Three Domains (the pilot version)

|                                 | Scale Mean if<br>Item Deleted | Scale Variance if<br>Item Deleted | Corrected Item-<br>Total Correlation | Cronbach's Alpha<br>if Item Deleted |
|---------------------------------|-------------------------------|-----------------------------------|--------------------------------------|-------------------------------------|
| <i>Domain I: Beliefs in SCP</i> |                               |                                   |                                      |                                     |
| r1                              | 54.3704                       | 15.5499                           | .1801                                | .5926                               |
| r2                              | 53.9259                       | 15.9174                           | .1958                                | .5859                               |
| r3                              | 53.5926                       | 17.7123                           | -.0468                               | .6149                               |
| r4                              | 53.6667                       | 16.3846                           | .4375                                | .5667                               |
| r5                              | 54.7407                       | 18.2764                           | -.1671                               | .6579                               |
| r6                              | 53.7407                       | 15.2764                           | .4848                                | .5443                               |
| r7                              | 54.4074                       | 15.8661                           | .1600                                | .5951                               |
| r8                              | 53.4815                       | 15.1054                           | .6027                                | .5337                               |
| r9                              | 53.7407                       | 15.2764                           | .5722                                | .5389                               |
| r10                             | 54.0370                       | 14.9601                           | .3534                                | .5554                               |
| r11                             | 54.1111                       | 17.9487                           | -.1249                               | .6500                               |
| r12                             | 53.6667                       | 15.6154                           | .4498                                | .5524                               |
| r13                             | 53.6667                       | 16.7692                           | .2197                                | .5832                               |
| r14                             | 54.2963                       | 15.9088                           | .2211                                | .5808                               |
| r15                             | 53.5926                       | 15.7982                           | .2894                                | .5697                               |
| r16                             | 53.7407                       | 14.6610                           | .4957                                | .5334                               |



|   | Scale Mean if<br>Item Deleted | Scale Variance if<br>Item Deleted | Corrected Item-<br>Total Correlation | Cronbach's Alpha<br>if Item Deleted |
|---|-------------------------------|-----------------------------------|--------------------------------------|-------------------------------------|
| <i>Domain: Perceived Institutional Support for Implementing SCP</i> |                               |                                   |                                      |                                     |
| h1  | 43.9643                       | 20.036                            | .666                                 | .672                                |
| h2  | 44.3571                       | 18.757                            | .447                                 | .686                                |
| h3  | 44.4643                       | 20.999                            | .223                                 | .720                                |
| h4  | 43.9643                       | 21.517                            | .282                                 | .708                                |
| h5  | 43.7143                       | 23.323                            | .044                                 | .727                                |
| h6  | 44.6786                       | 19.041                            | .436                                 | .688                                |
| h7  | 44.0357                       | 22.258                            | .193                                 | .716                                |
| h8  | 44.0357                       | 21.295                            | .362                                 | .700                                |
| h9  | 44.1071                       | 24.099                            | -.114                                | .750                                |
| h10   | 44.0000                       | 19.852                            | .564                                 | .676                                |
| h11   | 43.7500                       | 21.528                            | .346                                 | .702                                |
| h12   | 43.9286                       | 17.772                            | .709                                 | .647                                |
| h13   | 43.6429                       | 21.349                            | .427                                 | .696                                |
| h14   | 44.3929                       | 21.877                            | .168                                 | .723                                |

|  | Scale Mean if<br>Item Deleted | Scale Variance if<br>Item Deleted | Corrected Item-<br>Total Correlation | Cronbach's Alpha<br>if Item Deleted |
|--|-------------------------------|-----------------------------------|--------------------------------------|-------------------------------------|
| <i>Domain: Self-Efficacy in Practicing SCP</i> |                               |                                   |                                      |                                     |
| m1   | 66.4643                       | 55.369                            | .595                                 | .933                                |
| m2   | 66.3929                       | 55.951                            | .649                                 | .933                                |
| m3   | 66.3571                       | 55.497                            | .750                                 | .931                                |
| m4   | 66.4643                       | 55.443                            | .667                                 | .932                                |
| m5   | 66.4286                       | 54.698                            | .520                                 | .936                                |
| m6   | 66.6786                       | 55.263                            | .634                                 | .933                                |
| m7   | 66.5714                       | 55.884                            | .506                                 | .935                                |
| m8   | 66.4286                       | 52.921                            | .820                                 | .929                                |
| m9   | 66.5000                       | 55.889                            | .522                                 | .935                                |
| m10  | 66.5000                       | 53.963                            | .746                                 | .930                                |
| m11  | 66.5357                       | 54.851                            | .710                                 | .931                                |
| m12  | 66.5714                       | 56.328                            | .413                                 | .938                                |
| m13  | 66.5357                       | 53.369                            | .802                                 | .929                                |
| m14  | 66.5357                       | 54.999                            | .555                                 | .934                                |
| m15  | 66.5357                       | 53.295                            | .811                                 | .929                                |
| m16  | 66.3929                       | 55.284                            | .643                                 | .932                                |
| m17  | 66.5714                       | 53.513                            | .775                                 | .930                                |
| m18  | 66.4643                       | 54.999                            | .574                                 | .934                                |
| m19  | 66.2857                       | 57.101                            | .599                                 | .934                                |

*Appendix D*

## The Teacher Survey Questionnaire

(Post-Pilot Version, in English)

Dear Teachers:

You are invited to participate in a survey study. This study is being conducted to fulfill requirements of a doctoral dissertation of Teachers College, Columbia University. The purpose of the dissertation is to evaluate the progress of implementation of the China's New Curriculum Reforms in basic education in schools.

The risk of this study is minimum. There's no need for you to identify yourself. The questionnaire and the consent form will be collected and filled separately. At the beginning of the questionnaire, there are items asking about your school and teaching background. That information will be incorporated into group level analyses only. Copies of the questionnaire will be transported out of China shortly after the data collection is completed. Thus, eliminate the likelihood of a breach of confidentiality. In the very unlikely case of such breach, your disagreement with the school's policy objectives could be deemed a sign of your lack of cooperation.

This study has no direct benefit to your participation, either. However, findings of this research can provide valuable insights into the curriculum reform efforts that are going on at your schools.

The survey is not a test. There is no right or wrong answer to the items. Please read the questions carefully and choose the answers that are closest to your real feelings.

Answering this questionnaire will take about 20 minutes of your time.

Results of the survey will be used only for this dissertation and, if possible, related publication in the future.

Thank you very much for your participation.

By signing underneath, you declare that you have read the above statement and agree to participate in this survey study. You know that the aforementioned participation is voluntary. You reserve the right to withdraw at any time.

Signature by the Participant\_\_\_\_\_

Date\_\_\_\_\_

(Note: Your signature will be kept completely confidential.)

## Section A: Background Information

*Direction: Please provide information on your background.*

1. Gender: M\_\_\_\_\_ F\_\_\_\_\_
2. Name of your school: \_\_\_\_\_
3. Address of your school: \_\_\_\_\_  
\_\_\_\_\_
4. Highest degree at the time of the survey:  
\_\_\_\_\_
5. Total years of teaching experience:  
\_\_\_\_\_
6. Grade level that you teach (If you teach more than one grade level, list all grades; current grade goes first): \_\_\_\_\_
7. Main subject you teach (If you teach more than one subject, list all):  
\_\_\_\_\_
8. Size of your class (number of students) for the main subject you teach (If you teach more than one class, report the *highest* and *lowest* numbers of students in class):  
\_\_\_\_\_
9. Think of the average student in your class. What is the typical income range? Choose *one* answer.
  - a. <1,100 RMB/year
  - b. 1,100—10,000 RMB/year
  - c. 10,001—30,000RMB/year
  - d. >30,001RMB/year
10. Will the main subject that you teach be tested on the Gaokao?
  - a. Yes\_\_\_\_\_
  - b. No\_\_\_\_\_

## Section B

*This survey will ask you about your feelings and beliefs about Student-Centered Pedagogy (以学生为中心的教学) or teaching that centers on students. In the survey, Student-Centered Pedagogy is referenced as SCP.*

*Direction: The following items are in the format of a statement. Please circle the number corresponding with your response to the statement according to their designated meaning below:*

1= Strongly Disagree; 2= Disagree; 3= Neutral; 4= Agree; 5= Strongly Agree

| Item  | Response Scale |   |   |   |   |
|---|----------------|---|---|---|---|
|   | 1              | 2 | 3 | 4 | 5 |
| 1. SCP trainings were disconnected from my daily practice.                                  |                |   |   |   |   |
| 2. I was provided with textbooks for SCP implementation.                                    |                |   |   |   |   |
| 3. My colleagues at school tend to apply lecture-based methods.                             |                |   |   |   |   |
| 4. I have been provided with materials that help me implement student-centered activities.  |                |   |   |   |   |
| 5. Teachers in my school regularly discuss issues faced in SCP implementation.              |                |   |   |   |   |
| 6. My colleagues support SCP.   |                |   |   |   |   |
| 7. Professional development programs in SCP were directly applicable to my classroom.       |                |   |   |   |   |
| 8. My colleagues at school are practicing SCP.  |                |   |   |   |   |
| 9. My colleagues and I meet informally to discuss issues encountered in SCP implementation. |                |   |   |   |   |
| 10. I am using the textbooks suited to traditional lecture-style teaching.                  |                |   |   |   |   |
| 11. Financial resources for SCP practice have been provided to me.                          |                |   |   |   |   |
| 12. In-service trainings in SCP were continued for a sufficient time.                       |                |   |   |   |   |
| 13. My colleagues at school tend to control classroom activities tightly.                   |                |   |   |   |   |

## Section C

Scale:

1= Strongly Disagree; 2= Disagree; 3= Neutral; 4= Agree; 5= Strongly Agree

| Items  | Response Scale |   |   |   |   |
|--|----------------|---|---|---|---|
|  | 1              | 2 | 3 | 4 | 5 |
| In my classroom, I believe that...   |                |   |   |   |   |
| 1. Students' achievement is best tested with standardized or multiple choice tests.                            |                |   |   |   |   |
| 2. Grouping of students should promote social interaction among students and teachers.                         |                |   |   |   |   |
| 3. The teacher should serve as the facilitator of student learning.  |                |   |   |   |   |
| 4. The teacher's work should be to mainly transmit knowledge to students.                                      |                |   |   |   |   |
| 5. The teacher should encourage students to think in depth.  |                |   |   |   |   |
| 6. Students should be assessed in a variety of ways, such as projects, essays, multiple choice, or portfolios. |                |   |   |   |   |
| 7. Assignments such as projects help students learn more.  |                |   |   |   |   |
| 8. The teacher should teach students discovery methods of learning.  |                |   |   |   |   |
| 9. In comparison with teacher-centered pedagogy, SCP is better.  |                |   |   |   |   |

## Section D

Scale:

1= Strongly Disagree; 2= Disagree; 3= Neutral; 4= Agree; 5= Strongly Agree

| Items   | Response Scale |   |   |   |   |
|---|----------------|---|---|---|---|
|   | 1              | 2 | 3 | 4 | 5 |
| 1. I can effectively teach using mixed groups of students.  |                |   |   |   |   |
| 2. I can effectively use a variety of assessment methods.   |                |   |   |   |   |
| 3. I find it hard to assist all students in one class.  |                |   |   |   |   |
| 4. I can effectively facilitate class discussions.  |                |   |   |   |   |
| 5. I can effectively ask questions that make students think in depth.                                       |                |   |   |   |   |
| 6. I can effectively design projects that are appropriate for students' developmental stages.               |                |   |   |   |   |
| 7. I can effectively make myself available to all students.   |                |   |   |   |   |
| 8. I am not very good at raising probing questions.   |                |   |   |   |   |
| 9. I find myself having difficulties in designing projects that are appropriate for my students.            |                |   |   |   |   |
| 10. I am confident in helping students in discussions that enter into unfamiliar areas.                     |                |   |   |   |   |
| 11. I am not very good at assessment methods other than written testing (e.g., projects, portfolios, etc.). |                |   |   |   |   |



## Section E

*Please indicate the degree to which your school emphasizes the following elements during staff/teaching evaluations. Please circle the number corresponding to your best response to the statement according to their designated meaning below:*

1=Very low/not at all; 2=Low; 3= Moderate; 4= High; 5= Very High

| Items  | Response Scale |   |   |   |   |
|--|----------------|---|---|---|---|
|  | 1              | 2 | 3 | 4 | 5 |
| Indicate how much emphasis your school places on the following when evaluating teachers/staff: |                |   |   |   |   |
| 1. Students' test scores on standardized tests such as "Gaokao"                                |                |   |   |   |   |
| 2. Teacher absenteeism   |                |   |   |   |   |
| 3. Students' grade retention rates   |                |   |   |   |   |
| 4. Students' scores from other standardized tests  |                |   |   |   |   |
| 5. Teachers' knowledge of subject matter   |                |   |   |   |   |
| 6. Teachers' classroom teaching performance  |                |   |   |   |   |
| 7. Students' college admission rates   |                |   |   |   |   |
| 8. Teachers' class management  |                |   |   |   |   |
| 9. Students' graduation rates  |                |   |   |   |   |
| 10. Teachers' skills in planning of lessons  |                |   |   |   |   |
| 11. Students' scores from teacher-made assessments   |                |   |   |   |   |
| 12. Students' regularity/attendance  |                |   |   |   |   |
| 13. Teachers' communication skills   |                |   |   |   |   |

## Section F

*Please indicate the frequency with which you implement the following practices. Please circle the number corresponding to your best response to the statement according to their designated meaning below:*

1= Never or rarely; 2= sometimes; 3= Often; 4= Very Often

|  | 1 | 2 | 3 | 4 |
|--|---|---|---|---|
| 1. Interactive learning                                |   |   |   |   |
| 2. Allowing students to help plan classroom activities |   |   |   |   |
| 3. Lecturing   |   |   |   |   |
| 4. Facilitating discussions                            |   |   |   |   |
| 5. Flexible grouping of students                       |   |   |   |   |
| 6. Encouraging memorization and rote learning          |   |   |   |   |
| 7. Student questioning                                 |   |   |   |   |
| 8. Using probing question during teaching              |   |   |   |   |
| 9. Controlling teaching plans                          |   |   |   |   |
| 10. Using portfolios to track student development      |   |   |   |   |
| 11. Using different kinds of assessments               |   |   |   |   |
| 12. Designing class activities myself                  |   |   |   |   |

13. Please comment on the implementation of the new curriculum reforms:

---



---

*Appendix E*

## The Teacher Survey Questionnaire

(Post-Pilot Version, in Chinese)

尊敬的老师： 您好！

您被邀请参加一个问卷调查. 这项调查研究是针对哥伦比亚大学教育学院对其博士论文的要求而执行的. 此论文的目的在于调查中学校执行中国基础教育新课程改革的情况.

参与这个问卷调查对你的风险是极小极小的. 你的参与是完全自愿而且匿名的. 你不需要表明自己的身份. 调查同意书和问卷会被分开收集. 在问卷第一部分, 你会被要求提供一些学校及教学等方面的背景信息, 这些是考虑到研究的需要. 这些信息会被综合到集体信息层面. 所有的问卷会在数据收集结束后迅速被带离中国. 以上这些措施基本排除你身份被披露的风险. 尽管这一类身份暴露的现实可能性是微乎其微的, 但是作为研究方, 我有责任向你指出如果出现身份泄露的情况, 您对学校教育改革政策的不同意见有可能被认为是不合作的迹象.

这个问卷也不会给你带来直接的好处. 但是它的研究结果可以对正在学校进行的新课程改革提出有价值的建设性的认识.

这个问卷不是考试所以答案没有对与错. 只需要提供您认为最符合事实的答案. 因此请仔细阅读问题, 选择最接近你真实想法的选项.

回答本次问卷大约需要 20 分钟。

问卷的结果只会被用于本论文以及, 如果可能的话, 和本论文相关的文章发表.

衷心感谢你的参与。

如果您在下面签字, 就表示您已经阅读了以上叙述文字而且同意参加本次问卷调查. 您知道您的参与是自愿的. 您有权力在任何时候撤回对此信息使用的许可。

问卷参与者签字: \_\_\_\_\_

时间: \_\_\_\_\_

(备注: 您在此的签名保证是完全保密的。)

## 第一部分:背景信息

11. 性别: 男\_\_\_\_\_ 女\_\_\_\_\_

12. 学校名称: \_\_\_\_\_

13. 学校地址(全): \_\_\_\_\_

---

14. 所获得的最高学位(到目前为止):

\_\_\_\_\_

15. 教学年限: \_\_\_\_\_

16. 所教年级(如果跨年级教学, 请列出所有年级, 现在教的年级列在第一):

\_\_\_\_\_

17. 所教科目(如果教多个科目, 请列出所有科目):

\_\_\_\_\_

18. 所教主要科目的班级大小(如果跨班教学, 请给出最多的和最少班的人数):

\_\_\_\_\_

19. 所教学生平均家庭经济状况. 请从以下选项中选择 一个答案:

e. 年收入低于 1,100 元

f. 年收入介于 1,100 元与 10,000 元之间

g. 年收入介于 10,001 与 30,000 元之间

h. 年收入高于 30,001 元

20. 您所教的主要科目是属于高考科目吗?

b. 是\_\_\_\_\_

b. 不是\_\_\_\_\_

## 第二部分

提示: 接下来的问题都以陈述句的形式出现. 右边的 1 至 5 的数字代表对此陈述可能的意见 (见下面的定义), 请在最符合您的真实想法的数字上画圈.

数字定义: 1=强烈反对; 2= 反对; 3= 中立; 4= 同意; 5= 强烈同意

| 问题题目   | 1 | 2 | 3 | 4 | 5 |
|--|---|---|---|---|---|
| a. 我得到了与以学生为中心的教学模式相关的专业技能进修与培训.             | 1 | 2 | 3 | 4 | 5 |
| b. 以学生为中心的教学模式培训与我的日常实践相脱离.                  | 1 | 2 | 3 | 4 | 5 |
| c. 对执行以学生为中心的教学有帮助的教科书提供给了我.                 | 1 | 2 | 3 | 4 | 5 |
| d. 我的同事倾向于使用教师讲课的方式授课.                       | 1 | 2 | 3 | 4 | 5 |
| e. 学校为我提供了对执行以学生为中心的教学活动有帮助的材料.              | 1 | 2 | 3 | 4 | 5 |
| f. 我学校的教师定期开会讨论在执行新课改过程中遇到的问题.               | 1 | 2 | 3 | 4 | 5 |
| g. 我的同事们都支持以学生为中心的教学方式.                      | 1 | 2 | 3 | 4 | 5 |
| h. 以学生为中心的教学专业技能进修可以直接运用到我的课堂教学.             | 1 | 2 | 3 | 4 | 5 |
| i. 我学校的同事们都在执行以学生为中心的教学模式.                   | 1 | 2 | 3 | 4 | 5 |
| j. 我和同事们会非正式的聚在一起讨论在执行以学生为中心的教学模式过程中遇到的问题.   | 1 | 2 | 3 | 4 | 5 |
| k. 我使用的教科书更适合用于传统的教学方式.                      | 1 | 2 | 3 | 4 | 5 |
| l. 我得到了实行以学生为中心的教学方式的财政支持 (比如可自由支配资金, 奖金等等). | 1 | 2 | 3 | 4 | 5 |
| m. 我的同事倾向于严密控制课堂教学活动.                        | 1 | 2 | 3 | 4 | 5 |

## 第三部分

数字定义：1=强烈反对；2=反对；3=中立；4=同意；5=强烈同意

| 问题   |   |   |   |   |   |
|--|---|---|---|---|---|
| 我相信，在我的教室里.....                                  |   |   |   |   |   |
| 1. 学生的成绩最好通过标准化考试来评定。                            | 1 | 2 | 3 | 4 | 5 |
| 2. 学生小组活动的学习方式应该促进学生和学生，学生和教师之间的互动。              | 1 | 2 | 3 | 4 | 5 |
| 3. 教师应该是学生学习的帮助者。                                | 1 | 2 | 3 | 4 | 5 |
| 4. 教师应该培养学生的思考能力。                                | 1 | 2 | 3 | 4 | 5 |
| 5. 教师的主要工作就是传授知识给学生。                             | 1 | 2 | 3 | 4 | 5 |
| 6. 对学生的评估的方法应该是多种多样的，比如说学习项目，小论文，多项选择，或者学生成长档案袋。 | 1 | 2 | 3 | 4 | 5 |
| 7. 学习研究项目之类的作业可以帮助学生学到更多的东西。                     | 1 | 2 | 3 | 4 | 5 |
| 8. 教师应该教会学生发现知识的学习方法。                            | 1 | 2 | 3 | 4 | 5 |
| 9. 以学生为中心的教学比以教师为中心的教授法要好些。                      | 1 | 2 | 3 | 4 | 5 |

## 第四部分

数字定义：1=强烈反对；2= 反对；3= 中立；4= 同意；5= 强烈同意

| 题目   |   |   |   |   |   |
|--|---|---|---|---|---|
| 1. 我能有效的运用学生混合小组的方式进行教学。                     | 1 | 2 | 3 | 4 | 5 |
| 2. 我能有效的使用多种方法来测量学生的学习。                      | 1 | 2 | 3 | 4 | 5 |
| 3. 我发现在一堂课内要帮助所有的学生是非常困难的。                   | 1 | 2 | 3 | 4 | 5 |
| 4. 我能有效的推进课堂讨论。                              | 1 | 2 | 3 | 4 | 5 |
| 5. 我能有效的提出让学生进行深度思考的问题。                      | 1 | 2 | 3 | 4 | 5 |
| 6. 我能有效的设计出适合学生发展程度的学习小项目。                   | 1 | 2 | 3 | 4 | 5 |
| 7. 我能有效的让所有小组都能得到我的帮助。                       | 1 | 2 | 3 | 4 | 5 |
| 8. 我不是很能提出盘根究底的问题。                           | 1 | 2 | 3 | 4 | 5 |
| 9. 我感觉自己在设计适合学生的学习小项目这方面有一些困难。               | 1 | 2 | 3 | 4 | 5 |
| 10. 我有信心去帮助学生，即使他们的讨论进入了我并不熟悉的领域。            | 1 | 2 | 3 | 4 | 5 |
| 11. 我不是很擅长运用除书面考试之外的测量方法，比如说学习小项目，学生成长档案袋等等。 | 1 | 2 | 3 | 4 | 5 |



## 第五部分

请指明您的学校在做教师考核的时候对以下成分的强调程度。请根据右边数字代表的程度，在最符合您的真实想法的数字上画圈。

数字定义：1=非常低 / 从未强调；2= 低；3= 一般；4= 高；5= 非常高

| 题目                         |   |   |   |   |   |
|----------------------------|---|---|---|---|---|
| 请指明您的学校在做教师考核的时候对以下成分的强调程度 | 1 | 2 | 3 | 4 | 5 |
| 1. 学生在中考中的考试成绩             | 1 | 2 | 3 | 4 | 5 |
| 2. 教师的出勤率                  | 1 | 2 | 3 | 4 | 5 |
| 3. 学生的留级率                  | 1 | 2 | 3 | 4 | 5 |
| 4. 学生在地方性考试中的成绩            | 1 | 2 | 3 | 4 | 5 |
| 5. 教师专业知识                  | 1 | 2 | 3 | 4 | 5 |
| 6. 教师的教学表现                 | 1 | 2 | 3 | 4 | 5 |
| 7. 学生的高中入学率                | 1 | 2 | 3 | 4 | 5 |
| 8. 教师的课堂管理                 | 1 | 2 | 3 | 4 | 5 |
| 9. 学生的毕业率                  | 1 | 2 | 3 | 4 | 5 |
| 10. 教师的备课能力                | 1 | 2 | 3 | 4 | 5 |
| 11. 学生在教师自身设计的考试中的成绩       | 1 | 2 | 3 | 4 | 5 |
| 12. 学生的出勤率                 | 1 | 2 | 3 | 4 | 5 |
| 13. 教师的沟通能力                | 1 | 2 | 3 | 4 | 5 |

## 第六部分

请指明您执行以下行为的频率度。请根据右边数字代表的程度，在最符合您的真实想法的数字上画圈。

数字定义 1=从来没有 / 非常非常少； 2= 有些时候； 3= 经常； 4= 非常频繁

|                     | 1 | 2 | 3 | 4 |
|---------------------|---|---|---|---|
| 1. 互动式教学            | 1 | 2 | 3 | 4 |
| 2. 允许学生参与策划课堂活动     | 1 | 2 | 3 | 4 |
| 3. 讲课               | 1 | 2 | 3 | 4 |
| 4. 推进课堂讨论           | 1 | 2 | 3 | 4 |
| 5. 灵活的学生分组          | 1 | 2 | 3 | 4 |
| 6. 鼓励记忆和背诵式学习       | 1 | 2 | 3 | 4 |
| 7. 讲课时回答学生提问        | 1 | 2 | 3 | 4 |
| 8. 教学时运用盘根问底式问题     | 1 | 2 | 3 | 4 |
| 9. 独自准备教案           | 1 | 2 | 3 | 4 |
| 10. 运用成长档案袋追踪学生学习进展 | 1 | 2 | 3 | 4 |
| 11. 运用多种评定方法测量学生    | 1 | 2 | 3 | 4 |
| 12. 自己单独设计课堂活动      | 1 | 2 | 3 | 4 |

13. 请对新课程改革的执行情况作出评论：

---

*Appendix F*

## Examination of Model Fitness of Regressions: Casewise Diagnostics for Outliers

| Case Number | Std. Residual | SCP_IMP | Predicted Value | Residual |
|-------------|---------------|---------|-----------------|----------|
| 10          | -2.337        | 12.000  | 21.26           | -9.26    |
| 86          | 2.024         | 30.000  | 21.98           | 8.02     |
| 109         | 2.158         | 31.000  | 22.45           | 8.55     |
| 119         | -2.372        | 13.000  | 22.40           | -9.40    |
| 122         | 2.725         | 32.000  | 21.20           | 10.80    |
| 125         | 2.036         | 28.000  | 19.93           | 8.07     |
| 130         | 2.462         | 32.000  | 22.24           | 9.76     |
| 138         | 2.460         | 32.000  | 22.25           | 9.75     |
| 140         | -2.101        | 14.000  | 22.33           | -8.33    |
| 200         | -2.414        | 13.000  | 22.57           | -9.57    |
| 207         | 2.333         | 32.000  | 22.75           | 9.25     |
| 212         | -2.054        | 16.000  | 24.14           | -8.14    |
| 225         | -2.013        | 15.000  | 22.98           | -7.98    |
| 228         | -2.197        | 15.000  | 23.71           | -8.71    |

a Dependent Variable: SCP\_IMP

b When values are missing, the substituted mean has been used in the statistical computation.

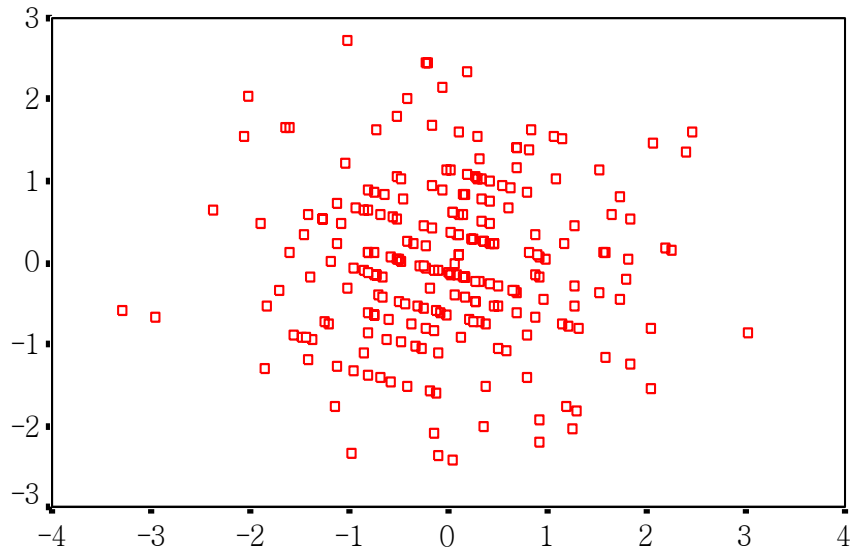
*Appendix G*

## Examinations of Basic Assumptions of Multiple-Regression Models

Examinations of the Normality of Residuals and the Assumption of Linearity for Multiple-Regression Models (the \*ZPRED and \*ZRESID Graph)

Scatterplot

Dependent Variable: SCP\_IMP



Regression Standardized Predicted Value