# ALIGNMENT OF CLASSROOM HISTORY ASSESSMENTS AND THE 7TH NATIONAL CURRICULUM IN KOREA: ASSESSING HISTORICAL KNOWLEDGE AND REASONING SKILLS

by

Mi-Sun Kim

B.A., Chonnam National University, 1986

M.A., Carlow College, 2002

Submitted to the Graduate Faculty of

The School of Education in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2005

UNIVERSITY OF PITTSBURGH

SCHOOL OF EDUCATION

This dissertation was presented

by

Mi-Sun Kim

It was defended on

13 June 2005

and approved by

Kathryn S. Atman, Ph.D., Instruction and Learning, Associate Professor,
University of Pittsburgh

Noreen B. Garman, Ph.D., Administrative and Policy Studies, Professor,
University of Pittsburgh

Ann Jannetta, Ph.D., History, Associate Professor,
University of Pittsburgh

Suzanne Lane, Ph.D., Psychology in Education, Professor,
University of Pittsburgh
Dissertation Director

ALIGNMENT OF CLASSROOM HISTORY ASSESSMENTS AND
THE 7TH NATIONAL CURRICULUM IN KOREA:
ASSESSING HISTORICAL KNOWLEDGE AND REASONING SKILLS

Mi-Sun Kim, Ph.D.

University of Pittsburgh, 2005

This study examined the extent to which middle and high school classroom history assessments align with the educational objectives outlined in the 7th National Curriculum in Korea. In particular, the alignment between the classroom assessments and the educational objectives focused on the level of cognitive reasoning skills and the breadth of historical knowledge. The technical quality of the classroom assessment items, and the extent to which teachers had professional development activities related to the design, use, and interpretation of assessments were also examined. Korean history assessments for the 2004 school year from 22 middle schools and 10 high schools were collected and analyzed. The classroom assessments and the educational objectives were analyzed to examine their alignment with respect to the depth of understanding, breadth of knowledge, and balance of representation. An item writing guideline developed by Haladyna, et al. (1989, 2002) was used to examine the technical quality of the items. A brief survey of history teachers was conducted to obtain information about their assessment related professional development activities.

The results of the study indicated that a relatively large percent of the assessment items from both middle and high schools tended to measure lower levels of historical reasoning than those required by the objectives, resulting in a small percent of items being consistent with the cognitive level of objectives. The distribution of the test items was not balanced across the

objectives, rather they tended to emphasize factual knowledge, and the assessments did not thoroughly cover the span of knowledge represented in the curriculum. There were little differences across different levels and types of schools. However, multiple-choice test items from high school were more likely to assess higher levels of historical understanding than middle school test items. In contrast, the performance assessment tasks for middle school students provided more opportunities to use higher level thinking skills. Most of the items were well developed in terms of formatting and writing test item stems and alternatives. The teacher survey suggested that teachers had little professional development related to the design, use and interpretation of assessments in both their training courses and activities before and during their professional careers.

## ACKNOWLEDGEMENTS

Throughout my life as a history teacher in Korea, the path I have traveled has been graced by the presence of joy in teaching and learning history with my wonderful students who, always eager and curios, posed many questions. The one question that I myself, as a history teacher, have always raised is how could I be the kind of teacher such students deserve; how can I become the best teacher I can for my students. Throughout my doctoral study, I have been privileged to have indispensable support and help from many gifted people at the University of Pittsburgh—people who opened their doors and their hearts to me, to cause lights to go on and to expand the horizon of my knowledge of history education. Now That I have completed my dissertation, I would like to take this fortunate to have the opportunity to thank the many individuals who influenced this study and provided me with the extraordinary chance to pursue my doctorate.

First of all, I would like to express my deep appreciation to my research advisor, Dr. Suzanne Lane, who provided me with the privilege of hands-on experience that helped shape the ways in which I think about classroom history assessments and reasoning skills for this study. I am truly honored to have had her as the chair of my dissertation and received her invaluable guidance and clear direction for keeping my study on the right track. Her breadth of professional knowledge and perspectives inspired me to accomplish a study I thought previously to be unclear. I also wish to extend my sincere thanks to Dr. Kathryn Atman, my academic advisor, who must deserve this study. From the moment of my admission to this university, she has been a mentor and friend, caring about me as a member of the academia family and trusting in my

work, and also keeping me grounded in the realities of education. She has been extremely tolerant and has placed no limits on answering all my questions, providing her warm thoughts and guidance all along the path of my doctoral study. I am also indebted to the impressive members of my dissertation committee and others, al of whom lent their expertise and time. I am honored to have had on my committee, Dr. Noreen Garman, who offered her honest criticism in helping me clarify the ideas presented in this study; Dr. Ann Jannetta, the professor who gave me a greater sense of inquiry-based history teaching through her class, Japanese Pre-modern History, which influenced the ideas of this study. Dr. Gaea Leinhardt, the great scholar who, throughout my doctoral training, enhanced my learning experiences by including me in her professional community to venture into the field of history education. Her high standards for the academic performance of her students, including myself, caused me to stretch beyond my reach, and directed my thoughts along many unexpected avenues and were a memorable challenge and joy of learning.

I owe a considerable debt of gratitude over the years to Professor Anthony Petrosky and his wife, Professor Ellen Bishop. I couldn't have begun my doctoral study without the financial support of *Teacher Education Program*, which is directed by Dr. Petrosky. As a graduate student assistant, I was extremely fortunate to work, specifically, for the supervisors of the Digital Video Project for pre-service teachers, which broadened my instructional perspectives. From the beginning of my new life in the United States, Dr. Petrosky and Bishop have also offered me their warm friendship and solid support, which allowed me to adjust to a foreign culture in a short period time, and helped me complete my doctorate study.

I thank my friend, Hyo-shim Kim, a high school history teacher who analyzed part of the data in order to provide the interrater reliability of this study, and who shared her deep insights

for the analysis of test items with me. Also, I want to express my sincere appreciation to the administrators and thirty-two history teachers of middle and high schools in Korea, who willingly provided me with their assessment materials and shared their ideas on assessments and history education, and from which I have learned a great deal about actual practice of classroom assessments in history.

I would also like to mention my family in Korea—my sisters, brother, sisters-in-law, and brothers-in-law, who never lost faith in me. It is truly difficult to put into words my thanks for their constant financial and emotional support, especially when my husband and I faced difficulties with our life here. Especially, I would like to thank my father-in-law, Kum-sop Yeom, who has grown to love and trust in me as his daughter, and who deserves my sincere respect and gratitude. And I think of my parents, Gwang-yol Kim and Ya-ryo Yang, now deceased, who devoted their lives to me and had great dreams for my life. I dedicate this study to them from the bottom of my heart; they have been always my constant, steady guiding force and their unwavering love and belief in me and have been a comfort and an inspiration throughout my doctoral study.

No one deserves more thanks, however, than my life companion, Min-Ho Yeom. He was the one who gave me the great opportunity to begin anew, and to challenge myself with a new life in the United States, and who taught me stay open to amazing possibilities as a researcher. Throughout the study, he has been my true mentor and friend, enriching my learning experiences academically, supporting me emotionally so that I could overcome the obstacles I faced, and clearing the way for me to complete all the challenges of learning and working in a second language. Without his forbearance and assistance, as ever, this study would not have been written. He is the gift in my life, for which I will ever be grateful.

Dedicated to my family, friends, and students and to history teachers and teacher educators in Korea

## TABLE OF CONTENTS

## LIST OF TABLES

# 1. INTRODUCTION

## 1.1. Background of the Problem

In day-to-day instruction, teachers spend a great deal of their time on the activities related to assessments. Prior to and during the instruction, or after several instructional segments, they assess over time how well students have learned from instruction, or how much they have mastered the knowledge or skills that are expected from instructional plans or curriculum. The results of classroom assessments can be, in general, used for instructional planning, grading students' understanding, and diagnosing students' capacities for enrolling in advanced scholastic levels. As classroom assessments play important roles for student learning, the adequate validity of the assessments has been considered by educational practitioners and researchers.

Regarding history education, there have been ongoing discussions surrounding the matter of how children acquire historical knowledge (Ashby & Lee, 1987; Ashby, Dickinson, & Lee, 1997; Fines, 1987; Leinhardt, 1994b & 2000; Sansom, 1987; Seixas, 1996; Seixas, Stearn, & Wineburg, 2000; Sinatra, Beck, & McKeown, 1992; VanSledright & Brophy. 1992; Wineburg, 1991, 1994, 2001). That is because the basic epistemological structure of history is very different from that of mathematics or science in general, in terms of the development of discourse-based reasoning and explanation based on historical evidence, which is tentative, ambiguous, and uncertain. Historical knowledge also includes multiple cause-effect relationships, various aspects of historical events, and explanation of the past interwoven into a social and political context. Therefore, the agenda to assess children's acquisition of historical

knowledge and their thinking processes for the understanding of history has been evident in a lot of studies in history education (Baker, 1994 & 2000; Bartul, 1993; Blackey, 1993; Grant, 2003; Macintosh; 1987; Scott, 1993).

Seventy years ago, Carl Becker (1931) challenged the absolutism of historical knowledge in which history is merely a collection of discrete knowledge of the past, as the following:

> At all events they go together, so that in a very real sense it is impossible to divorce history from life: Mr. Everyman can not do what he needs or desires to do without recalling past events; he can not recall past events without in some subtle fashion relating them to what he needs or desires to do. This is the natural function of history… In this sense all living history, as Croce says, is contemporaneous: in so far as we think the past (and otherwise the past, however fully related in documents, is nothing to us) it becomes an integral and living part of our present world of semblance (p. 226)

Becker's "Everyman His Own Historian" is a statement of historical subjectivity or relativism that asserts that history is a creation of time and place, based upon men's perceptions of events. Becker means that neither is everyone fully skilled in assaying history nor that everyone is fully equipped with the capacity to use historical evidence in constructing compelling narratives. Rather, he argues that each one is called upon to construct one's own stories of the past, and that these stories reflect the view that one takes from the perspective of one's own place in society.

Our knowledge and understanding of the past is always partial, incomplete, and uncertain. Historical knowledge can be the result of an imaginative reconstruction of past patterns of thought or perceptions, filling in the gaps between evidence and facts about the past (Collingwood, 1971). To reconstruct the past, historians may consciously or unconsciously reflect their own thought, experiences, or approximate imaginations through their presumptions. The past becomes more or less than the reality of the past. The function of historical inference or imagination is to create, as nearly as possible, the reality of the past that actually did exist.

However, the proximity of understanding the past can reduce our attempt of understanding the reality of the past.

Furthermore, historical knowledge can be constructed by individuals who are situated within (the possibilities of) their use of language or discourse (Foucault, 1972). In other words, history does not provide absolute knowledge or truths, rather it is a product of discourse that is controlled, selected, organized, and distributed to the novice. Here, historical knowledge can be placed within core questions about people, power, and struggles in particular historical context. Who decides which interpretations students should learn, and how do we know which historical interpretations are more important for students than other historical interpretations? Why do we hear voices only from a certain historical group of people, but not from others? In this way of regarding history, the reconstruction of the past is always an interpretive and selective process in which particular events, people, or dates in a certain historical interpretation are emphasized and remembered while other events, people, or dates are forgotten, excluded, or deemphasized (Boix-Mansilla, 2000; Levstik, 2000; Seixas, 1996).

In contemporary society, schools have been the major site for learning national collective memory. With compulsory schooling providing history class, students may be exposed to a particular version of history. In any historical narrative, insofar as school history engages with and shapes a collective memory, it is mainly political and selective regardless of the agenda that is focused on national progress or the struggle for human rights (Frisch, 1990; Fulbrook, 1999; Stearns, Seixas, & Wineburg, 2000). As Levstik (2000) claims, nationalism, which may be established through school history, may be shaped by the political, social, and cultural context in which people live and transmitted to the social members in various ways, that is, especially the case for national history. Learning national history, as Gerstle (1997) and Fulbrook (1999)

argue, tends to imply boundaries against outsiders. Thus, defining values of social members and boundaries of self from others may be a key feature of the construction of collective memory.

The significant changes in the view on historical knowledge that have taken place have also been accompanied by the changes in the theories of learning history. In earlier years, in cognitive learning theory, Piaget (1958) extrapolated how the learner acquires knowledge and how the learning environment is constructed, examining social, cultural, and historical factors as well as the role of the instructor. Piaget believes that a learner under 14 years of age should not be taught history in abstractive form; ideas that ancient history should be taught as a concrete subject have been criticized by researchers in the field of cognitive studies in history education (Booth, 1987). However, the Piaget theory implies that a learner's various levels of thinking is part of the process of building, creating, and making mental structures rather than merely absorbing or reproducing products and has played a role cultivating historical research in the cognitive sciences.

In history education, there has been a substantial body of research investigating children's historical knowledge (Ashby & Lee, 1987; Ashby, Dickinson, & Lee, 1997; Sinatra, Beck, & McKeown, 1992; Shemilt, 1987; VanSledright & Brophy, 1992). These efforts have influenced schools, but not as a site that delivers classroom history as a single-version of transmitted collective memory. Ashby and Lee (1987) and Ashby, Dickinson, and Lee (1997) conducted studies examining children's interpretations and reasoning through empathy. They found that children who can use a higher level of thinking recognized differences between past and present mind-set. In the study done by VanSledright and Brophy, (1992), it was also found that children were able to construct imaginative stories about the past, seeing some patterns in some of the stories they created. One study, Schools Council Project 'History 13-16' in 1973

(Shemilt, 1987), even provides the framework spanning the range of historical conceptualization by using historical evidence (from levels 1 to 4). In this study, children at level 3 started to distinguish historical interpretations from the past, and children at level 4 started to become aware of historicity of evidence and to understand historical interpretations as contextual knowledge.

In the last decade, a number of studies about historical learning and teaching have been conducted concerning a more coherent consciousness of the nature of history. For example, in his studies, Wineburg (1991a, 1991b, 2000) provides the framework of habit of minds based on historical reading, using skills to source the identification of historical accounts, to corroborate historical information, and to contextualize historical events. Leinhardt et al. (1994b, 1997) found that instructional explanations are based on rules of evidence, shared experiences, and texts. These involve four general occasions: actions, agents, and causal connections (events); rhetorical stance (themes); expository and descriptive language rather than narrative (structures); and analysis, hypothesis, synthesis, taking perspective, and interpretation (metasystems). The authors believe that students will be empowered to reflect on their own reasoning skills when one or another occasion is emphasized in history class. A number of studies have asked students to generate their own histories, allowing students to explore connections of the past (Bain, 2000; Greene, 1994; Holt, 1990; Voss & Wiley, 1997, 2000). Specifically, argument-based writing tends to invite students to historians' tasks, using historical evidence and building their own positions in the context of a past event (Greene, 1994; Leinhardt, 2000; Voss, & Wiley, 1997 & 2000). Moreover, studies that involve using multiple documents (Perfetti, Britt, & Georgi, 1995; Perfetti, Britt, Van Dyke, & Gabry, 2000; Voss & Wiley, 2000) enhance students' ability to

identify authorship, to evaluate the consistencies of historical information among texts, and to resolve contradictory issues.

These studies have been concerned with classroom instruction that focuses on inquiry-based teaching and learning, which involves students in tasks that promote students to use historical evidence and to construct sound interpretations and perspectives based on informed decisions in contemporary society. These learning environments allow students to problematize the nature of history and to challenge their own view on history, encouraging literacy-based history instruction that goes beyond uncovering discrete facts and simple transmission of historical interpretations or narratives.

As learning theories have changed concerning the epistemological structure of historical knowledge, new tools and instruments for classroom assessment have been developed by researchers and educators in order to measure children's deep understanding of subject knowledge by involving them in real settings (Darling-Hammond & Ancess, 1996; Darling-Hammond, Ancess & Falk, 1995; Wiggins, 1989). Therefore, the changes in classroom assessment techniques must be used to measure children's complex thinking and reasoning of history, including portfolio and performance assessment. Many history classes may still practice historical knowledge as objective facts and basic information to be mastered. Through testing, it may be determined how much students have mastered or absorbed the historical information given in the classroom. Scientific interpretation of history has still influenced the discipline of history based on an analytical approach to historical evidence, and cannot ignore the nature of history's uncertainty leading to interpretive, selective, and imaginative reconstruction. Based on these premises, historical understanding cannot be measured by the memorization of names, events, people, and information provided in the classroom.

Beginning with the 2001 school year, schools in Korea have implemented the 7th National Curriculum. For history education, the curriculum basically emphasizes a heuristic learning environment, focusing on historical thinking and reasoning skills in order to encourage the children to be informative, creative, and global citizens (Ham, 2003; Ju, 2001; Ministry of Education: MOE, 1998; Seo, 2000). In order to accomplish this, the curriculum focuses on children's involvement in the decision-making and problem-solving processes, and in the analyzing, evaluating, and synthesizing of a variety of historical materials. Regarding history education, the curriculum basically intends that students integrate many different cognitive skills, apply knowledge to the real world, and contextualize tasks through constructing their own meanings (Ju, 2001; Seo, 2000).

Applying the constructivist's view to learning theories, the 7th National Curriculum in Korea implies that the student's learning results from a personal interpretation of historical knowledge and is an active process in which meaning is developed on the basis of one's experiences or real world situations (MOEHRD, 1999, 2001; Presidential Commission for New Education Community: PCNEC, 2000). The learning process involves a concern with the experiences and contexts that encourage the student to be able to learn and to facilitate extrapolation or fill in the gaps beyond the information provided (Ham, 2003). Therefore, the curriculum is characterized as grasping meaning by the learner's own construction based on the inquiry of history, rather than being provided with his or her own experiences.

Regarding classroom assessment, the curriculum implies that children's achievement based on heuristic learning and historical reasoning skills should be measured through authentic assessment in order to capture students' diverse perspectives or thinking processes. The curriculum also encourages that teachers can use a variety of methods, such as paper-pencil

assessment, observation, questioning, interviews, etc., and that teachers develop complex test items in order to assess higher order thinking skills (Ministry of Education and Human Resources Development: MOEHRD[1], 1999 & 2001).

## 1.2. The Statement of the Problem

This study was conducted to determine whether middle and high school classroom assessments for Korean history align with the educational objectives embodied in the 7th National Curriculum in Korea; to examine whether middle and high school classroom history assessments are well-developed; and to investigate whether pre-service and in-service programs related to the design and use of assessments were provided to student teachers and practicing teachers. In order to specify this problem, this study answered the following research questions.

## 1.3. Research Questions

1. To what extent do middle and high school history assessments align with the educational objectives outlined in the 7th National Curriculum in Korea?
    - To what extent do the historical reasoning skills required by classroom assessments align with the historical reasoning skills demanded in the objectives of the curriculum?
    - To what extent do classroom history assessments cover the span of historical knowledge that is representative of the curriculum?

[1] In 1999, as a preparation for the arrival of the knowledge-based society, Korean government transformed the Ministry of Education into the Ministry of Education and Human Resources Development, reconceptualizing formal education to encompass human resource development.

- How is the historical knowledge measured by classroom assessments distributed in terms of the targeted objectives in the curriculum?

2. What is the quality of Korean middle and high school teacher developed history assessments?
3. Are there any differences in Korean middle and high school classroom history assessments and the curriculum based on teachers' current teaching assignments (i.e., middle or high school and public or private school)?
4. To what extent have Korean middle and high school history teachers been involved in professional development programs and student teacher experiences regarding classroom assessment?

## 1.4. Delimitations and Limitations of the Study

Delimitations

- Assessments were collected from 22 of 70 middle schools and 10 of 44 academic high schools in a metropolitan city located in the southern area in South Korea.
- Classroom assessments regarding Korean history were delimited to selected-response tests, constructed-response tests, and performance assessments (performance assessments included only a brief requirement and explanation).
- A brief survey about teachers' professional development activities and their pre-service courses in college regarding assessment was included to understand the results of test analyses.

Limitations

- Differences in teachers' concepts of assessment and awareness of the discipline of history might be partially responsible for the depth of historical understanding, breadth of knowledge, and balance of representation regarding the 7th National Curriculum.
- Differences in teachers' education and teaching experiences could limit the results of the alignment between assessments and curriculum.
- Korean educational system—national curriculum, a big class size, or College Entrance Examination—might affect the choice of assessment instruments, levels of historical understanding, or measure of historical knowledge.
- The degree of alignment between classroom history assessments and the curriculum may not be generalizable to entire schools in the city with different socio-economic situations.
- The sampling of classroom history assessments may not reveal complete results of the alignment between assessments and the curriculum.

## 1.5. Educational System, National Curriculum, and History Education in Korea

### 1.5.1. Background of Korean Education

The starting point of the modern Korean education system can be traced to the end of the nineteenth century when Korea started to open its doors to Western forces and to Japan. Education systems in Korea were initially set up during the Japanese colonial period and were Americanized after Korean's independence in 1945 (Kim, 2000b). Today Korean education shares characteristics found in other countries. Leaders of the various educational institutions

and corresponding governmental positions, in large part, have been influenced by the configuration of the education in the United States[2] (An, et al., 1995; An, et al., 1998; Jeong & Armer, 1994). Generally, the objective of education in Korea, like other nations, is to encourage a sense of self-worth in students, to pass on the heritage and goals of the nation, and to create an educated citizenry that will continue the development of the Korean civilization (ECNE, 2000; MOE, 1998b; Kim, 2000b; Shim, 1998).

According to the Education Law[3] promulgated in 1949, the school system in Korea is comprised of six-year elementary schools, three-year middle schools, three-year high schools, and four-year colleges (or universities) (MOE, 1996; Jin, 2003). A four-year college also offers graduate courses leading to the PhD degree. In addition to these schools, there are two-year junior colleges and vocational colleges. At the first level of education in Korea is the national compulsory system, which includes grades one to six. The elementary-to-middle-school transition rate in 2000 was 99.9% according to a statistic provided by the Ministry of Education and Korean Educational Development Institute (MOE & KEDI) in Korea (2000, p. 90). The next level of the educational system is middle school, which involves grades seven to nine.

[2] After the liberation of Korea from Japan in 1945, Korea was ruled by the United States Military Government for three years. Since then, the United States has deeply influenced the Korean education system. In 1946, the basis of syllabi was introduced from the United States to all levels of the school system (MOEHRD, 1999; Yu, 1995). After the Korea War, the United States education mission launched a three-year project for in-service teacher training. In 1954, the government planned to upgrade all teachers' schools to two-year colleges (Seth, 2002). During this time, the government also prepared the 1st National Curriculum that was implemented from 1955 (Lee, 1986). At that time, progressive educational theories by John Dewey were introduced to the field of education (MOEHRD, 1999, Shin & Huh, 1991; Seo, 2000). During the 3rd National Curriculum, Bruner's educational theories influenced Korean curriculum development. Moreover, by the 1980s, American-trained scholars dominated the Ministry of Education (MOE) and the Korean Educational Development Institute (KEDI)—a national research institute designed to advise the MOE—as well as research institutes and educational departments in Korean universities (Lee, 1986). They have influenced Korean education to reform like American system (Seth, 2002).

[3] The Education Law reads as follows: "All citizens have the right to receive education according to their ability; all children should receive at least elementary education and such education as may be prescribed by law; compulsory education is guaranteed in such manner as shall be presented by law; the state is responsible for promoting of school education; and basic matters related to the management of systems of school education and life-long education, the financing of schools and the status of teachers are prescribed by law. (MOE, 1996, p. 48)"

Ninty-nine point five percent ninth grade middle school students in 2000 entered high schools (grades ten to twelve), a figure that includes both general and vocational school (p. 147). Among them, 69% of the students in 2000 were enrolled in general high school, 31% in vocational high schools. General high schools include academic and other specialized high schools that concentrate on science, the arts, foreign languages, and other fields[4]. Currently, tremendous pressure is placed on students to be admitted to a foreign language and science high school. Such high schools offer a greater advantage of being admitted to a college rather than to an academic high school (Seth, 2002).

One of the salient features of education in Korea is that the school system is uniform regardless of whether the institution is private and public (An, et al., 1998; MOE, 1996). Such an education system has resulted in the idea of equalitarianism, often expressed by the term "equalization of education" (Seth, 2002, p 145), meaning that the school system is not just open to all, but also is uniform in content and standard. In the 1960s, rapid economic growth created vastly wealthy families whose children could acquire the best educational benefits at better-ranked schools. Public attitude toward education was influenced by a strong belief that a small group of industry bureaucrats was amassing great wealth while the poor were falling behind. This spurred efforts to create an equitable educational system (Seth, 2002; Yun, et al., 1991). By the campaign of the Mother's Association and the press, the Ministry of Education (MOE) started the practice assigning all students to elementary schools in 1966, to middle schools in 1969, and to high schools in 1974 by lottery[5] (MOE, 1996). Except only one or two high schools

[4] 95% students enrolled in academic high schools and 5% students in specialized high schools (MOE & KEDI, 2000); total high school enrollment was 628,644; 413,091 in academic, 5,184 in arts, 1,160 in physical, 1,226 in science, and 6,231 in foreign language high schools (MOE & KEDI, 2000).

[5] When students transfer to a high school, they need Home School Records (HSR: *naesin*) to determine whether they enroll in vocational schools or academic schools, including foreign language and science high schools. These

in the area, schools became universal for students, illuminating entrance examinations that had caused severe competitions toward better institutions and mental and physical burden for students (Yun et al, 1991).

This uniformity in education includes the establishment of teacher uniformity by regularly rotating teachers, generally in a four-year rotation cycle without any problem, so that certain schools and districts cannot, in general, acquire a reputation for having the best instruction and better qualified teachers. This practice began in the late 1950s with early confrontation between parents and officials. In the 1960s, this system began to take place on a fairly regular basis across the schools (An, et al, 1998; MOE, 1996; Seth, 2002). These efforts of education equalization have brought both quantitative and qualitative improvements and normalization in the schools. The United Nations Development Program ranks Korea as a country of 'high human development,' higher than 80% of the 162 listed countries (UNDP, 2002).

### 1.5.2. College Entrance Examination

With the abolition of the middle school entrance examination and high school equalization policy, parents' desires to have their children receive better education appeared with the College Entrance Examination (CEE). Education in Korea became the new vehicle for people to move up the social ladder by graduating from highly selective universities and taking up influential positions in society (Chung, 1991; Kim, et al., 1994). Since then, the most serious and comprehensive problem in education in Korea has been that all aspects of students activities

---

schools require 80 % of students' HSR, 5% of attendance, 5% of attitude, and 10% of extra activities. Foreign language schools generally require students' HSR in English, Korean language, and mathematics. Science high schools require students' HSR in Korean language, mathematics, and science. Arts schools require 50% of students' performances and 50% of HSR. Athletic high school requires 70% of students' performance and 30% of HSR.

in secondary schools are under the total influence of the CEE. Subject content, teaching methods, evaluation practices, and students' motivation are all determined with regard to the CEE (Bong, 2003; Kim & Lee, 1998; Shin & Huh, 1991). Especially in high schools, principals, teachers, parents, and even students are concerned only with only obtaining acceptance to a college. For both parents and students, the concept of success is highly related to superior academic achievement and the admission to few selected universities (Sorensen, 1994). As time has gone on, the starting point for college preparation gradually moved to lower levels, and most parents sent their children to private institutes and after-school classes in primary school (Seth, 2002; Sorensen, 1994).

From 1945 to 1990, the college entrance examination was reformed nine times, carrying out only a Home School Records System (*naesin*) or combination of written test and a Home School Record System (HSRS) (Yun, et al., 1991). Throughout the 1990s, the government endlessly tinkered with the examination system, changing the rules almost annually, reflecting public opinions (Seth, 2002). Since 1981, the government has managed the college entrance exam, and replaced the exam with the Higher Education Ability Test (HEAT) in 1991 (on the model of the Scholastic Aptitude Test in the United States) emphasizing on the mastery of broader skills rather than the memorization of facts. Now, depending on the college needs, 60-70% of HEAT and 30-40 % of HSRS have been implemented for college entrance.

However, these reforms have brought no fundamental changes of examination systems for the student. In Korea, the fate of students' success in society is dependent upon how well they do on a series of high school tests and whether they enter the best university. School education is merely test taking, nothing more. As critics have argued, the college entrance exams have been driving the entire educational system, which have been reduced to little more

than the constant preparation for and the taking of multiple-choice and short-answer exams. This situation has stifled creativity, hindered the development of analytical reasoning, made schooling a process of rote memorization of meaningless facts, and drained all the joy out of learning (An, et al., 1998; Shim, 1998).

### 1.5.3. National Curriculum and Social Studies

The general picture of the Korean education system represents a highly centralized structure that is run by the Ministry of Education and Human Resources Development (MOEHRD) (An, et al., 1995; Seth, 2002; Yun, et al., 1991). The MOEHRD administers all public education, and controls and manages schools and tertiary education areas. The MOEHRD has the overall responsibility in controlling the national curriculum for elementary and secondary education. Therefore, the formal curriculum of Korean schools is basically uniform across the nation, although school activities or extra curricula are somewhat different, depending on the school level (Lee, 1993; Shin, & Huh, 1991). Individual schools do not have enough autonomy to decide which subjects are taught or even which teaching strategies are used. The right to decide what is taught in schools has historically belonged to the central government (Hwang, 1998). In this educational context, the most pressing issue has been how well the educational goals embodied in the curriculum can be actualized in the teaching and learning processes in each school.

In 1946 when Japanese colonization of Korea ended, the field of social studies was first introduced to Korea by the American military administration as a temporarily constructed. They adopted the social studies system used in Colorado in the United States (MOEHRD, 1999). All levels of school system were managed on the basis of set syllabi (Yu, 1995; Ham, 2003). During

this period, social studies in secondary school involved civics, history, and geography. The subjects were intended to teach the learner how to be a democratic citizen. The most important goal in Korea after liberation from Japan was to cleanse imperial educational systems and to teach democratic ideas (Ham, 2003; Yu, 1995). The three subjects within social studies were independent from one another, so that educational objectives were provided separately from one another. In history and geography classes in middle school, before learning Korean history and geography, students learned about world history and geography was one of characteristics of the curriculum. This learning structure has been maintained so far under the national curricula. In high school, social studies involved: politics, economy, and ethics and philosophy (civics); introduction to geography, human geography, and economic geography' and the history of human races, our culture, and life and literature (history) (Ham, 2003; MOEHRD, 1999, 2001; Yu, 1995).

Since the establishment of the government of the Republic of Korea in 1948, curricula in Korea have been revised seven times, to date, in order to reflect the emerging needs of a changing society and the results of empirical research that the previous curriculum did not appropriate for the current educational circumstances (Ham, 2003; MOE, 1996; Ju, 2001). Table 1 and 2 illustrate the changes of national curricula over time regarding social studies including Korean history.

When the 1st revised curriculum was implemented in 1955, progressive educational ideas developed by John Dewey influenced the field of education and took children's lives and interests into consideration (MOEHRD, 2001; Shin & Huh, 1991; Seo, 2000). Also, this revised curriculum focused on subject matter (subject-centered) based on traditional subject boundaries (Yu, 1995). During this period, educational goals were intended to enhance anti-Communist

ideas in order to resolve the ideological chaos following the Korean War (Ham, 2003; MOE, 1996). In high school, moral education was regarded as important, so it became an independent subject within social studies. General sociology became the subject that taught politics, economy, society, and culture, while geography, history, and moral education remained separate. In addition, in order to restore the nation after the war, social studies both in middle and high schools played the role as one of the major subjects that devoted to social and national

Table 1: Changes of Middle School Social Studies in Korean National Curricula

| Department / Curricula | Social Studies | Korean History | Moral Education |
|---|---|---|---|
| Basic Syllabus (1946-1954) | Geography, History, Civics | | |
| 1st Curriculum (1955-1963) | Geography, History, Civics | | |
| 2nd Curriculum (1963-1973) | Geography, History, Civics | | |
| 3rd Curriculum (1973-1981) | Geography, World history, Civics | Korean history | Moral education |
| 4th Curriculum (1982-1987) | Korean geography, World geography, World history, Civics | Korean history | Moral education |
| 5th Curriculum (1988-1995) | Korean geography, World geography, World history, Civics | Korean history | Moral education |
| 6th Curriculum (1996-2000) | Korean geography, World geography, Korean history, World history, Civics | | Moral education |
| 7th Curriculum (2001- ) | Korean geography, World geography, Korean history, World history, Civics | | Moral education |

Table 2: Changes of High School Social Studies in Korean National Curricula

| Curricula \ Department | Social Studies | Korean History | Moral Education |
|---|---|---|---|
| Basic Syllabus (1946-1954) | Geography, History, Civics | | |
| 1st Curriculum (1955-1963) | Korean history,* World history Geography, General sociology* Moral education* | | |
| 2nd Curriculum (1963-1973) | Korean history* World history* Geography I* & II, General sociology,* Politics & economy, National ethics* | | |
| 3rd Curriculum (1973-1981) | World history, National geography, Human geography, Politics & economy,* Society & culture | Korean history* | National ethics* |
| 4th Curriculum (1982-1987) | World history,* Geography I & II, Sociology I & II | Korean history* | National ethics* |
| 5th Curriculum (1988-1995) | World history, Korean geography* World geography Politics & economy* Society & culture | Korean history* | National ethics* |
| 6th Curriculum (1996-2001) | *Common required subjects* Korean history* Common sociology* (General sociology, Korean geography) *Elective subjects* World history, World geography Politics, economy, Society & culture | | National ethics* |
| 7th Curriculum (2002- ) | *Common basic curriculum* Korean history*, *Elective-centered curriculum* General elective course (Human society and envir.) Intensive elective course | | Moral education* |

* indicates a required subject.

development (MOEHRD, 1999, 2001).

The 2nd revision of the curriculum occurred in 1963, and is termed a 'life or experienced-centered' curriculum (Ham, 2003; MOEHRD, 1999, 2001; Shin & Huh, 1991; Yu, 1995). The Ministry of Education (MOE) officially defined the curriculum as relating to "the total amount of experiences that the students undergo by the guidance of school" (Shin & Huh, 1991, p. 163). During this period, mottos in education included the establishment of national subjectivity, modernization, anti-Communism, and reunification (Ham, 2003; MOEHRD, 2001; Yu, 1995). In middle school, geography for the first grade (7th grade), history for the second grade (8th grade), and civics for the third grade (9th grade) were allotted. This curriculum was structured to help students approach social phenomena based on their spatial and temporal understanding of society (MOEHRD, 1999). In high school, politics and economy became one subject separated from general sociology, so that general sociology was reduced to one subject rather than one domain. While this curriculum was in place, a criticism that history education had been neglected provided an opportunity to regard Korean history as an independent subject for the next curriculum (MOEHRD, 1999). However, the 2nd revised curriculum was as not closely related to life experiences as it professed, so classroom instruction and learning heavily depended on rote memorization of historical knowledge (MOEHRD, 1999; Seo, 2000).

The 3rd curriculum revision occurred in 1973. Because it was influenced by Brunerians in the 1960s in the United States, the 'structures' and 'basic concepts' of history education were emphasized as the main instructional methods (MOEHRD, 1999; Seo, 2000). This education reform was directed toward the educational goals of producing self-directed and future-oriented democratic citizens (Ham, 2003; MOE, 1996). Moreover, regarding history as inquiry, the 3rd revised curriculum required schools to apply heuristic instructional methods emphasizing the

educational environment where the learner participates in the process of problem solving and decision-making (discipline-centered) (Ham, 2003; Jong, 2001; Ju, 2001; MOEHRD, 1999). These educational environments were influenced by the political and economic philosophy of the government that emphasized on anti-Communism, economic development, and national identify. In an effort to nurture national strength and enhance the quality of people's lives, the curriculum placed increasing importance on subjects, such as national ethics and national history (Yu, 1995). Therefore, national ethics and Korean history became independent subjects separated from social studies. Korean history was emphasized more than ever, because of its importance in establishing national identity. Moral education (middle school) and national ethics (high school) also focused on enhancing anti-Communism and national awareness (Ham, 2003; MOEHRD, 1999, 2001).

The 4th revised curriculum occurred in 1981, and was termed as 'humanistic-oriented,' emphasizing the education of the whole person and the integration of subjects (Ham, 2003; MOE, 1996; MOEHRD, 1999; 2001; Shin & Huh, 1991). This curriculum emphasized 'the education of the whole person,' balancing the development of sound body and mentality, the enhancement of intellect and skills, the establishment of moral character, and the development of the awareness of national community (Ham, 2003; MOEHRD, 1999, 2001). Also, this curriculum for high school students integrated subjects, such as politics, economy, society, and culture into sociology I and II, and national, world, and human geography into geography I and II (Ham, 2003; MOEHRD, 2001). In middle school, civics and Korean geography for the first grade (seventh grade), world geography and history for the second grade (eighth grade), and world history and civics for the third grade (ninth grade) were established in order to obtain the educational goals embodied in the curriculum (Ham, 2003; MOEHRD, 1999; Yu, 1995).

The 5th revised curriculum in 1987 was called future-oriented, implying a strong social demand for an information society in the future (Ham, 2003; MOE, 1996; Shin & Huh, 1991). The curriculum, however, maintained the objectives, content, and methods of the 4th revised curriculum (Yu, 1995). Instead of employing a certain political ideas, this curriculum focused on changing the educational methods, the content of inquiry, human and national development, and future orientation. When this curriculum was implemented, Korea hosted the 24th World Olympic Games in 1988 in Seoul. As the nation became increasingly involved in global events, the public strongly demanded education to equip students with the abilities needed in a society of the future (Ham, 2003; MOE, 1996; MOEHRD, 1999). For example, social studies in this curriculum was oriented toward the enhancement of heuristic learning and decision-making skills to allow students to voluntarily participate in informative and global communities in the future. In addition, the characteristics of social studies for each school level were strongly established, expanding students learning experiences from special concepts to social awareness and experiences. However, in high school, integrated subjects, such as politics and economy, society and culture, and Korean and world geography, were divided as in the 3rd revised curriculum, (MOEHRD, 2001).

The implementation of the 6th revised curriculum began in 1996. This curriculum focused on the enhancement of morality and the development of creativity in order to cope with the rapid social changes and current educational problems brought on by democratization, information socialization, high industrialization, internationalization, and reunification with North Korea (Ham, 2003; MOE, 1996; MOEHRD, 1999, 2001). Under the ideas presented in this curriculum, social studies experienced many changes. For example, Korean history was formally integrated into social studies again, although it was taught by its own independent

textbook. In high school, a subject, *common sociology*, that integrated the basic concepts of social studies, appeared as a required course, involving general sociology and Korean geography. Except common sociology and Korean history, students could have more choice in selecting subjects that depended on their interests. The characteristics of social studies presented a synthetic and integrative domain, emphasizing the abilities of rational judgment and decision-making. In terms of instruction and learning methods, this revised curriculum emphasized the processes of thinking and learning and the development of knowledge and skills (Ham, 2003; MOE, 1996; MOEHRD, 1999, 2001).

Lastly, the 7$^{th}$ revised curriculum, termed the learner-centered curriculum was introduced to middle school in 2001 and high school in 2002 with several distinctive features that differentiate it from previous curricula. This curriculum has introduced a national basic common curriculum for grades one to ten and an elective curriculum for grades eleven and twelve (PCNEC, 2000). In high school, tenth grade students learn Korean history and sociology as common required courses (Ham, 2003). Eleventh and twelfth grade students can choose general elective courses, such as human society and environment in social studies. They can also learn social studies by choosing subjects from intensive elective courses, such as Korean geography, world geography, economy geography, Korean modern-current history, world history, law and society, politics, economy, and society and culture. The curriculum is also differentiated to better meet the individual's different learning abilities and his or her needs through differentiated instructions. In general, the curriculum represents a more democratic and decentralized version, which can be best described as school-based and student-centered (MOE, 1998), giving offer students skills to face global and information societies that require self-directed and creative national and global citizens (MOEHRD, 1999; 2001; Ham, 2003).

### 1.5.4. The 7th National Curriculum in History Education[6]

Because the curricula in Korea lead the direction of education and determine the level of educational content or characteristics, they have become benchmarks of education implementations with respect to school education. The 7th National Curriculum is defined as a 'learner-centered curriculum' that is based on developing "a self-directed and creative Korean able to lead a global and informative society in the 21st century" (MOEHRD, 1999, p. 10). This curriculum has been revised based on the changes in learning and knowledge theories and the changes in the quality of democratic citizens with respect to a society that is characterized as informative, diverse, and global (Ham, 2003). Learner-centered education is supported by a differentiated curriculum, which is one of its main features (Choi, 2000; Ju, 2001). For example, in social studies there are three different levels of instructions and assessments that can be applied, depending on whether students' abilities are basic, intensive, or supplementary. This curriculum intends that, based on the content learned at the basic level, differentiated instruction should be offered for students at the intensive level in order to extend and improve a higher level of thinking (historical thinking) skills and for students at the supplementary level in order to make up for a class deficiency (Choi, 2000; MOE, 1998; MOEHRD, 1999, 2001; Ju, 2001).

Another main feature of the 7th National Curriculum in history is that the national common compulsory curriculum is taught from the first to the tenth grades, and the elective curriculum is taught in the eleventh and twelfth grades. Korean history is a required subject for eighth and ninth grade middle school students and tenth grade high school students; and modern Korean history is an elective course for eleventh grade (Ju, 2001, MOEHRD, 1999, 2001). The

[6] The review of the 7th National Curriculum is generally based on Social Studies Curriculum (MOE, 1998), Middle School Curriculum Commentary (MOEHRD, 1999), and High School Curriculum Commentary (MOEHRD, 2001).

subject of history became part of social studies that enhances students' ability to be democratic citizens, although it is taught with its own independent textbooks, not integrated into the content of social studies (Choi, 2000; Yang, 2001). The time allotted for history class has been decreased from six to four hours per week for the tenth grade and from two to one hour per week for the eighth grade (Ham, 2003; Yang, 2001). In the higher-grade levels, students learn intensified knowledge and skills of Korean history and a higher level of application to current issues based on the synthesis of the contents learned at their middle and elementary school levels (MOE, 1998; MOEHRD, 1999, 2001).

The overall objectives of the curriculum for social studies, including Korean history, are the following:

- Basic knowledge and abilities of social phenomena
- Heuristic ability of basic concepts and principles
- Synthetic comprehension of the features of the society and the world
- Creative and relevant problems-solving of contemporary social issues through different types of information
- Ability of participation in a community
- Improvement of the quality of democratic citizens contributing to the development of the nation, society, and the world as well as the individual (MOEHRD, 1998, pp. 29)

The last element is the general learning target in social studies that can be achieved through the accomplishment of the first five elements (Ju, 2001). In order to enhance the quality of citizens, the curriculum emphasizes that students first must be able to use a variety of sources and comprehend basic knowledge and skills, then be able to solve problems that have occurred in contemporary society (MOE, 1998**)**. Based on this basic and synthetic awareness of social phenomena, students can become democratic citizens who successfully participate in their social, political, and cultural communities. Under these overall objectives for social studies, the curriculum offers five general objectives of Korean history and four or five specific objectives

for each unit of Korean history (MOEHRD, 1999, 2001).

The learning environment that the 7th National Curriculum emphasizes is heuristic learning based on students' self-directed learning (Choi, 2000; Ju, 2001). This self-directed learning offers students opportunities to become aware of their own thinking, to make effective learning plans by themselves, to use necessary resources for the class, and to evaluate the effectiveness of their own learning actions (Marzano, 1993). In the learning environment that requires self-realization in their own social lives, students are expected to improve their abilities to solve problems and make decisions applying to real world contexts. In order to accomplish these educational goals, the curriculum emphasizes the ability to use higher-order thinking skills to analyze, evaluate, and synthesize a variety of materials, such as maps, charts, films, statistics, chronology, newspapers, media, pictorial materials, etc. (MOEHRD, 1999, 2001; Ju, 2001). The learning for thinking skills is presented in the description of social studies for the middle school curriculum in teaching and learning methods:

> Social studies class, for the improvement of thinking skills, will involve clarifying the dispositions of concepts logically; discovering principles through reflective thinking; verifying facts through the principles discovered; solving problems creatively; and making a decision by exploring alternatives. Through these kinds of learning processes, students can enhance logical, critical, and creative thinking abilities…. [Teachers] should appropriately decide or select the thinking skills to be emphasized for each part of the learning unit depending on the topic, and it should be considered from the course of instructional plan (MOEHRD, 1999, p. 324).

In the domain of history, the curriculum emphasizes historical imaginative understanding in order to comprehend the meaning of intended historical actions or purpose, and the process of historical decision-making. In order to help imaginative understanding, empathy can be used as a way of teaching and learning of history (Ju, 2001; MOEHRD, 1999, p. 354). For elementary and middle school students, using historical materials/documents and learning about historical figures can also be effective for teaching skills (Ju, 2001).

In light of instructional environments and learning objectives embodied in the curriculum, the methods of student assessment should be aligned with learning goals and conditions. Evaluation standards involve thirteen elements (MOE, 1998, Appendix A) that emphasize classroom evaluation as one of the processes of learning, validation of individual achievement based on evaluation standards, and differentiated evaluation (Choi, 2000). The overall evaluation standards for Korean history for the tenth grade are the following:

- Evaluate students based on assessment elements regarding to the objectives embodied in curriculum.
- For paper-pencil assessment, in order to correspond to differential curriculum, a diverse way of evaluation, such as knowledge attainment, concept comprehension, historical thinking skills, and problem solving, should be used.
- Develop complex test items containing pictorial and statistic materials, maps, or charts, and focus on assessing higher-order thinking skills involving interpretation and analysis of the meaning of the materials.
- Use accumulated observations, questionings, and checking homework as an alternative assessment for participation in the class and attitudes and use them as a tool of instructional improvement.
- Focus on assessing basic historical concepts, comprehension of historical knowledge, and problem solving skills through the heuristic learning and historical thinking skills based on analysis, evaluation, comparison, inferences, and reasoning of historical facts (MOEHRD, 2001).

The curriculum implies that students' achievement of heuristic learning and historical thinking skills can be measured by authentic/performance assessment, because traditional assessment methods are not sufficient to capture students' diverse perspectives or thinking processes (Ju, 2001). However, the curriculum also encourages teachers to use multiple-choice, short answer, and essay test items measuring higher-order thinking skills (MOEHRD, 1999, 2001).

With respect to the educational environment and conditions, the 7$^{th}$ National Curriculum involves the view of constructivists on knowledge, heuristic learning, and performance assessment (Choi, 2000b; Ju, 2001; PCNEC, 2000). According to Ju (2001), the perspectives of constructivism on knowledge, heuristic learning, and performance assessment are related to one

another systematically, thus, these three areas must be balanced in order to accomplish their learning goals. Constructivists view knowledge as resulting from a personal interpretation and constructed in an active learning environment. They believe that learning should be situated in realistic settings, thus, testing should be integrated into the task, and not be considered a separate activity (Spiro, Feltovich, Jacobson, & Coulson, 1992). In short, the main educational intention of the 7th National Curriculum in Korea is to develop students' high-level cognitive abilities, which are necessary in the current social conditions.

### 1.5.5. General Features of History Education

One of the goals for a national history education in Korea has been to teach students nationalism, which the Korean nation has inherited since man-god *Tan Gun* founded the *Choson* Kingdom in 2,333 B.C (MOEHRD, 2003a, 2003b). Through the centuries, Koreans have preserved this legend and it has become a foundation of Korean culture. Because the characteristics of Korean society are generally homogeneous, Korean national history education is devoted to establishing a strong national identity based on the awareness of a national community and the understanding of its heritages (Seo, 2000). It celebrates national achievements, venerates the Asian tradition, and emphasizes a shared Korean experience (Lee, 1998).

As mentioned earlier, students are required to learn Korean history in grades eight and nine in middle school and ten in high school. Through the 3rd, 4th, and 5th National Curricula, Korean history was regarded as an independent subject. However, since the 6th National Curriculum in 1995, history has been integrated into social studies in an approach to incorporate real world experiences and current social issues (MOE, 1998, MOEHRD, 1999, 2001; Ham,

2003). Students in grades eight and nine mainly learn political Korean history, which provides a comprehensive content based on the brief learning in elementary history. For students in grade ten, history textbooks are composed of thematic, cultural, and social Korean history about the pre-modern era and are based on the students' comprehensive understanding of the content knowledge learned in their grades eight and nine (MOE, 1998; MOEHRD, 1999, 2001). Modern Korean history consists of elective course for students in grade eleven. In the mandatory curriculum for history, students have been implicitly restricted from learning a lot of modern history because Korea is the only country in the world that remains divided by two different political-economic ideologies—Capitalism and Communism. These political and ideological conflicts have affected the content of Korean history and have been often regarded as issues that are too serious to teach in history classrooms where it could cause controversy in the public eye (Kim, 2000a; Seo, 2001). Therefore, the content of Korean history textbooks still strongly remains servants of political orthodoxy.

Under the nation-wide education systems, textbooks and teacher guides are the primary materials available for history education. The present textbook publishing system distinguishes between a first class and second class (MOE, 1996)[7]. The copyright of Korean history textbooks distinguished as the first class is held by the Ministry of Education; the content of the textbooks is determined by a national organization (Korean History Publishing Committee). As schools and teachers only use government designated or confirmed textbooks, Korean history education has encouraged a textbook-centered instruction and a one-sided history that the government assumes is desired (Choi, 2001; Seo, 2000). Moreover, teachers and pupils heavily depend on

[7] There are three types of textbooks in Korea. The first class of textbooks is the one whose copyright is preserved by the Ministry of Education. Korean history and Korean language belong to this type. The second class of textbooks is authorized by the Ministry of Education. The third class of textbooks is recognized by the Ministry of Education or the superintendent of school districts. The second and third class of textbooks is published by private companies

textbooks because internal and external assessment is limited to the content of approved textbooks (Shin & Huh, 1991). The content of the textbooks for high school students also functions as the main source for college entrance examinations. Fortunately, since the implementation of 7th National Curriculum, for world history, schools have been given an opportunity to choose their textbooks from the second-class ones that are published by private companies. Through this choice, schools can respond somewhat to the different needs and requests of students (Choi, 2000).

Because of government controls, the content of Korean history textbooks has been a long standing issue among Korean scholars in history. Regarding the substantial body of research conducted in the field of history education, two recent studies have provided significant insight into the improvement of history education. In their studies, C. Song (1999a) and I. Song (1998) analyzed the research of history education published from 1963 to 1998 and after the Korean liberation (1945) to 1996, respectively. As Seo (2000) argues, what is of most interest to history education scholars is looking at to what extent history textbooks reflect the results of studies. These studies found that most of the studies in history education heavily rely on textbook analysis. In C. Song's (1999a) study, 31.3% of the 377 studies published in history education journals have analyzed the content of history textbooks, regarding the accomplishment of research in political, social, economic, and cultural history while 1.3% of the studies are concerned about the improvement of assessment. Also, in the study done by I. Song (1998), 51.8% of the 143 articles published in academic journal circles studied history textbooks, including comparative studies between the textbooks of Korea and foreign countries and between curriculum and textbooks, and an analysis of Korean history textbooks. However, in the journal, *History Education* from 1955 to 1996, only 2.4% of the 165 studies focusing on the methods or

improvement of assessments, and no studies of assessment were done in other academic journals (Song, 1998). In general, these two studies showed that the frequently conducted research areas included historical theories, instructional methods, and the general theories of history education.

In addition to history education in Korea, there have been two organizations devoted to the practice of theory in school settings: one is the Society of History Educational Study (SHES) established in 1955 and the other is National History Teacher Organization (NHTO) established in 1988 (Kim, 2000a). According to Kim (2000a), SHES has a practical origin, and has developed the theories of history education and applied them to school settings. The society has produced theories and methods of history education, school instruction, the analysis of national and international history textbooks, as well as opened a symposium discussing the current issues of history education. Thus, it can be said that this society is the foundation for research in Korean history education (Kim, 2000a). NHTO started its organization in order to practice 'real history education' in schools that produce a sound democratic citizen devoted to social development because it is believed that Korean history has been distorted by the power that enhances one-sided historical perspectives. This organization is devoted to developing a variety of instructional materials and methods, and a case of successful classroom instruction (Song, 1999b). Thus, it can be said that the teacher organization has played a central role for history teachers (Kim, 2000a).

Yet, there remains a big gap between history in academia as a discipline and history in school as a subject. According to a recent study, conducted by Research Committee for the Revision of Social Studies Curriculum of Korean Teacher University in 1997 before implementing the 7$^{th}$ National Curriculum (in Jong, 2001, p. 39), only 10.1% in elementary schools and 1.0% in secondary schools of 439 subjects (teachers and educational stakeholders)

agreed that they were practicing heuristic and problem-solving instruction. In secondary schools, only a few teachers had been practicing role-play or simulations in the classroom, using historical materials. History still might have been taught as the subject delivering discrete facts to the student. The educational environment for college examination, the lack of instructional materials, the content of history textbooks controlled by the government, and teachers' low level of professionalism are important factors that discourage discipline-centered and learner-centered instruction, which the national curricula have emphasized. However, in order to make history in both academia and schools more congruent, the results of research in history education conducted in academia should be applied in school settings in more practical ways.

# 2. LITERATURE REVIEW FOR A CONCEPTUAL FRAMEWORK

## 2.1. Classroom Assessment

### 2.1.1. Introduction

In the past decades, clear trends have emerged in classroom assessment that is associated with the changing theories of teaching and learning, of psychology, and of epistemology. New perspectives on classroom assessment require that teachers have a close understanding of students' learning processes, mediate the development of their intellectual abilities, and enhance the construction of knowledge, forming students' meaning in their lives. These changes are different from the scientific assessments that measure students' achievement through objective methods, reducing the subjectivity while scoring. They are also different from the assessment that measures students' fully mastered knowledge, matching the congruency of observable behaviors with the goals of learning. Rather, the changes imply an active involvement by teachers in monitoring and evaluating students' achievement of learning through the use of various assessment systems that identify what pupils are really doing in the classroom and how they use their learning in the context of real situations. This section reviews the current issues of learning theories with respect to classroom assessment, assessment roles, and assessment tools that measure students' performance, and models for examining the alignment between curriculum and testing.

### 2.1.2. Changing Learning Theories

The changes in learning theories have given significant implications for both curriculum and assessment. Since the early 1900s, the process of learning has been consistently defined. In mid-1950s, behaviorists in educational psychology defined the notion of the learning process as consisting of cognitive associations that result in observable changes in behavior (Cizek, 1997; Glaser, 1984; Shepard, 1991). Shepard (2000) has provided a broad overview of early behavioristic approaches on the perspective of learning as the following:

- Concept of mind replaced by stimulus-response associations
- Accumulation of atomistic bits of knowledge
- Learning tightly sequenced and hierarchical
- Limited transfer, each objective taught explicitly
- Test-teach-test to ensure learning… (p. 5)

According to Shepard (2000), these early learning theories required scientific measurement of ability and achievement, assessing each skill mastered at the desired level, in order to ensure the social efficiency necessary for a students' future role in society. After the late 1970s, when the development of cognitive sciences significantly influenced the theories of learning, these early theories of learning have been strongly criticized by researchers (Glaser & Silver, 1994; Good & Brophy; Resnick & Resnick, 1992; Shepard, 1991; 2000) for their unexpected effects on teachers' instructional planning, students' learning, and assessment practices. For them, the behavioral theories of the mid-20$^{th}$ century influences on teaching and learning do not adequately describe and assess complex thinking processes, problem solving skills, and decision-making processes.

Good and Brophy (1986) suggest that "…learning is an internal, cognitive event that cannot be equated with observable performance… The performance potential acquired through learning is not the same as its reproduction or application in any particular performance

situation" (p. 134). Learning is now viewed as a relative reorganization of permanent knowledge and skills, involving the process of cognitive activities. There has been a substantial body of research evidence in a variety of dimensions related to the cognitive version of learning, regarding knowledge as the subject to be interrogated by the learner. For example, research has studied dimensions of thinking regarding curriculum and instruction (Marzano, et al. 1988); metacognition and metacognitive processing (Costa, 1991; Fountain & Fusco, 1991); scaffolding higher-order thinking skills (Quellmalz, 1991); inquiry-based learning (Collins & Stevens, 1982); differences in expert and novice problem solving (Chi, Glaser, & Rees, 1982); and critical thinking in instruction (Ennis, 1987, 1991; Paul, 1990; Paul & Elder, 2001, 2002).

Cognitive-based learning includes the view that new learning is shaped by prior knowledge (Marzano et al., 1988; Shepard, 2000) that is defined as *schema*. In other words, the learner, when faced with new information, tends to assimilate it into the existing structures or patterns similar to the new ones. If the learner does not have a stored knowledge about similar situations or does not recognize similar topics or patterns, it may be not easy to obtain or solve the problems that he or she faces in the classroom or the real world. According to Marzano et al. (1988), the existing schemata of the learner help recognize the pertinence and importance of new information, allowing inferential skills to fill in the gaps between existing information and allowing orderly searches of memory that the learner needs to recall.

Influenced by the cognitive tradition, the constructivist paradigm focuses more on social interactions and cultural meanings in the tradition of anthropology. In his study, Shepard (2000) offers the constructivist version of principles for learning. According to him, the development of intellectual abilities is socially and culturally constructed. That is, cognitive abilities are developed through socially mediated learning conditions that are guided by parents or other

significant adults, leading them in their interactions with the environment. Also, Shepard uses the earlier study done by Vygotsky, who provided a theoretical model for understanding how social interactions between adult and child could supply both a model of expertise and the opportunity for guided practice. According to Vygotsky in the study by Shepard (2000), the zone of proximal development is "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 19). In addition, learning occurs not only within a social context, but is shaped by cultural perspectives (Resnick & Klopfer, 1989; Shepard, 2000). Constructivism also posits that learners create knowledge from new information in light of their previous experiences (Resnick & Klopfer, 1989). This learning theory provides the view that learners construct their knowledge and understanding within the form of identity and images of possible selves shaped in a certain cultural context; that the learner with exiting knowledge can reason more profoundly, elaborate their study more properly, learn more effectively. In this sense, for the learner, the content knowledge goes from simple to complex, hierarchical to heterarchical, determinate to indeterminate, linear causality to mutual causality, and objective to perspective (Marzano, et al. 1988).

These meanings of learning socially and culturally negotiated cannot be separated from the aspect of the contextualization of skills and knowledge development in a particular community of practice. Reid and Stone (1991), who contrasted the roots of the cognitive perspective with behavioristic approaches, describe, "Students are no longer regarded as empty vessels to be filled with knowledge. They are viewed as inherently active 'apprentice learners' … who benefits from participation in goal-oriented collaborative activities" (p. 8). This

perspective of learning is similar to the situated learning by Lave and Wenger. Lave and Wenger's (1991) concept of legitimate peripheral participation is that learning and development of an identity of mastery occur together as a newcomer becomes increasingly adept at participating in a community of practice. In other words, children are able to use their knowledge not in a new setting, but in the original community of practice, which provides both meaning and support for knowledge use.

Cognitive perspectives focus more on cognitive structures, abstracts of representations, and generalized principles that enable knowledge use in new situations whereas constructivist perspectives depend on learning to interact with an external world where the learner actively engages, constructs, interprets, and brings experiences to bear. Learning occurs within an individual's mind and through the social interactions with people and the environment. Although individual learners internalize their knowledge and understandings by themselves, using cognitive reasoning skills, learning should be understood within a social context and through the ways of thinking in a community of practice. With respect to the changes in these learning theories, classroom assessment must be changed in fundamental ways: The form of classroom assessment must measure representative important thinking and problem solving skills in each of the disciplines. Tasks to be assessed must offer learners opportunities to be real thinkers and producers in an external world that they can interact with.

### 2.1.3. Definition of Classroom Assessment

As classroom learning theories have changed during the last decade, the instruments of classroom assessment also have changed. The instruments now include portfolios, projects, and performance assessments. These changes in classroom assessment techniques are seen as efforts

by schools to closely align themselves with the instructional objectives of the lesson. In order to understand the nature of classroom assessment that is becoming an integral part of the instructional program (McTighe & Ferrara, 1998; Smith, Smith, & De Lisi, 2001), the meaning of classroom assessment first should be defined.

McTighe and Ferrara (1998), considering the term assessment from the Latin root *assidere*, refer to the meaning of assessment as "sitting beside," which includes informal methods, such as observing, questioning, and asking students what to do in order to understand and describe what students know and can do. The original meaning of assessment differs from justifying grades, sorting, and selecting students that schools have placed too much emphasis on.

Cizek (1997) suggests four roles related to the new definition of assessment:

- Any definition of assessment must be applicable to existing, emerging, and future conditions, formats, and contexts.
- It would be desirable for a definition of assessment to convey an attitude that is embraced by educators.
- A definition that recognizes that assessments should serve, as opposed to drive, instruction would be preferable.
- A definition of assessment should provide a link to educational processes that seek the welfare of each student (pp. 9-10).

Considering these conditions, Airasian (1994) suggests that assessment should include "the full range of information teachers gather in their classrooms: information that helps them understand their pupils, monitor their instruction, and establish a viable classroom culture" (p. 5). Borrowing the word from Brown, McTighe and Ferrara (1998) assessment can be defined as "any systematic basis for making inferences about characteristics of people, usually based on various sources of evidence; the global process of synthesizing information about individual in order to understand and describe them better" (p. 2). Moreover, Nitko (1996) provides the meaning of assessment as "a broad term defined as a process for obtaining information that is used for making decisions about students, curricula and programs, and educational policy" (p. 4).

Therefore, assessment refers to a planned process of gathering and synthesizing information in order to accomplish a specific instructional objective and to make decisions about students and educational purposes with the primary benefits being students' understanding of learning.

The term assessment is often used with the terms test and evaluation. However, testing is one form of assessment and is used to determine how much understanding learners possess after instruction (McMillan, 1997), using paper-pencil formats such as multiple choice and essay. Evaluation is the process used to make value judgments about the quality of students' products or performances based on established criteria (McMillan, 1997; McTighe & Ferrara, 1998; Nitko, 1996) such as grading. Moreover, classroom assessment differs from large-scale assessment, such as standardized tests. Large-scale assessment focuses on groups of students with unformed tests, tends to be conducted before and after instruction by outside experts, tends to cover a large domain of content, and is used for teacher and school accountability (McMillan, 2000).

Classroom assessment includes: diagnosing individuals as well as group strengths, weaknesses, and needs; informing students and parents about learning processes; using the process of grading to motivate students to focus on valued knowledge and skills; using standards to judge the quality of student work; and planning instructions with reflections (McMillan, 1997 & 2000; McTighe & Ferrara, 1998; Nitko, 1996). Classroom assessment tends to involve direct implications of instructional quality and curriculum and offer a more personalized feedback of students' performances while large-scale assessments tend to have high-stakes associate with them and are used for the purpose of consequential decisions such as promotion, graduation, admission, certification, or evaluation.

### 2.1.4. Alternative Assessment

As classroom assessment is defined, and learning theories are changed, how do teachers gather and synthesize the information about students' performance, helping students' intellectual growth and their genuine understanding? What is the more authentic assessment—placing the students as a major focus of the assessment tool?

In recent years, an increasing number of schools is turning to assessment practices that ensure all students learn in meaningful ways, resulting in high levels of performance. To overcome the pitfalls of traditional assessment, which are norm-referenced, rote-oriented, and decontextualized tasks, the new assessment strategies have been developed by researchers, policymakers, and educators. These alternatives are referred to as 'authentic assessments' because they look directly at students' products and performances produced in real-life situations (Darling-Hammond & Ancess, 1996; Darling-Hammond, Ancess & Falk, 1995; Wiggins, 1989). However, the term for these assessments is not unified, so that they are sometimes called performance assessment or alternative assessment. Alternative assessments are performance-based assessments, such as portfolios, exhibitions, demonstrations, experiments, group projects, and other methods.

Alternative assessments are defined as the "system that emerges out of a concept of teaching that place students at the center of the learning environment" (Darling-Hammond, Ancess, & Falk (1995, p. 22), different from "the multiple-choice, timed, 'one-shot' approaches that characterize most standardized and some classroom assessment" (McTighe & Ferrara, 1998, p. 3), and able to assess "the complex structuring of multiple skills and knowledge, including basic and higher-order skills, embedded in realistic or otherwise rich problem contexts that require extended or demanding forms of reasoning and judgment" (Messick, 1996, p. 3). Rather

than simply testing isolated skills or retained facts that do not effectively measure a student's capabilities, alternative assessments accurately evaluate a complex extended process that students have learned (Lane & Stone, in press), requiring the student to apply his or her relevant and heuristic skills to real-world situations. Moreover, these assessment strategies require students to create their own knowledge within the context of their own experiences. The perspectives of assessment involve students in learning, not by passively receiving knowledge, but by being actively engaged in doing authentic tasks through personal knowledge socially constructed within an active and collaborative learning environment.

Indicating limitations of current assessment instruments, Wiggins (1993) suggests some postulates for a more thoughtful assessment system that includes:

- Assessment of thoughtful mastery should ask students to justify their understanding and craft.
- The student is an apprentice liberal artist and should be treated accordingly, through access to models and feedback in learning and assessment.
- An authentic assessment system has to be based on known, clear, public, nonarbitrary standards and criteria.
- An authentic assessment makes self-assessment central.
- We should treat each student as a would-be intellectual performer.
- An education should develop a student's intellectual style and voice.
- Understanding is best assessed by pursuing students' questions, not merely by noting their answers.
- A vital aim of education is to have students understand the limits and boundaries of ideas, theories, and systems.
- We should assess students' intellectual honesty and other habits of mind (pp. 47-63).

In short, alternative assessment requires students to develop responses rather than select from predetermined options; requires cognitive and reasoning skills; allows for the possibility of multiple human judgments; directly evaluates holistic projects; and stems from clear criteria known to students.

### 2.1.5. Role/Purposes of Assessment

Whether done prior to, during, or after teaching, the first step in classroom assessment will be to define the purposes of gathering information about students' outcomes. There are other widely used terms of assessment, such as formative and summative assessment, that refer to the purposes of assessment activity and how assessment information is used. Formative assessment refers to any ongoing diagnostic assessment that provides information to help teachers identify students' weaknesses, strengths, or learning characteristics and to improve students' performance and teachers' instruction (Black, 1998; McTighe & Ferrara, 1998; Nitko, 1996; Shavelson & Stern, 1981). This assessment occurs before, during, and after instruction. For example, teachers' short answer questions before instruction, oral quizzes, and classroom discussions about course materials during instruction are included in formative assessment. Brookhart (2001) argues that the use of formative assessment as a way of improving student performance is the central role, but assessment has focused on the teachers' role. In order to involve diagnostic aspects of learning for students as well as teaching for teachers as the key point of formative assessment, self-assessment is considered to be an essential tool for measuring the progress of learning (Brookhart, 2001). According to Black (1998), feedback is sometimes called formative assessment. If feedback takes place as an ongoing process of individual assignment, formative assessment can play a role in integrating assessment into classroom instruction.

In contrast, summative assessment evaluates students' outcomes after a certain period of a learning sequence by accumulating evidence over time (Brookhart, 2001; McMillan, 1997; Nitko, 1996), such as a final exam and dissertation defense (McTighe & Ferrara, 1998). Summative assessment can be used to determine whether the student is promoted or retained, or

whether the student is to be accepted into or rejected from an educational program (Cunningham, 1998). In those terms, these two forms of assessment are different in that formative assessment is internally used to provide initial indicators to improve instructional activities to better students' understanding of learning, while summative assessment is used to offer students' formal grades for parents or external uses, such as acceptance into another educational program.

However, borrowing words from Gipps, Brookhart (2001) avoids making a clear distinction between the two types of assessments. With respect to the changing learning theories, classroom assessment as an ongoing process should be primarily formative, rather than judging students' achievement at the end of the semester or school year. However, Shepard (2000) and Brookhart (2001) emphasize that both formative and summative assessment should be included in the classroom routine as a natural part of the learning process, providing more valuable information with which to make decisions for the next instruction. In fact, in the study of high school students' perceptions of formative and summative aspects of assessment, Brookhart (2001) found that successful students integrated the two assessments for their individual development, summing up their accomplishments to date and realizing learning as ongoing process. They did not use two different assessment systems in explicitly separate ways.

Stiggins and Conklin (1992) studied teachers' practices of classroom assessment; in particular, about the purposes of assessment. They analyzed journal articles written about diverse assessment situations provided by elementary and secondary teachers. The teachers exclusively used three assessment purposes: 'assign grades,' 'diagnosing individual and group needs,' and 'mastery of instruction.' Compared to secondary school teachers, elementary teachers focused less on 'assign grades,' and more on 'mastery of instruction.' Specifically, elementary school teachers tend to use formative assessment, and secondary school teachers tend

to use summative assessment.

As classroom assessment is to provide useful information for instruction and learning, the purposes of assessment might not be exclusively dichotomous. McMillan (2000) provides a wide range of the purposes of classroom assessment in order to collect and synthesize information for intended uses.

- To identify if students have mastered a concept or skills
- To motivate students to be more engaged in learning
- To get students to learn the content in a way that stresses application and other reasoning skills
- To help develop a positive attitude about a subject
- To communicate to parents attitude what students know and can do
- To communicate expectations to students
- To give students feedback about what they know and can do
- To show students what they need to focus on to improve their understanding
- To encourage student self-evaluation
- To determine report card grades
- To evaluate the effectiveness of instructional approaches (p. 5)

Once purposes of assessment are defined, accurate methods/tools to assess reasoning skills or subject-matter knowledge should be selected. For example, the assessment conducted in the classroom will provide strong evidence of what students have done or can do in the class rather than assigning homework or conducting multiple-choice tests (McMillan, 2000).

### 2.1.6. Tools and Forms of Assessment

A broader range of assessment tools is needed to capture important learning goals and to more directly connect assessment to ongoing instruction. In order to achieve that, assessment tools and forms should provide more open-ended performance tasks requiring students to reason critically, solve complex problems, and apply knowledge in real-world contexts. McTighe and Ferrara (1998) provide a framework of classroom assessment approaches and methods that

involve selected-response formats and constructed-response formats. Selected-response formats include multiple-choice, true-false, and matching items. Constructed-response formats are divided into two categories, brief constructed response and performance-based assessment. Brief constructed response involves open-ended questions, problems, prompts, short written answers, or visual representations (concept map, flow chart, or graph). Performance-based assessment includes products, performances, and process-focused assessments. Based on the assessment framework developed by McTighe and Ferrara (1998), the methods and types for classroom assessment in this section will be reviewed.

Selected-response formats

*Multiple-choice item*: Multiple-choice tests are the most common form of assessment in the classroom. There are a number of criticisms of multiple-choice items. Researchers (Mitchell, 1992; Wiggins, 1992; Wood, 1977) argue that the weaknesses of multiple-choice items is that students respond to a fixed list of options rather than generating their own ideas or solutions; that poorly written multiple-choice items can be superficial, trivial, and limited to factual knowledge; that exclusive use of multiple-choice testing for important or high-stakes assessments may shape education in undesirable ways. On the other hand, there are a number of advantages to multiple-choice items. Multiple-choice items can be used to assess a greater variety of learning content in a limited amount of time; students' knowledge of facts, concepts, and principles effectively and objectively; and a wide range of cognitive skills (Cunningham, 1998; McTighe & Ferrara, 1998; Nitko, 1996; Smith et al., 2001; Wood, 1977). Also, for multiple-choice items, the computer scoring process is objective, allowing teachers to obtain test results quickly for students' feedback.

Smith et al. (2001) provide two examples of multiple-choice items:

*Q1*: It is 4 p.m. and you are going to have dinner at 5 p.m. You are hungry now, and you reach for the cookie jar. Then you put the cookie jar back, telling yourself if you have a snack you may spoil your appetite. Freud would say which of the following most influenced your decision?

a. id
b. ego
c. superego
d. libido

*Q2*: Which of the following controls our guilt feelings?

a. id
b. ego
c. superego
d. libido (p. 58)

Question 2 directly asks factual information, requiring recall or recognition to determine which options fits the definition while question 1 requires students to understand the definition of the superego as "the conscience or the location of our guilt feelings when we mess up" (p. 58).

Regarding the multiple-choice item for history, Scott (1993) offers a variety of information with respect to establishing the content and skills to be tested, the process of writing items, and diverse multiple-choice styles. One of them is to categorize the generalization of historical events or characteristics.

*Q*: Classify these characteristic situations according to this scheme:

N = If the situation was typical of a new, growing dynasty
P = If the situation was typical of a dynasty at its peak of power
D = If the situation was typical of a dynasty in decline

(1) A Committee of Regents conducts imperial affairs in the name of an infant emperor.
(2) Literature and painting are encouraged by Imperial subsidies.
(3) Dikes are repaired and irrigation ditches are built (p. 67).

This question can be solved through the analysis of each Imperial dynasty in the instruction, finding the similarities of signs of dynastic growth, strength, and decline. For developing items

that assess higher-order thinking and reasoning skills for multiple choice-items, both Bloom's and Quellmalz' taxonomies have been widely used.

Researchers (Haladyna & Downing, 1989; Haladyna, et al., 2002; Nitko, 1996; Robert, 1993) have summarized rules for multiple choice item writing. Five basic skills of the craft of multiple choice items provided by Nitko (1996) are "1) to focus items to assess specific learning targets; 2) to prepare the stem as a question or problem to be solved; 3) to write a concise correct alternative; 4), to write distractors that are plausible; and 5) to edit the item to remove irrelevant clues to the correct answer" (p. 142).

Haladyna, et al. (2002) conducted a study validating the taxonomy of multiple-choice item writing guidelines through reviewing 27 textbooks on educational testing and 27 research studies. Based on the study, they suggest that in order to write good multiple-choice questions that avoid trick items, it is necessary to use simple vocabulary; put the central idea in the stem; employ positive, not negatives, in the stem; and avoid using the phrases none of the above or all of the above. In their study, Haladyna, et al. (1989; 2002) argue that the accusation of trick items on a test may be an excuse by students who lack knowledge of the subject matter and therefore perform poorly on a test. Difficult vocabulary also places some students at risk; the vocabulary should be appropriate for students being tested. Specifically, while constructing test items, we should state the central idea of the test item in the stem. Although Downing et al. (1991) found that the effects of unfocused items have no significant differences in difficulty or discrimination, they suggested that students studying for a profession have enough experience with multiple-choice items to ameliorate the effects of poor item writing.

*Other selected-response items*: Matching items are used mostly to assess students' ability to associate between terms, definitions, dates, events, individuals, etc. Matching items consist of

'premises,' which are typically numbered and listed down the left-hand column of the test, and 'responses,' which are often letters listed down the right-hand column. True-false items are frequently used to assess the acquisition of facts and principles (Cunningham, 1998). One of the biggest advantages of true-false and matching items is that they can include a large number of items and cover a lot of content within a limited amount of time. It is easy to construct and score true-false and matching items. However, for these two types of assessment, it is not easy to construct items that measure higher-order thinking skills. In order to improve true-false questions, Cunningham (1998) suggests several roles: assessing 'higher level' cognitive abilities; reducing the effect of guessing by underlining a word or clause in the statements; requiring students to correct false statements; grouping short true-false items under a common questions or statement heading. In order to create matching items, he suggests: including content that is homogeneous; providing no more than ten pairs of premises and responses; including short response statements; and offering more matching exercises than premises.

These selected-response items tend to provide an unintended educational message, selecting 'one correct answer,' as the primary goal of education. They also do not adequately measure creative or critical thinking, oral communication, and social skills in real-world situations (Cunningham, 1998; McTighe & Ferrara, 1998; Nitko, 1996). Moreover, constructing items that measure reasoning skills is time-consuming. By recognizing the limitations and strengths of selected-response items, teachers should incorporate them appropriately with other approaches, thus providing a balance of adequate learning outcomes.

Constructed-response formats

*Constructed response items*: Constructed response items involve short answer questions and extended response questions, including essay tests. *Short answer questions* appear to be less

subjective than extended response questions. The question provides teachers with exact information about what students know or do not know. However, without preparation in constructing a good set of short answer questions, these types of questions tend to measure only the acquisition of trivial factual information rather than reasoning processes. In order to construct better short answer items, teachers should avoid items that measure only trivia, and, instead, write each question or statement about a specific problem. In this category, McTighe and Ferrara (1998) include open-ended questions, problems, prompts, and visual representations that require judgment-based evaluation (criteria), which can be extended response questions.

*Extended response questions* (essay formats) tend to be used effectively in measuring instructional objectives, knowledge of issues, and ideas within subject matter (Resnick & Resnick, 1992) as well as measuring divergent thinking. Divergent thinking is a form of cognitive functioning that involves generating many different answers to a single problem with the correctness dependent upon a subjective evaluation of answers as being abstract and flexible. The biggest disadvantage of essay tests is low reliability. According to Cunningham (1998), there are two sources of measurement error that create unreliability for essay tests: "1) lack of agreement among graders, and 2) low internal consistency among items" (p. 110). If there are several classroom teachers working together with one subject, they can minimize unreliability by carefully constructing test items and delineating scoring rubrics. In order to construct better essay questions, cognitive processes, as specified in the instructional objectives, should be assessed with novel materials and clearly defined questions (Cunningham, 1998).

Wilson, et al. (1993) and Blackey (1993) provide the general direction for historical writing on essay tests, which involve an introduction, body, and conclusion as in professional historical writings. According to the authors, the introduction should include brief descriptions

of time and place setting, and one or two sentences of thesis explaining major categories of evidence to support the thesis. In order for the essay to be a strong answer, the sentences in the body of the paper should include sound argument or opinions based on the relevant facts presented or generalizations from factual materials provided (Blackey; 1993) as well as coherent sentences that include transitions. The conclusion should provide the results of the argument or ideas presented in the body of the paper, rather than merely repeating the main point of the essay or providing a summary of it.

Performance-based assessment (alternative): Performance-based assessment is defined by Stiggins (1994) as an assessment system "involves students in activities that require the demonstration of certain skills and/or the creation of specified products" (p. 160). This assessment is required in order to direct judgments and evaluations of student' activities, so that it is more likely to reveal students' understanding of subject materials (Resnick & Resnick, 1992). In the light of these characteristics, the tools for performance assessment could include observations, clinical interviews, reflective journals, oral presentations, work samples, projects, and portfolios.

McTighe and Ferrara (1998) divide performance-based assessment into products, performances, and process-focused assessments. Products include " written products (essays, research papers, laboratory reports), visual products (e.g., two-and three-dimensional models, displays, videotapes), aural products (e.g., an audiotape of an oral presentation)" (p. 16), and other types of products that show learners' proficiency or ability. Performance assessment is the evaluation system by which teachers can directly observe students' actual activities, including oral presentations, demonstrations, and inquiry-based classroom discussions. Process-focused assessment is used in order to obtain information or gain insight into students' cognitive

processes and learning strategies. This form of assessment can employ oral questions or require students to 'think out loud' while making explicit students' thinking or decision-making processes.

A construct performance task consists of the performance task itself and a clearly defined scoring scheme or rubric (Nitko, 1996; Marzano, et al., 1993). In order to properly evaluate the quality of student performance, Marzano, et al. (1993) suggest the technique for constructing the task and rubrics. First, content standard should be identified, such as declarative or procedural knowledge. For example, "understands that war forces sensitive issues to surface and causes people to confront inherent conflicts of values and beliefs" (p. 27). The second step is to structure the task around complex reasoning skills, such as comparison, decision-making, or problem solving. Next is to identify standards of information processing that require students to access accurate and valuable information. Then, habits of mind and collaboration/cooperation standards should be identified. Finally, the task should include effective communication techniques for students' final reports, such as written reports, demonstrations, or performances. These procedures have to be revised whenever new standards of categories are added. Once a performance task is constructed, a scoring rubric can be developed. According to Jacobs (1997), a scoring rubric should be used as a form of feedback rather than as a grading system. As indicated in an example of performance task in Nitko's book (1996), Jacobs also argues that students should use the rubric to edit and analyze their own work in order to improve. The important intention here is that assessment should play an important role assisting students in developing the knowledge and skills for their understanding. This process can be a means of self-assessment for the student. Students are more likely to be able to perform well if they know what constitutes quality performance (Jacobs, 1997; Shepard, 2000).

According to the supporters for performance-based assessments (Hambleton & Murphy, 1992; Linn & Gronlund, 1995; Marzano, et al., 1993; Shepard, 1991, 2000; Stiggins, 1994; Wiggins, 1990), the advantages are that: 1) performance-based tasks include more explicit complex learning targets than do selected response tasks; 2) these tasks are associated more with new learning theories that emphasize assessing prior knowledge and complex reasoning skills and involving inquiry-based activities; 3) these tasks tend to integrate students' knowledge, skills, and abilities from a variety of sources; and 4) these tasks assess students' thinking processes and broaden their learning boundaries. However, there have been discussions about the disadvantages of performance-based assessments (Hambleton & Murphy, 1992; Linn & Gronlund, 1995; Ruder & Boston; Stiggins, 1994). First, crafting high quality performance tasks and scoring rubrics take a great deal of time and effort. Completing performance tasks requires a great deal of time and effort. In addition, performance tasks may discourage less able students from completing the tasks, and performance tasks cannot cover all learning targets.

### 2.1.7. Sources of Validity Evidence

It is important to note that assessment results require sound interpretations for a given use of test scores, based on logical and empirical evidence; that is, validity evidence (APA, AERA & NCME, 1985; Crocker, 2003; McMillan, 2000; Messick, 1989; Nitko, 1996; Linn, et al., 1991; Quellmalz, 1991). Validity concerns the accuracy of inferences from test scores rather than the characteristics of the instruments to be assessed (McMillan, 2000; Nitko, 1996). In order to make correct decision about using the results of assessment, the information being gathered should be valid. Airasian (2000) offers key aspects of assessment validity.

- Validity is concerned with this general question: To what extent will this assessment information help me make an appropriate decision?

- Validity refers to the decisions that are made from assessment information, not the assessment approach itself. It is not appropriate to say the assessment information is valid unless the decisions or groups it is valid for are identified. Assessment information valid for one decision or group of pupils is not necessarily valid for other decisions or groups.
- Validity is a matter of degree; it does not exist on an all-or-nothing basis. Think of assessment of validity in terms of categories: highly valid, moderately valid, and invalid.
- Validity is always determined by a judgment made by the test user (p. 20).

With respect to analyzing test items for classroom assessment, here, two criteria presented in literature to judge the validity will be reviewed.

*Content validity evidence*: Content validity refers to the evidence that judges the "relevance of the test content to the content of a particular behavioral domain of interest" and "representativeness with which item or task content covers that domain" (Messick, 1989, p. 7). Content validity focuses on the content taught in the classroom, represents the curriculum framework that school districts state, or contains important specifics of the target subject (Nitko, 1996). Linn, et al. (1991) includes content quality and content coverage/breadth as content validity. Moreover, validity standards offered by APA, AERA & NCME (1985) consider content validity to include themes, wording, and format of the test items, tasks, or questions on a test. Poor quality of assessment and instructional materials can produce invalid results of assessment.

Content validity, therefore, provides judgmental evidence in support of domain relevance and representativeness of the content of the test instrument (Messick, 1989). Content representativeness may be the only aspect of validation that can be completed prior to administering the test and reporting the results (Crocker, 2003). In preparing a test blueprint, teachers can review whether the test items represent the content domain and make a decision as to which test items should be included or excluded. Thus, before administering a test, each test

task should be examined in order to determine whether it matches learning targets and important content; each test task should be reviewed for its relevance, representativeness, meaningfulness, and accuracy; and the assessment procedures should be judged for quality of the tasks as a whole (Nitko, 1996). Specifically, content validity should be considered in order to compare what has been taught with what is being assessed. This is called 'instructional validity' (Anderson, 2003; McMillan, 2000). In order for classroom assessment to be valid, the proportions of test items should correspond with the emphasis given to instructional objectives and content of units.

*Substantive validity evidence*: Several researchers consider the types of thinking skills and processes that appear in the assessment as one of the criteria that judges the validity of assessment (Linn, et al., 1991; Messick, 1995; Nitko, 1996; Quellmalz, 1991). Substantive validity refers to evidence that judges whether the assessment measures the cognitive complexity and processes emphasized in the framework of school curriculum. If a test instrument covers a sufficiently wide range of thinking skills and processes, the assessment can be valid in terms of cognitive complexity and worthy learning. Current learning theories focus on students' thinking abilities and processes that can be applied in real world situations, specifically for performance assessment. Also, skillful learners can construct their meaning by calling upon a variety of thinking strategies. In the case of history, classroom assessment should advocate the depth of historical understanding, and measure reasoning skills, such as the ability to hypothesize or evaluate significant past events.

### 2.1.8. Models of Content and Alignment

For accurate inferences about students' learning and their growth over time, classroom assessment must measure the knowledge and skills deemed valuable and described in state

curriculum or standards. Alignment is defined by Porter (2003) as "the degree of agreement between a state's content standards for a specific area and the assessment(s) used to measure student achievement of these standards" (p. 21). Here, assessment includes classroom tests and state-, district-, or school-developed tests; assessment also involves homework assignments, portfolios, interviews, observations, projects, and presentations. Alignment is also defined by Webb (2001) as "the degree to which expectations and assessments are in agreement and serve in conjunction with one another to guide an education system toward students' learning what they are expected to know and do" (p. 4). Therefore, alignment is a quality of the relationship between learning expectations and assessments. The study of alignment can illustrate the relationship between expectations and assessments and can be a valid indicator of school improvement by modifying either the expectations of students or the assessments.

There have been studies conducted about the alignment of states' content standard with large-scale assessments, classroom assessment, and instruction. Webb (2001) conducted a study determining the alignment of assessment systems (math, science, and language arts), proposing five categories as criteria. The categories and criteria include content, articulation, equity and fairness, pedagogical implications, and system applicability. Content standards include the criteria of depth of knowledge, categorical concurrence, range of knowledge, structure of knowledge, balance of representation, and dispositional consonance. Depth of knowledge refers to consistency between the cognitive demands of the standards and the cognitive demands of assessment items. Webb (2001) further develops four levels of cognitive complexity in the category of depth of knowledge: recall, or using simple skills or abilities (level1); comprehension and some subsequent mental processing (level 2); strategic thinking including reasoning and planning (level 3); and extended/higher order thinking (level 4). Webb suggests that, in order to

meet adequate cognitive complexity, at least 50% of the test items should correspond to higher cognitive levels presented in the objectives.

Range of knowledge criterion refers to judging the agreement between multiple dimensions of standards and the assessment intended to evaluate the standards. If a standard requires multiple dimensions, a test item measuring only one dimension would not be aligned with the standard for the criterion of range of knowledge. Webb suggests that more than 50 % of the objectives for a standard should meet at least one matching test item. Lastly, the balance of representation criterion refers to "the degree to which one objective is given more emphasis on the assessment than another" (p. 10). That is, this criterion judges the distribution of the assessment items across the objectives. Webb describes the index he used to judge the balance of distribution of assessment. If an index value is close to one, most of the objectives are measured by the equally balanced test items, and the assessment is balanced. If an index value is close to zero, one or two objectives are measured, and the assessment is not balanced. Webb suggests that an index value should be .7 or higher in order to meet this criterion.

Another model, developed by Achieve (Rothman, Slattery, & Vranek, 2001), uses the following four categories: content centrality, performance centrality, challenge, and balance and range. The content centrality criterion examines the degree of concurrence between the content of the test items and the content of the related standards. The performance centrality criterion evaluates the alignment between the cognitive demands of the test items and the related standards. This criterion is consistent with the depth of knowledge category studied by Webb (2001). Challenge category includes two factors for evaluating sets of test items: source of challenge and level of challenge. The criterion for source of challenge judges the difficulty of individual items because of its level of knowledge and skills. This factor also evaluates other

reasons that are not related to the subject matter, such as unfairness regarding students' background knowledge. Level of challenge focuses on whether the set of items span an appropriate level of difficulty for the target group of students. Balance examines whether the test items assessed emphasizes a certain objective presented in the standards. This criterion differs from the balance of representation in the Webb study (2001), which evaluates the evenly distribution of test items across target objectives. Range evaluates the extent of coverage or breadth, that is, it examines whether an assessment contains test items that measure knowledge and skills representative of the content domain documented in the standards.

Measuring content and alignment, Porter (2002) has developed a model that uses an alignment index, which calculates the value of content emphasized in standards and the value of the number of items that appear in the assessment regarding the same content. Considering the values of the index range (from 0 to 1.0, with 1.0 indicating perfect alignment), Porter describes, "tests that are not a sample of items from a domain, whereas standards represent the domain. Thus, perfect alignment should not be expected" (p. 6). In other words, assessments are not expected to cover every content discussed in the standards, rather they are expected to cover a sample of the content domain representative in those standards. Bhola et al. (2003) argue that using more a complex model to align assessments to standards will provide a less likely match between test items and the standards.

There are several examples developed by Project 2061 (2002) that evaluate the degree of alignment between assessments or textbooks and content standards. Among the models, Stern and Ahlgren (2002) analyzed assessments presented in curriculum materials, with respect to how well these assessment tools contribute to the attainment of specific ideas in benchmarks and standards. In order to analyze nine widely-used middle school science curriculum materials,

they used three criteria: aligning goals, testing for understanding, and informing instruction. The criterion aligning goals evaluates the degree of alignment between the test items in the curriculum materials and the goals documented in the benchmarks or standards. This procedure examines the degree to which the assessment has the potential to reveal whether students have attained the necessary and sufficient number of ideas of presented in the benchmarks or standards. The second criterion, testing for understanding, focuses on cognitive complexity and whether test items aid students in their understanding of a sufficient amount of subject knowledge, including both familiar and novel tasks. According to Stern & Ahlgren, criteria one and two evaluate based on the number of assessment items that meet the criteria of the indicators; thus, reviewers consider whether all benchmarks are assessed and how many test items are included for each of the standards. The last criterion, informing instruction, evaluates whether information gathered from students' responses can be used to inform or modify subsequent instruction. Stern & Ahlgren found that with the exception of one instructional material, all other materials do not help students provide important science literacy ideas and offer poor information on alignment, understanding, and informing instruction.

In their studies, Lane, Parke, & Stone (2002) and Lane, Parke, Stone, Hansen, & Cerrillo (2000) provide schemes that seek the impact of the Maryland performance assessment program. The former study, conducted by Lane, et a. (2002), examines the degree of alignment between instructional practices in math and the Maryland Learning Outcomes (MLOs) and Maryland School Performance Assessment Programs (MSPAP). The latter study, conducted by Lane et al. (2000), investigates whether social studies classroom activities align to MLOs and MSPAP. According to Parke & Lane, MSPAP was developed to assess students' academic growth toward the MLOs. To examine the extent to which mathematics or social studies classroom practices,

including instructional materials, reflect the MLOs and MSPAP, this study used a three-coding scheme: instruction, assessment, and test preparation activities involving several components. Moreover, MLOs for math and social studies and the format and content of MSPAP were provided as one of the coding schemes. For example, these studies include the analysis of the process learning outcomes defined in the MLOs in math and social studies: problem solving and reasoning skills for math; skills and process; valuing self and others; understandings and attitudes; and explaining social studies with no reading and with reading for social studies. These authors evaluated overall similarity of classroom practices to MSPAP, using a 5-level scale for social studies and a 6-level scale for math ranging from not at all similar MSPAP to very similar to MSPAP. They found that teachers use MSPAP-like classroom practices more in their instruction than in their assessments, and more in their test preparation activities than in their instruction.

Taken all together, the alignment models mentioned above tend to look at the relationship between assessment and standards, considering content, cognitive complexity, breadth of knowledge, balance of representation, and consistency between the number of items used on the assessment and emphasis in standards and benchmarks. However, Bhola et al. (2003) consider the problems that have been presented in aligning tests. These problems occur when alignment criteria: look at holistic interpretations rather than at particular content standards; classify performance into more than two categories; and include training reviewers of alignment tests. They are cautious that many standards are multidimensional and require categorical concurrence defined by Webb (1997). Therefore, the set of assessments must measure across the full range of performance categories, and training reviewers can resolve the failure of matching alignment evaluations.

## 2.2. Taxonomies of Higher Order Thinking and Reasoning Skills

### 2.2.1. Introduction

Reasoning skills are cognitive or mental operations using knowledge that employs one's understanding, deriving a conclusion from certain premises. The cognitive reasoning process employs knowledge to approach the task of solving problems and making a decision, or to engage in critical thinking and generalization. Several frameworks for assessing cognitive and reasoning skills have been developed. Commonly used frames of reference are Bloom's Taxonomy of the cognitive domain (Bloom, Englehart, Furst, Hill, & Krathwohl, 1956), Quellmalz's framework for evaluating the reasoning process (Quellmalz, 1987), and Marzano's core thinking skills derived from psychological and philosophical literature and learning standards for the framework of performance assessment (Marzano, 1988, 1992, 1993).

### 2.2.2. Bloom's Taxonomy of the Cognitive Domain

> Arguably, one of the most influential educational monographs of the past half century is the *Taxonomy of Educational Objectives, The Classification of Educational Goals, Handbook I: Cognitive Domain*. Nearly forty years after its publication in 1956 the volume remains a standard reference for discussions of testing and evaluation, curriculum development, and teaching and teacher education. A search of the most recent Social Science Citation Index (1992) revealed more than 150 citations to the Handbook… (Anderson & Sosniak, 1994, p. vii).

Since its publication, "Bloom's Taxonomy," as it is frequently referred to in deference to Benjamin Bloom, has contributed significantly to educational theory and practice and has been used by educators in every subject area at every grade level. Bloom's Taxonomy is one of the most widely accepted models of cognitive abilities and educational objectives in teaching. The model has been a framework for classifying statements of teachers' expectations of what students must learn and how they must be evaluated (Krathwohl, 2002). The taxonomy reflects

the influence of behaviorism that characterized both educational and psychological theory in the 1950s.

> We are reading about an attempt to build a taxonomy of educational objectives. It is intended to provide for classification of the goals of our educational system. It is expected to be of general help to all teachers, administrators, professional specialists, and research workers who deal with curricular and evaluation program (Bloom, et al., 1987, p. 1).

According to Morzano (2001), Bloom's Taxonomy had a strong influence on evaluation, but it had a minimal effect on curriculum. By 1970, Ralph Tyler established the model of evaluation that presented an objective-based evaluation in which instructional experiences were evaluated based on the accomplishment of explicit goals. The more explicitly the goals were stated, the more precisely the instruction could be evaluated. Bloom's taxonomy proved a powerful tool for objective-based evaluation (Marzano, 2001).

In the 1980s, Bloom's Taxonomy was widely used as the model for designing statewide testing items that measure higher level skills because its six categories were interpreted by most educators to be hierarchical (Anderson & Krathwohl, 2001; Krathwohl, 2002). Bloom (1987) also describes, "the whole cognitive domain of the taxonomy is arranged in a hierarchy, that is, each classification within it demands the skills and abilities which are lower in the classification order" (p. 121). The categories were ordered from simple to complex and from concrete to abstract: Mastery of simpler categories was a prerequisite for accomplishing the next, more complex one. Much research has been conducted on the model, and it has been found to transcend age, type of instruction, and subject matter (Hill & McGaw, 1981; Kottke & Schuster, 1990; Kunen, et al., 1981; Paul, 1985). Since its inception, the model has influenced curricular development, educational research, and the construction of tests (Kunen et al., 1981).

Bloom's Taxonomy outlines six categories of cognitive processes. The first category of

cognitive learning is *knowledge*, which includes:

> those behaviors and test situations which emphasize the remembering, either by recognition or recall, of ideas, materials, or phenomena. The behavior expected of a student in the recall situation is very similar to the behavior he was expected to have during the original learning situation. In learning situation the student is expected to store in his mind certain information, and the behavior expected later is the remembering of this information (Bloom, et al., 1987, p. 62).

This form of knowledge involves the recall and recognition of specifics (terminology and facts), of ways and means of dealing with specifics (conventions, trend and sequences, classification, categories, criteria, and methodology), and universals and abstractions (principles, generalizations, theories, and structures) (Bloom, et al., 1987).

The second category of Bloom's Taxonomy is *comprehension*, which represents the lowest level of understanding of information.

> Here we are using the term "comprehension" to include those objectives, behaviors, or responses which represent an understanding of the literal message contained in a communication. In reaching such understanding, the student may change the communication in his mind or in his overt responses to some parallel form more meaningful to him. There may also be responses which represents simple extensions beyond what is given in the communication itself (Bloom, et al., 1987, p. 89).

Comprehension involves three different ways of understanding: translation, interpretation, and extrapolation. Translation is comprehension that is paraphrased or rendered from one communication to another. Interpretation is the explanation or summarization of a communication involving reordering or rearrangement. Extrapolation is the extension of trends or tendencies beyond the given information to determine implications, effects, or predictions (Bloom, et al., 1987).

The third category of cognitive skills is *application*, which uses abstractions in particular and concrete situations. The abstractions involve general ideas, rules, methods, or principles and technical principles, ideas, and theories, which must be remembered and applied to new

situations. Bloom explains that the application of abstraction can be used when no mode of solution is specified.

> A problem in the comprehension category requires the students to know an abstraction well enough that he can correctly demonstrate its use when specifically asked to do so. "Application," however, requires s step beyond this. Given a problem new to the student, he will apply the appreciate abstraction without having to be prompted as to which abstraction is correct or without having to be shown how to use it in that situation (Bloom, et al., 1987, p. 120).

The fourth cognitive category is *analysis*, which breaks down information into its constituent elements or parts so that the relationship among the elements of an idea is made clear. Bloom describes that there are no clear lines between analysis and comprehension, or between analysis and evaluation.

> Analysis, as an objective, may be divided into three types or levels. At one level the student is expected to break down the materials into its constituent parts to identify or classify the elements of the communication. At the second level he is required to make explicit the s among elements to determine their connections and interactions. A third level involves recognition of organizational principles, the arrangement and structure, which hold together the communication as a whole (Bloom, et al., 1987, p. 145).

The fifth category found in the taxonomy is *synthesis*, which puts together elements and parts from different sources in order to form a whole and to produce unique patterns or structures. Synthesis skill requires the demonstration of ability to relate knowledge from several areas to create new or original work. The process of synthesis is defined as

> Working with elements, parts, etc., and combining them in such a way as to constitute a Pattern or structure not clearly there before. Generally this would involve a Recombination of parts of previous experience with new material, reconstructed into a new and more or less well-integrated whole (Bloom, et al., 1987, p. 162).

The sixth category identified in the cognitive domain is *evaluation*, which is the last stage in a complex process that involves some combination of all other behaviors (Bloom et al, 1987).

Bloom, however, explains that evaluation is not necessarily the last step that requires all the other categories of behavior in thinking and problem solving. Evaluation is defined as:

> the making of judgment about the value, for some purpose, of ideas, works, solutions, methods, materials, etc. It involves the use of criteria as well as standards for appraising the extent to which particular are accurate, effective, economical, or satisfying. The judgments may be either quantitative or qualitative, and the criteria may be either those determined by the student or those which are given to him (Bloom, et al., 1987, p. 185).

The behavior of evaluation is based on reasoned evidence involving both internal criteria (evidence as logical accuracy and consistency) and external criteria (references selected or remembered). Bloom explains that evaluation also deals with the major links of the affective behaviors, such as values, but represents the more cognitive behaviors.

In the last decades, although Bloom's Taxonomy has been practiced in teaching, its model has come under criticism. According to Furst (1994), one of the most common criticisms is that the taxonomy has oversimplified the nature of thinking and its relationship to learning. There has been evidence found in research that higher levels of the taxonomy did not seem to involve the more difficult cognitive processes than did lower levels of the taxonomy (Fairbrother, 1975; Poole, 1972). That is, there is no cumulative hierarchical structure (simple to complex behavior) in a philosophical perspective (Furst, 1994). It has been also found that there are no differences in the comprehension of instruction using the lower level or the higher level of the taxonomy (Barker & Hapkiewicz, 1979). Moreover, Furst (1994) argues that Evaluation should not be placed higher than Synthesis, but should be, at least, parallel with it, and that there is not a broad category of 'understanding' as one of the primary goals of learning that "is the tacit knowledge students acquire by transforming the details of formal instruction into interpretive schemata or categories" (p. 30). In addition, Huh in the study by Chung Bom Mo

(1994) argues that the Knowledge category corresponds with the ability to 'remember,' Application corresponds to 'simple application,' and Synthesis, to 'complex application.'

### 2.2.3. Revised Bloom's Taxonomy of Cognitive Domain

As education researchers have emphasized teaching higher levels of thinking, they have raised the awareness of a need for revision of Bloom's Taxonomy. Forty five years after the publication *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain* (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956), the revision of this framework was developed in the same ways by Anderson & Krathwohl, et al. (2001). In the revision, the authors divide the original categories into two dimensions: knowledge and cognitive process.

In the revised taxonomy (Anderson & Krathwohl, et al., 2001), knowledge is divided into from four categories: Factual Knowledge, Conceptual Knowledge, Procedural Knowledge, and Metacognitive Knowledge. Except for Factual Knowledge (two subcategories), each category involves three subcategories of knowledge from the original framework. Factual Knowledge requires the basic elements that the learner must acquire within a specific discipline (know what), involving knowledge or terminology and specific details. Conceptual Knowledge involves the interrelationships of how the basic components function or associate with the structure of entity, such as knowledge of classification, categories, generalizations, and theories. Procedural Knowledge demands that the learner know how to do something, including knowledge of specific skills or techniques that disciplines require and knowledge of algorithms. Metacognitive Knowledge refers to awareness and control of one's own thinking, such as self-knowledge, knowledge of contextual cognition, and strategic knowledge.

For the cognitive process dimension, Anderson & Krathwohl, et al. (2001) involve six categories: Remember, Understand, Apply, Analyze, Evaluate, and Create. Three categories were renamed and two were interchanged in switching the categories to verb form in order to include the uses of instructional objectives. The original category of Knowledge was renamed '*Remember*,' which is the ability to retrieve relevant knowledge from long-term memory, such as recognizing and recalling. Comprehension was renamed '*Understand*' that is the ability to determine the meaning of instructional messages, including oral, written, and graphic communication. The authors include the term 'understand' because it is a widespread synonym for the term comprehending. Application, Analysis, and Evaluation were retained as the verb forms '*Apply*,' '*Analyze*,' and '*Evaluate*.' Synthesis changed places with Evaluation and was renamed '*Create*.' Krathwohl (2002) argues that these cognitive categories are hierarchal in terms of their complexity, with *Remember* being more complex than *Understand*, which is less complex than *Apply*, and so on. He offers a two-dimensional table in which any instructional objectives, learning activities, and assessment can be represented: one dimension for knowledge and the other for cognitive process. Thus, the table can be used to classify objectives, activities, and assessments as well as to examine curriculum alignment.

### 2.2.4. Quellmalz's Taxonomy of Reasoning Skills

Quellmalz's approach reflects the contemporary influence of cognitive theory and provides a basis for assessing the constructive process of learning. The cognitive influence is reflected in the emphasis placed on the reconstruction of knowledge (Phye, 1997). Quellmalz (1985, 1991) conducted an analysis of the different ways of conceptualizing thinking and reasoning skills that have been proposed by psychologists, philosophers, and educators (Table 3).

Quellmalz found that philosophers' theories provide definitions of reasoning skills and criteria for the use of knowledge in light of grounds that justify it and its probable consequences. That is, reasoning skills require the learner to identify components of arguments, to judge the credibility of evidence, to use deductive and inductive thinking skills, and to make value judgments.

Table 3: Relationship Among Reasoning Skills Proposed by Psychologists and Philosophers (Quellmalz, 1991, p. 340)

| *Problem-Solving Strategies (Psychology)* | *Critical Thinking Skills (Philosophy)* | *Probable Dominant Cognitive Processes (Psychology)* |
|---|---|---|
| 1. Identify the problem (essential elements and terms) | 1. Clarification<br>• Identify or formulate a question<br>• Analyze major components<br>• Define important terms | 1. Analogical<br>• Analysis<br>• Comparison |
| 2. Identify appropriate information, content, and procedural schemata | 2. Judge credibility of support, the source, and observations | 2. Analogical<br>• Analysis<br>• Comparison<br>• Evaluate components |
| 3. Connect and use information to solve the problem | 3. Inference<br>• Deduction<br>• Induction<br>• Value judgment<br>• Fallacies | 3. Inferential—infer/interpret relationships among components |
| 4. Evaluate the success of the Solution | 4. Use criteria to judge the adequacy of solution | 4. Evaluate—evaluate the effectiveness of specific and general strategies |

Quellmalz (1985, 1991) also found that psychologists' theories of higher order thinking skills have been developed within the reasoning skills proposed by philosophers. Psychologists offered the definitions of higher order thinking skills based on the problem solving skills that

underlie cognitive processes, which require the learner to identify the type of problems, to identify and connect appropriate information, to judge the accuracy of the information, and to solve the problem. In addition, Quellmalz (1985, 1991) analyzed reasoning skills developed in the domain of curriculum that were identified by philosophers and psychologists. Curriculum theories offer to map the significant problem types and methods of inquiry within the subject matter, such as Bloom's Taxonomy, which has been the reference most frequently used by educators.

Quellmalz found that five common elements are derived from those many other cognitive structures: recall, analysis, comparison, inference, and evaluation. According to McMillan (1997), the reasoning skills developed by Quellmalz are easily applied to each different subject. The following is a summary of the five categories developed by Quellmalz (1985, 1991) and Stiggins, et al. (1988).

*Recall* is similar to knowledge and comprehension categories in Bloom's Taxonomy, and requires recognizing or remembering key facts, definitions, terms, and principles. Recall refers to the verbatim repetition, identification, or translation of knowledge. After rehearsing or mastering it, students can associate mastered subject content knowledge with other related concepts. Subsequent reasoning skills then arise, based on a mastered knowledge base. However, Quellmalz thinks that this level of thinking is different from the understanding of content knowledge.

*Analysis* is used in the same way in Bloom's taxonomy. It requires the student to divide a whole into component elements and to identify the elements, the relationships among the elements, and the relationships to its whole structure, such as the parts of cause/effect

relationships. With this skill, the student is able to break down, categorize, subdivide, or sort certain characteristics of objects or ideas, or the basic actions of procedures or events.

*Comparison* refers to the reasoning skill of similarities and differences. Simple comparisons require comparing and contrasting a few attributes or components, while complex comparisons require identification of the differences among many attributes or components. Some of the skills in the Bloom level of analysis are involved in this category. When students use distinct information processing, going beyond breaking the whole into parts in order to compare similarities and differences, these skills are involved in the Bloom level of synthesis.

*Inference* involves both deductive and inductive reasoning. In deductive tasks, the student is given a generalization and principles in order to construct a relevant conclusion. This category involves applications of rules that contain 'if, then' relationships. In inductive tasks, the student is required to relate and integrate evidence and details to form generalizations. This category involves hypothesizing, predicting, concluding, and synthesizing that require levels of application and synthesis.

*Evaluation* requires judging the quality and credibility of information according to established criteria, such as evidence, logic, or shared values. To evaluate, students should be able to explain the interrelationship of evidence and reasons in support of their conclusions. This task is similar to Bloom's levels of synthesis and evaluation categories.

### 2.2.5. Marzano's Core Thinking Processes and Standard-Based Learning

Marzano and his colleagues, Brandt, Hughes, Jones, Presseisen, Rankin, and Suhor (1988) have established a framework of thinking processes that involve multifaceted and complex thinking skills: concept formation, principle formation, comprehension, problem-

solving, decision making, research, composition, and oral discourse.

According to Marzano, et al. (1988), the first three processes—concept formation, principle formation, and comprehension—are used to acquire knowledge directly, while the next four processes—problem solving, decision making, research (scientific inquiry) composition—are used for knowledge production and application based on the first three processes. Oral discourse appears in the process of both knowledge acquisition and production in the form of communication.

Core Thinking Skills

The cognitive thinking processes as a complex operation defined by Marzano, et al. (1988) includes substantial thinking skills as an integral part. Marzano and his colleagues provide the definitions of core thinking skills, documented in various standards of psychological and philosophical research, and strategies, and comments on classroom applications of goal-oriented classroom activities. Marzano and his colleagues (1988) identify eight skills as core thinking: focusing, information gathering, remembering, organizing, analyzing, generating, integrating, and evaluating skills.

Marzano, et al. define *focusing skills* as involving the defining problems and the setting of goals, referring to the skills used to attend to selected information, while other information is ignored in order to sense a problem, issue, or a lack of meaning. Focusing skills can be used at the end of the process for problem solving, comprehending, or establishing any next steps. *Defining problems* refers "primarily to clarifying situations" early in the process of problem solving (Marzano et al., 1988, p. 70) while *setting goals* involve establishing direction or purpose in order to obtain outcomes as expected.

*Information gathering skills* refer to "the skills used to bring to consciousness the substance or content to be used for cognitive processing," which use the data already stored or newly collected (Marzano et al., 1988, p. 73), and involve observing and formulating questions. *Observing* is defined as the skill of obtaining new information from the environment, which involves complex operations such as classifying, hypothesizing, or inferring. *Formulating questions* involves the ability of attending to important information and generating new information, and clarifying issues and the meaning of inquiry (Marzano et al., 1988).

*Remembering skills* are activities that engage the learner in storing new information in long-term memory and retrieving it for use. Marzano and his colleagues (1988) argue that remembering skills are not activities associated only with rote memory, rather they are thinking activities involving *encoding* and *recalling*. They suggest that using keyword methods and activating prior knowledge can help the learner retrieve stored information.

*Organizing skills* are the activity "used to arrange information so it can be understood or presented more effectively" (Marzano et al., 1988, p. 80). These skills involve comparing, classifying, ordering, and representing. *Comparing* is the ability to identify similarities and differences between information collected. According to Feyerstein, in the study done by Marzano, et al. (1988), comparing involves complex cognitive operations such as precision, discrimination, and judgment of similarities and differences. *Classifying* is the activity categorizing items based on their attributes, and facilitating comprehension and retention of information. *Ordering* refers to the activity of sequencing things in a logical organization, for example, ordering a sequence of causes or effects in a historical event. *Representing* involves many forms such as visual, verbal, or symbolic, in order to show how ideas or objects are related.

*Analyzing skills* "clarify existing information by examining parts and relationships. Through analysis, we identify and distinguish components, attributes, claims, assumptions, or reasons" (Marzano et al., 1988, p. 91). Thus, these skills involve activities that identify attributes and components, relationships and patterns, main ideas, and errors. The authors argue that analysis functions as the core of critical thinking as defined by philosophers, looking at inside of ideas or objects.

*Generating skills* refer to the activity of constructing a new structure based on prior knowledge and new information or ideas. Marzano, et al. (1988) explain, "Generating is essentially constructive, as connections among new ideas and prior knowledge are made by building a coherent organization of ideas (i.e., a schema) that holds the new and old information together" (p. 98). Generating involves *inferring, predicting, and elaborating* new ideas that come into play in order to recast new structures.

*Integrating skills* use prior knowledge in combination with relevant ideas or aspects of a solution or principles in order to build new understanding. Marzano, et al. (1988) illustrate the function of integrating skills as the following:

> New information and prior knowledge are connected and combined as the learner searches for prior knowledge related to incoming information, transfers that knowledge to working memory, builds meaningful connections between incoming information and prior knowledge, and incorporates this integrated information into a new understanding (p. 104).

Integrating skills involve summarizing and restructuring. *Summarizing* is the activity of combining information efficiently into a coherent form, such as oral or written, including important elements and their relationships with supporting details. *Restructuring* refers to the activity of recasting past ideas by modifying, extending, reorganizing, or discarding old understandings based on the understanding that past beliefs or concepts are no longer valid.

*Evaluating* refers to detecting inconsistencies in or fallacies of ideas, including establishing criteria and verifying skills. *Establishing criteria* in a philosophical perspective refers to "setting standards for judging the value or logic of ideas," while in a psychological perspective it focuses on "effectiveness of particular learning strategies or achievement of learning goals" (Marzano, et al., 1988, p. 110). *Verifying* focuses on confirming or disconfirming the accuracy of ideas based on understanding that the nature of the evidence must be proved in order to claim it as true.

Standardized-Based Learning

Thinking processes have also been developed in standardized-based learning. Marzano and his colleagues (Marzano, 1992; Marzano, Pickering, & McTighe, 1993) have developed an instructional framework for organizing learning outcomes into two dimensions: content standards and life-long learning standards. This framework reflects the change in learning targets to a broader array of standards-based on learning, which provides an educational environment ensuring all students acquire expected specific knowledge and skills.

*Content standards* are divided into two types of knowledge: declarative knowledge and procedural knowledge. Declarative knowledge reflects knowledge skills about particular information and facts at the bottom level and concepts and generalizations at the top level. Concepts and generalizations help students develop a broad knowledge base. Declarative knowledge involves three phases of acquiring and integrating knowledge: constructing meaning, organizing, and storing. Students construct meaning by associating new information with prior knowledge (schema). They then organize the information in order to make salient important information and its relationships. After that, they store the information in long-term memory, linking new information with the old (Marzano, 1992; Marzano et al., 1993).

Procedural knowledge reflects strategies or skills that apply to a variety of situations, such as general strategies or skills of analyzing, interpreting, summarizing, or transforming information. When acquiring and integrating procedural knowledge, students must initially build a detailed model of the process involved, using algorithms, tactics (referred to as heuristics), or strategies. They then must shape the process by establishing a conceptual understanding, eliminating errors, and identifying the most efficient techniques for completing the process. Finally, students must practice the process in order to internalize it, so that they can perform it again at a later time (Marzano, 1992; Marzano et al., 1993).

Acquiring and integrating knowledge is a fundamental goal of schooling for learners. However, the purpose of learning is not a simple matter of obtaining content knowledge. Although facts are important, they are often meaningless in isolation. Marzano (1992) illustrates, "the most effective learning occurs when we continually cycle through information, challenging it, refining it" (p. 67).

*Life-long learning standards* involve complex thinking, information processing, effective communication, cooperation/collaboration, and standards of effective habits of mind. These standards include reasoning skills that extend and refine knowledge, and using them in meaningful ways in virtually all aspects of life.

Complex thinking standards involve reasoning processes, such as comparing, classifying, induction, deduction, error analysis, constructing support, abstracting, analyzing perspectives, decision making, definitional, historical, and projective investigation, problem solving, experimental inquiry, and invention. Comparison requires an analytic task that identifies and determines the similarities and differences of characteristics between things. Classifying refers to sorting information into definable categories on the bases of their characteristics. Induction is

making unknown generalizations and logical statements based on observation or analysis of various cases. Deduction refers to making unstated consequences and conditions from given generalizations and principles. Error analysis is identifying and articulating erroneous conclusions in reasoning. Marzano (1992) states, "One of the most common types of errors made every day falls under the category of confirmatory bias, which is the tendency to seek out information that confirms our hypotheses" (p. 87). Ways in which errors in reasoning are based on faulty logic, attack of a person or position, and weak references. Constructing support means creating sound persuasive argument based on evidence, elaboration, and qualifiers. Abstracting is identifying and drawing underlying general patterns or themes from information or situations. Analyzing perspectives involves "identifying your position or stance on an issue and the reasoning behind that stance and considering a perspective different from your own" (p. 98).

All these reasoning skills can be used for specific targets in meaningful ways, such as decision making, investigation, experimental inquiry, problem solving, and invention. Among these reasoning skills, decision making and problem solving strategies are frequently used in the classroom (McMillan, 1997). Problem solving tasks involve developing and testing a method or product for overcoming obstacles or constraints in order to achieve a desired outcome. The processes of problem solving involve specifying a goal, identifying the constraints, identifying alternative ways of accomplishing the goal, and selecting an alternative and trying it out. Decision making is similar to problem solving, but it may or may not involve obstacles and constraints. According to Marzano, et al. (1988), borrowing from the model from a study conducted by Wales, Nardi, and Stager, there are four steps involved in decision making: stating the goal, generating ideas, preparing a plan, and taking action. Each step involves three processes: identifying problems (analysis), creating options (synthesis), and selecting the goal

(evaluation). In decision making, students need to understand the desired goal or result, evaluate the alternatives in terms of criteria related to the situation, and either select a plan, task, course of action or make a choice on the basis of their evaluations. Investigation examines and systematically inquires about something. For example, historical investigation involves identifying why or how a past event occurred, in an attempt to understand the past. Experimental inquiry tasks refer to testing hypotheses that have been generated in order to explain a phenomenon, engaging in such questions as "How can I explain this?" and "Based on my explanation, what can I predict?" (Marzano, 1992, p. 116). Finally, invention tasks are for developing something unique or making unique improvements in a product or process (Marzano et al., 1988; Marzano, 1992; Marzano et al, 1993).

Information processing standards involve the use of various information gathering techniques and sources; the interpretation, synthesis, and evaluation of information; and the assessment of valuable and relevant information (Marzano, 1992; Marzano, Pickering, & McTighe, 1993; Marzano, 2001). Effective communication standards reflect effective communication with diverse audiences, in a variety of ways, and for a variety of purposes (Durst and Newell, 1989). Collaboration and cooperation standards work toward the achievement of group goals, use interpersonal skills effectively, contribute to group maintenance, and perform a variety of roles. After the studies conducted by Johnson & Johnson (1987) and Slavin (1983), collaboration and cooperation in education have received attention. Ennis (1987), Paul (1990), Costa (1991), Zimmerman (1990) found that human beings have the ability to control by using effective habits of mind. Habits of mind standards include self-regulation (awareness of one's own thoughts and of the uses of necessary resources, and making effective plans), critical thinking (seeking the accuracy, clearness, and clarity of information), and creative thinking

(generating new ways of viewing situations, generating trust, and maintaining one's own standards of evaluation) (Marzano et al., 1993).

Table 4: Comparison of Reasoning Frameworks

| | Definition | Major Components |
|---|---|---|
| Bloom's Taxonomy | Higher Order Thinking Skills | Application (apply in novel situations, predict effects)<br>Analysis (distinguish and check consistency)<br>Synthesis (combine information)<br>Evaluation (logical inconsistencies, fallacies, adequacy of evidence, judgment of quality or value of something) |
| Revised Bloom's Taxonomy | Higher Order Thinking Skills | Apply (executing, implementing)<br>Analyze (differentiating, organizing, attributing)<br>Evaluate (checking, critiquing)<br>Create (generating, planning, producing) |
| Quellmalz's Taxonomy | Higher Order Reasoning Skills | Analysis (identify components)<br>Comparison (contrast, relate; similarities, differences)<br>Inference (deductive/inductive thinking)<br>Evaluation (judgment) |
| Marzano's Core Thinking | Core Thinking Skills | Focusing (defining problems, setting goals)<br>Information gathering (observing, formulating questions)<br>Remembering (encoding, recalling)<br>Organizing (comparing, classifying, ordering, representing)<br>Analyzing (identifying attributes, components, relationships, patterns, main ideas, errors)<br>Generating (inferring, predicting, elaborating)<br>Integrating (summarizing, restructuring)<br>Evaluating (establishing criteria, verifying) |
| Marzano's Dimensions of Learning | Complex Thinking or Reasoning Skills | Extending and refining knowledge (comparing, classifying, inducting, deducting, error analysis, constructing support, abstracting, analyzing perspectives)<br>Using knowledge meaningfully (decision making, problem solving, experimental inquiry, and invention) |

Table 4 illustrates a comparison of reasoning and higher order thinking skills as defined by Bloom, Quellmalz, and Marzano and his colleagues. According to this comparison, several similarities appear among the reasoning and higher order thinking skills. *Analysis* in Bloom's Taxonomy is similar to: *analysis* in the revised Bloom's Taxonomy; *analysis* and *comparison* in Quellmalz's Taxonomy; *organizing* and *analyzing skills* in Marzano's core thinking skills; and *comparing, classifying, error analysis*, and *analyzing perspectives* in Marzano's complex thinking. *Synthesis* of Bloom's taxonomy is similar to: *create* [creating] of revised Bloom's Taxonomy; *inference* in Quellmalz's Taxonomy; *generating* and *integrating* in Marzano's core thinking skills; *induction* and *deduction* in Marzano's complex thinking. *Evaluation* in Bloom's Taxonomy is similar to: *evaluation* in Bloom's Taxonomy; *evaluation* in Quellmalz's Taxonomy; *evaluating* in Marzano's core thinking skills. These thinking skills can be used separately or for tasks such as *decision making*, *problem solving*, *experimental inquiry*, and *invention* as Marzano's using knowledge meaningfully.

## 2.3. Mapping Historical Reasoning for Historical Understanding

### 2.3.1. Introduction

Historical thinking is a cognitive operational ability that utilizes complex mental activities in order to achieve historical understanding. Learning history can be justified only if what is termed 'history' offers its recipient an introduction to historical thinking because the basic epistemological nature and structure of history is very different from that of mathematics or science. In general, historical knowledge itself as well as the development of discourse-based reasoning and explanation based on historical evidence are always tentative, ambiguous, and uncertain. Although there are an increasing number of ongoing issues concerning the premises

under the historians' work, it may be possible to discern several elements in historical thinking, defining the general characteristics of history. Also, in 1994, the National Center in the Schools at the University of California at Los Angeles recommended that schools follow the National Standards for United States and World History. According to the standards for history, the goals of history education focus on two main types: historical thinking skills (knowing how) and historical understanding (knowing what) (National Center, p. 2), and that the attainment of higher levels of historical understanding is derived from higher levels of historical thinking. Based on the standards for history and on substantial studies conducted by historical researchers, this section attempts to define the important elements of historical thinking skills derived from the nature or structure of history, and consider their implications in the school instruction.

### 2.3.2. Chronological Thinking

What separates history from every other kind of inquiry about human affairs is its fundamental concern with the concept of time. History views all human activity in the setting of particular times and sequences of event, from one time to the next. Therefore, history, more than other social science, is concerned with questions of change, repetition, continuity, development, or progress in society's affairs in a time frame (Daniels, 1972; Lomas, 1993). The time perspective accentuates history's concern with studying the unique characteristics of particular situations and events seen in their specific time settings and evaluated accordingly (Daniels, 1972; Stow & Haydn, 2000).

The National Center (1994) defines chronological thinking as being:

> at the heart of historical reasoning. Without a strong sense of chronology—of when events occurred and in what temporal order—it is impossible for students to examine relationships among those events or to explain historical causality. Chronology provides the mental scaffolding for organizing historical thought (p. 20).

In the study of history, the concept of chronology has a central place in the development of a child's historical understanding. The term chronology is derived from two Greek words: "*chronos*, meaning time, and *logia*, meaning a branch of knowledge" (in Stow & Haydn, 2000, p 87). Chronology is defined as "the arrangement of dates or events in order of occurrence; the determining of the proper sequence of past events" (Collins, 1995), or "the science of computing time or periods of time and assigning events to their time dates" (Oxford, 1993). Thornton and Vukelich (1988) consider that there are three main aspects in understanding time: clock time, calendar time, and historical time. Differentiating the first two, the authors describe historical time as "[requiring] one to depict a person, place, artifact, or event in the past, using some form of time language" (p. 70).

Oaken and Sturt, in Stow and Haydn (2000), saw three standards in historical time:

- A child's understanding of time—words and symbols such as are in use in everyday life;
- His power to form the conception of a universal time scheme… and his ability to use the dates by which such a scheme is symbolized;
- His knowledge of the characteristics of definite epochs in the time scheme, and his ability to place these epochs roughly in the correct order (p. 87)

Therefore, an understanding of chronology is the ability to match events and personalities to dates and historical periods and place them in order of occurrence (Hoodless, 1996). Chronological thinking is a fundamental thinking skill in learning history, providing a mental framework or map which gives significance and coherence to pupils studying history. Achieving such a level of understanding, however, requires a lengthy developmental process involving the elaboration, restructuring, and synthesis of children's knowledge (Masterman & Rogers, 2002). Wood (1995) argues that the ability to sequence is a basic feature of historical understanding, and that 'the past is chaos' to pupils, until sequenced (p. 11).

Chronological thinking is basically the ability to sequence events and objects and to use the appropriate everyday vocabulary to describe time in the past. It requires students to use "mathematical skills to measure time by years, decades, centuries, and millennia, to distinguish between past, present, and future time, and to interpret the date presented in time lines" (National Center, p. 20). Chronological thinking is also the ability to analyze patterns of historical duration, to make sense of the temporal structure of events unfolding over time, and to link between the various chronologies, such as the temporal connections between antecedents and their consequences (National Center, 1994; Holt, 1990; Stow & Haydn, 2000). In addition, chronological thinking involves knowing characteristic features of particular periods and societies, including the range of ideas, beliefs, and attitudes of people in a time frame.

Clearly the development of an ability to 'arrange dates or events,' 'determine a proper sequence of events,' or 'compute time or periods of time' touches on the development of an understanding of events of the past (Stow & Haydn, 2000, p. 87), or to what those periods refer and what their characteristic features are. Thus, chronological thinking in the classroom can legitimately be used to achieve a broader sense of understanding of historical time.

Friedman (1982) suggested that by the age of six, children are aware of regularities in time, and use the names of the more everyday elements of time. In their study, Thornton and Vukelich (1988) argue that from the ages of nine to eleven, children start to use period labels. Moreover, Stow indicated that six-year-olds were able to use period labels, while the majority of eight- to nine-year-olds were capable of using these with some confidence and understanding (in Stow & Haydn, 2000, p. 88). All these studies suggest that the cultural and educational contexts of time influence the pace at which a child develops an understanding of the language of time.

Using pictures to help make sense of time capitalizes on children's awareness of stereotypical images of time periods. According to Lynn (1993) and Harnett (1993), children at a young age are able to codify the relative periods to which the pictures belong. Andretti (1993) emphasizes the value of artifacts as a multi-sensory stimulus that helps develop children's sense of chronology and sequence. Buildings and the use of an historic environment are other essential aspects that enhance children's concept of a time period, such as architectural styles, building materials and shapes, styles of windows and doors, and even types of buildings.

The use of a timeline is often recommended for primary school children (Friedman, 1982; Hoodless, 1996). Researchers believe that timelines have become a strong aspect of preferred practice in understanding chronological thinking. Along with children's mathematical development, it is recommended that the numerical complexity of the calibrations also increase with children's age (Friedman, 1982; Hoodless, 1996). In addition, more complex timelines can also be used in presenting concepts of secondary history in order to ensure continuity and progression of history.

When teachers are confronted with the task of teaching chronological thinking, simply teaching the best story in the way the event unfolded is the basic strategy used to engage students in an understanding of the past (History Center, 1994; Perfetti, et al., 1994 & 1995; Shemilt, 2000). To learn history is to learn a story and to come to know the major characters, events, and simple causal relation between events (Perfetti et al., 1994). While building a progression of narrative frameworks, Shemilt (2000) attempted to develop pupils' understanding of history as a logical and evidence-based means of making sense of the past, that is "a chronologically ordered past" as the first level. According to Shemilt, it is necessary for pupils to acquire a basic

chronology that embraces the whole of the past and is represented in terms of significant phases of human history, such as the benefits of telling a simple story.

### 2.3.3. Historical Comprehension

The second category of historical thinking found in the National Standards for History is historical comprehension (National Center, 1994, p. 23). According to the standards for history, historical comprehension can be accomplished by learning narratives, stories, or biographies involving the major characters, events, and the causal relations among events. For historical comprehension, National Center (1994) offers the important role of historical narrative in understanding history:

> One of the defining features of historical narratives is their believable recounting of human events. Beyond that, historical narratives also have the power to discourse the intentions of the people involved, the difficulties they encountered, and the complex world in which such historical figures actually lived (p. 23).

The standards emphasize that in order to read historical accounts with comprehension, students must be able to read them imaginatively, to understand what the accounts uncover of "the humanity of the individuals and groups involved—their motives and intentions, their values and ideas…" and "avoid 'present-mindedness'" judging the past with contemporary values (p. 23).

In the Random House Dictionary (1983), history is defined as "a story, narrative, and a continuous, systematic narrative of past events as relating to a particular people, country, period, person, etc., usually written in chronological order" (p. 352). In a study conducted by Rosa (1994), the original Greek word "history [means] learning or knowing by inquiry, [or] narrative" (p. 225). In general, historical accounts represented in a narrative form involves a story with a beginning, middle, and end, a setting, characters, problems, and resolution (Levstik & Barton, 2001). Basically, knowing an historical story may be the starting point of understanding history.

Several researchers think that teaching the best story as the way in which the past event actually took place is the first strategy in engaging students in historical understanding, such as chronological understanding, defining causal relationships between events, or ethical issues surrounding the event (Levstik & Barton, 2001; Perfitti, et al., 1994 & 1995; Seixas, 2000; Shemilt, 2000). Even if historical experts differentiate a story from their professional works involving analytic investigations of the past, their explanations of history involve the basic story (Perfitti, et al., 1994 & 1995).

Among historians and philosophers of history, there has been an ongoing debate about the nature of historical narrative. In fact, historical accounts are represented in a variety of forms, such as narrative stories, chronicles, imaginative reconstructions, and formal analytical essays (Husband, 1996). Analytical essays resemble more closely to the professional work of historians, conveying the truth of the past, while narrative stories tend to illustrate the "lifelikeness" of the past (Bruner in Husband, 1996, p. 44; Lochman & French, 1978). For Danto (1965), historical narratives rely on ideal descriptions based on chronicles, so that historical narratives tell everything in objective ways, excluding the particular perspective interest taken by the historian. On the other hand, Ricoeur (1984) regards historical narrative as a kind of fiction, that is, there are common principles that can apply the way in which novelists write to the way in which historians reconstruct the past. Linking scientific and fictional history, White (1984) argues that all of historical writings follow one of the four basic types of plot: romance, tragedy, comedy, or satire. According to the author, the historian chooses one form or another among the four for exploring the representation of the past, then the account can be assessed ideologically, such as taking a perspective. Therefore, for White, historical reconstruction is situated between emplotment and ideology.

Lichtman and French (1978) differentiate historical narratives from fiction, accepting the idea that history involves a form of narrative. The authors suggest that, in order to reconstruct a story or narrative, a historian is constrained by the logical and empirical requirements of historical inquiry using evidence. Moreover, the authors argue that an historian may choose to explain history in a narrative form, applying the same principles a writer uses in constructing a story. That is, an historian may alter the actual flow of time, creating a tempo best suited to the narrative or to highlight a dramatic turning point in the event described. However, Lichtman and French report this difference between fiction and history:

> If the poet abandons the requirements of the sonnet, he still may write great poetry, but he will not have written a sonnet. If the historian abandons the constraints of his discipline, still may write great literature, but he will not write history. (p. 214)

As a result of those arguments, history is shaped by stored accounts and historical analysis. Historical accounts are more than chronicles (a list of events) and fictions. Historical narratives are represented by sorting events and organizing them around ideas in a way that is coherent and complex with rich connections among the events, containing multiple roles and full of multilayered interpretations.

For young children, National Standards for History (National Center, 1994) suggest that students should be given stories, narratives, biographies, autobiographies, and historical documents in order to comprehend the events, the lives and ideas of historical figures, and the historical motives. No historical narrative can be entirely objective, that is, historical accounts always involve interpretations by historians because the past is already dead and cannot be observed directly and repeated. Historians must use primary sources or artifacts in order to explain what happened and why. Thus, for historical comprehension regarding the nature of history, students must master the skills presented in National Standards for History (National

Center, 1994): "1) identifying the central question the historical narrative seeks to answer; 2) defining the purpose, perspective, or point of view from which the narrative has been constructed; 3) reading the historical explanation or analysis with meaning; and 4) recognizing the rhetorical cues that signal how the author has organized the text" (p. 23). In addition, students must be able to differentiate historical fact from interpretation, but also recognize the importance of their relationship (National Center, 1994).

As the standards emphasize, students are required to use the ability to read historical accounts imaginatively, understanding ideas of historical agents or motivations of the event in the context of the past. Collingwood (1972) outlines an idealistic philosophy of history based on the 'imaginative reconstruction' of past patterns of thought and perceptions, which can be a basic skill for historians. He defined historical knowledge as "an activity of thought… knowing mind re-enacts it and knows itself to be doing so. To historian, the activities whose history he is studying are not … but experiences to be lived through his own mind…" (p. 218). Imaginary reconstruction is grounded in historians' reenactment of the past based on their thoughts (hindsight), which go beyond historical fact, employing historical evidence in order to create the purposes that existed in the past.

By using Collingwood's term, Portal (1987) and Shemilt (1984) seek to link the 'rethinking ideas' to the process of 'historical empathy.' Thus, Portal and Shemilt believe that historians' intellectual operations, such as inference from both the historical evidence and their own backgrounds of knowledge, play a part in historical reconstruction. In other words, empathy as re-enacting or rethinking history plays an important role in bridging the gap between historical reconstruction and fact (Shemilt, 1984; Ashby & Lee, 1987; Portal, 1987; Foster, 1999; Yeager & Foster, 2001). The authors believe that in order to project themselves imaginatively

into the historical situations, students must learn to use their mind's eye, bringing into play historical perspective and intuitive observation because the past is an incomplete entity and contains limited views. In order to read historical narratives imaginatively, using historical empathy, Seixas (1996) is cautious of applying presentism—imposing students' own frameworks of meaning on others in the past.

As a school subject, historical narratives can play a role in shaping the collective memory of students as a member of society (Lerner, 1997; Fulbrook, 1999; Seixas, 2000) sharing collective historical experiences and common myths, traditions, and historical consciousness with their community. Also, historical narratives in school education can convey group identity for social cohesion and offer a moral framework for future action, such as human rights, race struggles, or gender equality (Mathien, 1991; Seixas, 2000). In order to accomplish these kinds of particular political uses for school history in general, history taught in the school tends to rely on the transmission of historical stories, thus limiting the offer of a different side of the story (Rosa, 1994). Therefore, the manner in which historical interpretations should be taught and which is the right version of the past should be thoroughly considered for history education, teaching history with historical accounts.

### 2.3.4. Historical Analysis and Interpretation

Just learning the past through historical stories and narratives does not engage students in the historical disciplines' mode of inquiry (Bain, 2000; Perfitti, et al., 1994, 1995; Seixas, 2000; Voss & Wiley, 2000) because the meaningful learning of history entails going beyond simple stories to interpret, reconstruct the literal meaning of a historical passage, and generally negotiate the uncertainties surrounding the event (Perfetti, et al., 1994; Rodrigo, 1994; Leinhardt, et al.,

1994a, 2000). A higher standard of learning applies to historical literacy, requiring the analysis of historical texts and interpretations. This is closely related to Seixas's 'doing discipline' of history that establishes a basic historical idea by practicing historical methods (Seixas, 2000). It implies the awareness that the construction of a story comes from records of various kinds, that history is grounded in historians' interpretation through evidence, and that there is a distinction between primary and secondary sources (Carr, 1961; Wineburg, 1991a; Perfetti et al., 1994; Seixas, 1999).

In National Standards for History, the ability for historical analysis and interpretation requires that students must recognize that:

> Historians may differ on the facts they incorporate in the development of their narratives and disagree as well as on how those facts are to be interpreted… Thus, "history" is … but written history is a dialogue among historians, not only about what happened but about why and how events unfolded (National Center, 1994, p. 26).

This idea is intrinsically related to the interpretive nature of historical knowledge based on survived evidence. History consciously or unconsciously reflects the thoughts and perceptions that the historian himself may take from the society in which he lives. As any working historian knows, he is engaged in a continuous process of molding his facts to fit his interpretation and molding his interpretation to fit his facts (Carr, 1961; von Wright, 1971). Such views allow that historical facts might be constructed in an infinite number of ways, depending upon which web of interpretations the historian favors.

Almost half a century ago, E. H. Carr (1961) explored the question, "What is history?" as "a continuous process' of interaction between the historian and his facts, 'an unending dialogue' between the present and the past" (p. 24). It implies that the interaction between past and present takes place through the process of a dialogue across time, one in which a historian carries on

with the facts of the past discovered through historical sources. Historical fact itself cannot explain the truth of the past regarding what and why something happened. As E. H. Carr (1961) states, "The necessity to establish these basic facts rest not on any quality in the facts themselves, but on an a priori decision of the historian…" (p. 6), a historical fact is relative to the purpose of the historical discourse and attentive to the current needs for the representation of the past (Rosa, 1994).

Historical accounts fundamentally involve the reconstruction of a past, thus, there exists a heavy subjectivity of the interpretation of historical events and negotiation of the tension between the past and the present. Historians are faced with having to decide which events, people, or dates among the chronicles are to be included or left out, when the story begins and ends, and which problems are resolved: A particular event, person, or date is emphasized as a significant event, person, or date while others are excluded or deemphasized (Boix-Mansilla, 2000; Rogers, 1987; Seixas, 1996). Therefore, there is the possibility of multiplying historical facts on different occasions by the number of interpretations offered by historians and of new interpretations that are challenged to uncover new voices from the past.

In a substance study of historical reasoning conducted by Kuhn, et al. (1994), the task of a juror or an historian are similar in that each must reconstruct a past event and make decisions about the role of the human action or intention in that event. Because of the incomplete, uncertain, and inconsistent nature of evidence about a past event, both the juror and the historian must examine the evidence, defining the credibility of it, its source, and the relationships among the evidence. These ideas, that history refers to the epistemic practice of producing scientific historical narratives that address human affairs, are related to the basis for analysis in order to enhance historical thinking, such as hypothesizing the influence of the past, understanding the

event in the broader context, and comparing unsupported opinions from interpretations grounded in historical evidence (National Center, 1994).

An example of what it means to expect students to have a sense of historical evidence comes from Wineburg (1991a, b), who conducted empirical studies comparing the way historians think about primary and secondary sources with the thinking processes of high school students and teachers. In his study of expert-novice approaches to reading documents, Wineburg identified three heuristics used by experts in reasoning from historical evidence that could be taught to students. Experts notice and evaluate the source of the document (sourcing), check the facts mentioned in the document against those in other documents (corroboration), and set events into a larger context (contextualization). These multiple strategies help construct a more complex and, ultimately, richer understanding of historical thinking and history (Wineburg, 1991a, b).

In their study, Perfetti, et al. (1994, 1995) also suggest methods: evaluating uncertainties by using interpretative skills involved in reading multiple historical texts; detecting the author; handling inconsistencies among texts; evaluating the incompleteness of texts; and resolving conflicting views. Using multiple texts can facilitate historical thinking as well as using strategies historians employ, for example, by encouraging the comparison of contents across text, creating awareness of the importance of source information, and recognizing the inconsistencies and biases that exist within a text (Perfetti et al., 1994, 1995, 2000). According to Perfetti, et al., a common feature of these techniques is that students are exposed to multiple texts presenting several facets of the same topic, provoking the reader to be contradictory. Studies conducted by Perfetti, et al. were based on the causal-template structure model that engages students in understanding the event in which the United States acquired the Panama Canal. The authors

found that students understood more the core event, increasing the number of core events reported after each reading while non-core events were not significantly reported. That is, students learned more from the causal structure of the historical texts.

In addition, another model for historical analysis is suggested in the study conducted by Stearn (2000). For his college students, Stearn used analytic processes of world history: comparing the differences and similarities of gender issues between China and India; comparing the changes in the political, social, cultural, and economic features of several countries over time; comparing two societies and their changes; finding causation of an event and interpreting it. These activities seemed to utilize cognitive improvement in relationship to the comparison across different regions or eras, seeking their changes and the decline or improvement that transcends regional and temporal boundaries.

Considering these philosophical issues surrounding the nature of history and the theories of acquiring historical understanding, the historical thinking for historical analysis and interpretation suggested by National Standards for History (National Center, 1994) are:

> Students should be able to:
> Identify the author or source of the historical document or narrative and assess its credibility; compare and contrast differing sets of ideas, values, personalities, behaviors and institutions; differentiate between historical facts from interpretations; consider multiple perspectives; analyze cause-and-effect relationships; challenge arguments of historical inevitability; compare competing historical narratives; hold interpretations of history as tentative; evaluate major debates among historians; hypothesize the influence of the past (p. 27).

### 2.3.5. Historical Research Capability

One of the most important aspects in improving students' historical thinking is to invite them to participate in the process of authentic historical inquiry that historians use. Learning history refers to the epistemic practice of producing historical narratives (Blanco & Rosa, 1997),

resulting in hypothesizing about history, testing it, and interpreting its response to a specific question about the past (Husband, 1996; Carretero & Voss, 1994). Rodrigo (1994) argues that in order use the research skills used by historians themselves, although a full command of skills is not required, students should have achieved a certain level of historical literacy presented in the historical thinking skills of interpretations. The National Center (1994) discusses the importance of developing historical inquiry skills:

> Perhaps no aspect of historical thinking is as exciting to students or as productive of their growth in historical thinking as "doing history." Such inquires can arise at critical turning points in the historical narrative presented in the text…. Worthy inquires are especially likely to develop if the documents students encounter are rich with voices of people caught up in the event and sufficiently diverse to bring alive to students the interests, beliefs, and concerns of people with differing backgrounds and opposing viewpoints on the event (p. 29).

Regarding improving historical research capabilities, the intensive study conducted by Leinhardt, et al. (1994a, b) is based on interviewing practicing historians. In the study, five clusters of ideas related to reasoning processes in history were defined: a motivational/purposive assumption, a compelling narrative, an evidential exhaustivity, a central hypothesis, and contextual interpretation. According to the study, Leinhardt's historians assumed that doing and constructing history involve purposive motivations that unfold the reality of the past itself or offer a source of liberation. The author states that the underlying assumption of doing history is to help "understand both what was and also what is" (1994b, p. 141). The historians interviewed agreed that history is constructed of a narrative form as a main requirement with internal coherence, chronology, and causality. Chronology and causality act as a framework for the story, including all evidence as either supporting or contradicting in order to build a coherent narrative. Moreover, historians develop hypotheses and support them with relevant evidence from their own cases in order to construct interpretations. The reconstructed interpretations

based on ethical issues are illustrated either in the original context or in the context of contemporary situations, defining the power relationships.

Informal writing or thinking on paper may allow students to explore the connections and speculations about historical phenomena, and develop the skills of the historian (Bain, 2000; Husband, 1996), specifically, problem-based essays that interpreted the task as historians do, could be invited to incorporate their own ideas with information from the sources (Greene, 1994, Leinhardt, 2000; Voss & Wiley, 1997, 2000). Greene (1994) found a similarity between students and historians in writing reports and solving problems. Students showed the primary difference in using a wide range of evidence in order to incorporate their knowledge in problem-based writing into report writing. Although students had difficulty in incorporating their knowledge in using specific evidence and in building arguments in the discipline of history, there was not much difference between historians and students in problem-based task. Both groups recognized that scholarly writing involves the task of weaving source information with the writer's knowledge in order to support or contrast their view of specific cases.

A case study, Lessons on Teaching and Learning in History from Paul's Pen, by Leinhardt (2000) shows how one student in an AP history class developed his own historical concepts, argument, and reasoning skills through writing historical essays based on multiple documents, and through communicating and discussing in class. This study shows one possibility where students could incorporate multiple perspectives, develop their standpoints, and engage in a type of imitation through argument in written essays or in a class as well as through consistent teacher feedback. The multiple-segment essay and argument-writing task tend to yield a higher proportion of transformed sentences (Leinhardt, 2000; Voss & Wiley, 1997 & 2000). In relation to the inference verification task that follows from using the text as evidence,

reading from multiple segments and writing argumentative essays can be a good way for students to recognize inference skills (Voss & Wiley, 1997, 2000).

The activities of student writing involve the same work the historian does, therefore, the work that the historian does can be applied to the same process that students use for their research projects in considering historical inquiry, such as questioning, collecting, processing, analyzing, and synthesizing historical information that is relevant to the event investigated (Husband, 1996). The activities proposed by the National Center (1994) for improving historical inquiry skills are presented as: "[formulating] historical questions from encounters, [obtaining] historical data from a variety of source, [interrogating] historical data by uncovering the social, political, and economic context, [identifying] the gaps in the available records, and [marshalling] contextual knowledge and perspectives of the time and place" (p. 30).

### 2.3.6. Historical Issues-Analysis and Decision-Making

In the standards for history (National Center, 1994), finally, historical thinking can be constructed by issue-centered analysis and decision making activities that lead students to historical dilemmas and problems in the past and the near present. The final goal of learning history can be to promote students' capabilities to be democratic and active citizens by involving them in the process of problem solving and developing their own moral judgment based on past historical events or actions. Although the past is already dead, it is alive for us as other values or meanings in contemporary society. As Thompson argues, "For we are saying that these values and not… other values, are the ones which make this history meaningful to us, and that these are the values which we intend to enlarge and sustain in our own present" (in Husband, 1996, p. 66). The activities of issue-centered analysis and decision making can promote students' ethical

judgment or personal involvement by "confronting the issues or problems of the time, analyzing the alternatives available to those on the scene, evaluating the consequences that might have followed those options for action that were not chosen, and comparing with the consequences of those that were adopted" (National Center, 1994, p. 32).

Many historians regard specific facts from the past as the instrument for understanding the present (Pratt, 1974; Smith, 1991; Carretero & Voss, 1994; Blanco & Rosa, 1997; Leinhardt, 2000; Lerner, 1997). This may be the most valid justification for the study of history. In fact, history's locus of historical inquiry is the complex relationship between the past and the present (Blanco & Rosa, 1997) because it is likely that something can be learned from how other societies in the past dealt with their problems. The past that is similar to the present as a frame of reference becomes a useful resource in our everyday lives in solving problems and issues. Situating ourselves in the stream of time may be a basic human need that helps shape and interpret the present (Boix-Mansilla, 2000; Wineburg, 2000), thus, history helps make sense of our own lives within the current social context.

However, as Wineburg (2001) and Foster (1987) argue, viewing the past as usable may encourage us to instantly consume history. We may discard or ignore the past when it contradicts our current needs or when it fails to align with them. Instead of stretching our understanding to learn from the past, we may contort the past to fit the predetermined meanings we have already assigned it. Moreover, if history is, for practical uses, to meet the needs of present, there would be a heavy selection of historical facts from among the mass of crude facts, thus, the past only to be interpreted for the present rather than being properly historicized and containing its own meanings.

The issues and problems defined in the past can be a compelling moral framework for contemporary society (Mathien, 1991; Beck & Mckeown, 1994; Rosa, 1994; Blanco & Rosa, 1997; Seixas, 2000). They offer a trajectory that ties individuals' decisions and actions in the present and for the future to the longer course of events, whether expressed in the struggle for human rights, sacrifice for the national good, moral uplift, or economic well-being (Seixas, 2000). The framework can also provide a danger to be avoided, and act as a consolation for an unavoidable fate, class struggle, or gender equality (Rosa, 1994). One may not mobilize for any social or national purpose without invoking history to support it. In reality, the idea of social change or even conscious ideal of social conservation makes no sense without a historical orientation in which to frame it (Beck & McKeown, 1994; Blanco & Rosa, 1997).

In his study of how people look to the past when they require justification for a claim or a state of affairs, Mathien (1991) found that the results of historical study can serve the purposes of morality. Historical study can be used to support attempts to see to it that the right things is done in particular cases, to provide evidence in arguments about what the right thing to do is in a certain circumstance, and to offer recommendations or to oppose a course of action. The uses of historical studies in offering certain moral purposes demand objectivity. However, as Mathien claim, such studies dealing with moral value fail to register objectivity because the results of decision making or problem solving are very much determined by personal beliefs or interests. Thus, Lichtman and French (1978) differentiate moral analysis from empirical study, stating "the moral values he chooses are a matter between himself and his conscience" (p. 75). In fact, people in the past lived under different circumstances (a number of historical concepts are quite specific and limited), but also experienced and interpreted the world through different belief systems. Thus, the concept of analogy, imagination, or empathy can be used to fathom worlds

unlike their own, contexts far from those they know, and ways of thinking and feeling that are alien to them (Husband, 1996; Lerner, 1997). However, applying analogous, imaginative, and empathetic concepts to understanding a society in the past always runs the risk of bringing a proximate understanding of the past through our presumptions. Such proximity of understanding the past would confuse our attempt to understand the reality of the present (Husband, 1996; Seixas, 1996).

Through ethical judgment of the past, history plays a prominent role in promoting citizenship (Wineburg, 1991; Hahn, 1994; Blanco & Rosa, 1997), leading the awareness of higher levels of citizen duties, increased participation, and an increased level of political efficacy. Hahn (1994) says that understanding history is important in preparing for citizenship in a democracy. Because the essence of democracy is decision making, the goal of preparing for citizenship is the development of informed, reflective citizens who have the will to participate as well as the skills of analysis and decision making and values respecting human dignity and rationality (Hahn, 1994). Without history, no nation can enjoy legitimacy or command patriotic allegiance (Lerner, 1997). The National Center (1994) emphasizes the methods regarding historical issues-analysis and decision making revealed in the historical record as:

> If well chosen, these activities also promote capacities vital to a democratic citizenry: the capacity to identify and define public policy issues and ethical dilemmas; analyze the range of interests and values held by the many persons caught up in the situation and affected by its outcome; locate and organize the data required to assess the consequences of alternative approaches to resolving the dilemma; assess the ethical implications as well as the comparative costs and benefits of each approach; and evaluate a particular course of action in light of all of the above… (p. 32).

In order to offer alternatives for current problems, to challenge others' evaluations, and to defend one's own positions, students must be able to identify relevant or appropriate historical antecedents (National Center, 1994). Yet, Pratt (1974) argues that there tends to be favorable

attitudes, beliefs, or values toward out-groups that are often juxtaposed with such expressions as neighborliness, cooperation, and international goodwill. Therefore, the practice of ethical judgment of the past event should not be used simply to reinforce prejudices or inject a lethal dose of 'us-themism' (Blanco & Rosa, 1997; Pratt, 1974).

In short, historical thinking requires the following abilities: to imagine oneself in situations unlike anything one is ever likely to experience; to develop a sense of temporal structure of events unfolding over time; to develop hypotheses about cause and effect and to assess how well one's hypotheses fit the facts; to construct sound historical arguments and perspectives which inform decisions about contemporary society; to recognize that there will always be counterarguments available that appear to contradict one's hypotheses; and to articulate one's own values precisely, making sure that one's conclusions follow logically from the evidence. Therefore, historical thinking requires the learner to problematize the concept of history and challenge his view of historical discipline, offering something beyond mere facts beamed through time, as the crucial step for capturing the past events in a person's learning. Historical thinking also offers the learner an historical literacy environment beyond the simple transmission of historical knowledge or narratives, and provides an opportunity to articulate his own values precisely for a meaningful learning of the past.

# 3. RESEARCH DESIGN AND METHODOLOGY

## 3.1. Research Design

This study involves a descriptive study looking at the practices of classroom history assessment in middle and high schools in Korea. This study illustrates the degree of alignment of classroom history assessments with national curriculum and depicts how history teachers construct and practice their assessments. The analytic and empirical work of this study was drawn from the analysis of assessments including multiple-choice test items, short answer questions, and performance assessments. This study first employs an alignment study that examines the relationship between classroom history assessments and objectives for Korean history outlined in the 7$^{th}$ National Curriculum. To accomplish this, the cognitive complexity and content of assessments were analyzed, including the extent to which the history assessments reflected the complex historical thinking processes and breadth of historical knowledge representative in the curriculum. Due to performance assessments provided from the schools did not assess the content in the objectives of the curriculum, these assessments were evaluated only with regard to cognitive demands presented in the assessment. Second, this study evaluates the quality of the assessments in terms of formatting concerns and writing the stems and choices of multiple-choice test items. Third, teachers' responses to a brief survey regarding teacher preparation courses and teacher professional development programs related to classroom assessments were also examined in order to understand the results of the analyses of test items and assessments. In addition, demographic information, such as teachers' current teaching

assignment (i.e. middle or high school and private or public school) was examined to determine whether there are differences in teachers' practices of classroom history assessment from the curriculum. This chapter discusses the process of data collection, the characteristics of data and participants, and the framework of data analysis.

## 3.2. Description of the Data and Procedures of Data Collection

The Data

For the collection of classroom history assessments, the targeted secondary schools were 32 (22 middle school assessments for ninth grade[8] and 10 high school assessments for tenth grade[9]). These schools are located in a metropolitan city in the southern part of Korea. In general, they are attended by students whose families are from the upper-middle, middle, and lower classes and from culturally homogeneous communities situated in residential, semi-residential, and commercial areas. The city includes 70 middle schools (one national, 42 public, and 27 private schools) and 44 high schools (one national, nine public, and 34 private schools) (MOE & KEDI, 2000[10]). Twenty-two middle schools and 10 high schools were randomly selected. It is important to note that students at the same grade level were assessed by one test developed by history teachers for the grade.

For this study, test items and assessments for Korean history that were developed in the 2004 school year were collected[11]. Each school provided four formal tests a year (one school

[8] Middle schools include seventh to ninth grades.
[9] High schools include tenth to twelfth grades.
[10] In Statistical Yearbook of Education, the number of middle school teachers is 3 in national, 45 in public, and 25 in private schools (total 73). The number of high school teachers is 2 in national, 17 in public, and 81 in private schools (total 100) (p. 128, 168, 214, & 280).
[11] School year in Korea is from March 1st to February 28th of the next year, and students have their final examination in December.

year consists of two semesters: two formal tests a semester), which included multiple-choice and short answer questions (60% to 70%), and performance assessments (30% to 40%). One test for middle schools had from 16 to 25 multiple-choice and short answer questions while one test for high schools has from 28 to 34 multiple-choice and short answer questions. The majority of the test items from those schools were multiple-choice questions (more than 98%); the rest of the test was composed of short answer questions (less than 2%). Performance assessments were generally implemented one or two times in a semester: depending on teachers and schools, the frequency of implementation of performance assessment was somewhat different. Performance assessments, in general, included 30% to 50% of students' performances (generally written products), 25% to 30% of students' attitude and class participation in the class, and 25% to 30% of homework or organization of class materials. This study does not include the assessments of students' attitude, homework, and class participation, rather it focused on assessment tasks for students' written products. Therefore, the requirements for and explanations of performance assessments were collected.

The results of the test are generally used to provide grades for students and parents and are recorded on the Home School Record (HSR: *naesin*) for enrolling in higher scholastic levels. Academic and vocational high schools require that ninth grade middle school students submit 100% of their Home School Records; 80% of the HSR[12] comes from subject grades that reflect 50% of the ninth grade record[13]. In addition, the HSR for high school students is also used to

[12] The rest of 20% of HSR comes from 5% of attendance, 5% of behavioral development, and 10% of extra curriculum, including 5% of volunteer activity. The admission for science and foreign language high schools require between 30 to 40% of HSR, but does not require the HSR of history. Athletic high schools require 30% of HSR, and art high schools require 50% of HSR. In case of art high school, the criteria for the selection of students are different depending on the program.

[13] 20% from seventh grade record and 30% from eighth grade record

gain admission from college; depending on the need of the college, 30-40% of HSR is considered.

Ninth (middle school) and tenth grades (high school) were chosen because this study examined whether there are differences between the two grades in terms of teachers' practices of assessment. Ninth graders in middle schools learn Korean history from the establishment of the Chosōn Dynasty (1392) to modern society and eight grade history covers up to the end of 14th century. In tenth grade, they learn Korean history again from the Stone Age to modern society by thematic history, such as political, social, cultural, and economic, and in a more comprehensive ways[14]. Because high schools prepare students to take the exam, Higher Education Ability Test (HEAT)[15] for college, students are required to demonstrate broader abilities based on the knowledge and skills learned in middle school history class, including the problem-solving, decision-making, and community participation. On the other hand, ninth grade middle school students are relatively free from the pressure of the College Entrance Examination, so that they are also expected to master a variety of knowledge and skills, including the comprehension of factual knowledge and basic concepts and principles of each domain and the acquisition of the ability to solve individual and social problems.

In addition, 28 of the 32 middle and high school history teachers who provided their assessment items responded on survey questions about pre-service programs in their colleges and in-service activities during their teaching profession related to classroom assessment. In general, students who prepare to become teachers are required to take teacher preparation courses, such as educational evaluation, educational psychology, educational administration, etc. Teachers are also required to participate in teacher professional development activities related to subject

[14] For eleventh grade, Korean history is an elective course.
[15] HEAT is based on the model of the Scholastic Aptitude Test of the United States.

matter during summer or winter vacation, 10 to 15 days a semester. Those coursework and teacher development activities are intended to improve the quality of instruction, to adopt advanced technology to the classroom, or to manage students' behavior at the school.

Procedures of Data Collection

Data was collected for approximately one and half months. In order to collect assessment data, the schools to visit among the list of schools were identified: 22 of the 70 middle schools and 10 of the 44 high schools were randomly selected. Then, consent letters were sent to the principals and history teachers of the targeted schools in order to obtain the permission for this study (one consent letter for principals, one consent letter for history teachers, and one supporting letter from research advisor: Appendix A) with survey questions. Two or three days after mailing the letters to each school, the researcher called to the schools to make an appointment for meeting with principals and history teachers at the school. After that, the researcher visited two or three schools a day and explained to the principals and history teachers the purposes of this study and its benefits for classroom history assessments in Korea. Then, the researcher received the copies of test items and assessments from history teachers with their survey answers.

## 3.3. Instruments for Data Analysis

### 3.3.1. Analysis for Alignment of Assessments and Objectives

This study for the analysis of test items had three criteria in order to judge the quality of alignment between actual classroom assessments for Korean history and the 7th National Curriculum: *depth of understanding*, *breadth of knowledge*, and *balance of representation*. In order to meet all these criteria, classroom assessments developed by history teachers were

examined according to the educational objectives required by the curriculum (Appendix C). These criteria were based on the framework for alignment between assessments and standards developed by substantial studies (Lane, et al., 2000; Parke & Lane, 1999; Portor, 2002; Rothman, Slattery, & Vranek, 2002; Stern & Ahlgren, 2002; Webb, 2001).

***Depth of understanding*** requires the alignment of depth of knowledge by examining the relationship between the cognitive thinking process required by the assessment and the curriculum (Webb, 2001). This criterion focuses on whether assessments measure what the expected cognitive skills are in the curriculum. This criterion relied on the framework, depth of knowledge, developed by Webb (2001) and National Standards for History of the United States (National Center, 1994) in terms of historical thinking, considering the three levels[16] of expected learning outcomes for differentiated education stated in the 7th National Curriculum for social studies in Korea.

*Historical understanding level 1***:** Level 1 requires students to recall basic historical knowledge or simple facts (Bloom, et al, 1987; Quellmalz, 1985, 1991; Marzano, et al, 1988, 1993; MOEHRD, 1999, 2001; Webb, 2001). This level does not measure complex mental thinking processes, but requires students to recite historical events, names, people, or places. Students are expected to remember and reproduce basic ideas presented in textbooks or in historical materials that are covered in the class in order to answer the questions. This level also includes general ideas of information assessed throughout the different historical eras and general historical concepts. Some of examples for assessing level 1 performance are:

- Identifying when the event occurred, who was involved in, or where it took place.

[16] The 7th National Curriculum provides three levels of different learning outcomes for social studies: the first level that students understand basic concepts of social phenomena and theoretical meanings; the second level that students use and apply the knowledge acquired to social situations; and third level that students predict the future based on current knowledge.

- Identifying basic historical concepts or general phenomena of each different era.

*Historical understanding level 2*: Level 2 indicates the test items that measure historical comprehension with some mental processing, asking beyond basic historical facts (Bloom, et al, 1987; Quellmalz, 1985, 1991; Marzano, et al, 1988, 1993; MOEHRD, 1999, 2001; Webb, 2001). At this level, students are expected to reach an understanding of the message presented in a historical account, narrative, story, or visual materials, so that they can interpret the course of the event with the understandings of which people were involved and results that followed (National Center, 1994; Perfitti, et al., 1994, 1995; Seixas, 2000). For this activity, students can also develop empathy as opposed to present-mindedness (National Center, 1994; Seixas, 1996). In order to comprehend, students are required to understand the significance of the past to their lives and society. Also, they are required to understand history as constructed interpretations by historians rather than as facts of the past. This level requires students' chronological thinking through identifying the temporal structure of a historical narrative or the temporal order of historical events (Hoodless, 1996; National Center, 1994; Stow & Haydn, 2000). Some abilities expected at this level are:

- Identifying and summarizing the historical event, identifying who was involved, where and when it happened, and what were causes and consequences.
- Identifying the central questions and viewpoints of historical passages.
- Interpreting main ideas presented in historical reading and visual materials.
- Explaining some patterns of historical progressions and duration in temporal structures.
- Imagining oneself in situations unlike anything one is ever likely to experience (National Center, 1994; Portal, 1987; Shemilt, 1984).

*Historical understanding level 3*: Level 3 involves some of complex historical reasoning processes and disciplinary activities (Bloom, et al, 1987; Quellmalz, 1985, 1991; Marzano, et al, 1988, 1993; MOEHRD, 1999, 2001; Webb, 2001). This level requires analytical thinking about the interpretations of a historical event, based on students' awareness that there are multiple interpretations existing about a historical event depending on historians' view points (Leinhardt, et al.,1994a, 2000; National Center, 1994; . Students are engaged in the analysis of historical interpretations competing with one another in terms of their perspectives, ideas, or values about a certain event (Perfitti, et al., 1994; Rodrigo, 1994). This analytic activity involves the analysis of the credibility of authorship and sources. At this level, students are also required to analyze the causes and effects of a historical event from a variety of social, economic, or political phenomena surrounding the event (Lienhardt, 1994b). Moreover, students are expected to construct their own argument challenging or supporting a certain viewpoint of a historical narrative. The following are some of abilities required at this level:

- Comparing and contrasting the perspectives and viewpoints presented in historical interpretations.
- Distinguishing interpretations grounded in historical evidence from unsupported ideas or opinions.
- Analyzing causes and consequences based on multiple causations.
- Identifying the credibility of the author or source of historical documents.
- Understanding the uncertainty and ambiguity of the past illustrated in historical interpretations.
- Inferring the historical data or evidence in order to form historical generalization (Marzano, et al., 1993; National Center, 1994; Wineburg, 1991a, 1991b, 2000).

*Historical understanding level 4*: Level 4 involves a complex historical reasoning process, requiring *s*tudents to demonstrate their deep understandings of history (Bloom, et al, 1987; Quellmalz, 1985, 1991; Marzano, et al, 1988, 1993; MOEHRD, 1999, 2001; Webb, 2001). Students are involved in their own interpretations of history, formulating historical questions from historical documents and using a variety of historical data as supporting details of their interpretations (Holt, 1990; Greene, 1994; Leinhardt, 2000; Voss & Wiley, 2000). They are expected to fill in the gap between available materials in order to reconstruct their own interpretations (Collingwood, 1972). Moreover, students are expected to identify issues and problems of the past, so that they can find alternative courses of action for the solutions for problems in the contemporary society (National Center, 1994). This level also requires students to make an ethical judgment of a past event, and about the actions of individuals and groups in terms of human rights, clarifying moral implications of their own lives (Blanco & Rosa, 1997; Mathien, 1991; Seixas, 1996, 2000). The required abilities for level 4 are:

- Reconstructing or supporting interpretations with relevant historical evidence, formulating questions from historical documents.
- Identifying the issues and problems of the past.
- Evaluating alternatives from the past event that contribute to the current problems.
- Making ethical decisions, considering the power involved in the past.

Alignment criteria for depth of understanding used seven criteria in order to assure whether test items meet the objectives of the curriculum. Table 5 illustrates criteria that were intended to analyze the degree of depth of understanding presented both in test items and in curriculum objectives. The first column provides the % of items that are inconsistent with the content of objectives (Item≠objective). The second column provides the % of items whose levels of

historical understanding are aligned with the levels of objectives (Item level=objective level). The third column provides the % of items that assess at 2 levels higher than the objectives (Item is 2 level>objective). The fourth column provides the % of items that assess 1 level higher than the objectives (Item is 1 level>objective). The fifth column provides the % of items that assess 1 level lower than the objectives (Item is l level<objective). The sixth column provides the % of items that assess at 2 levels below the objectives (Item is 2 level<objective). The seventh column provides the % of items that assess at 3 levels below the objectives (Item is 3 level<objective). This analysis also depicts the results by different school types and grades.

Table 5: Analysis Tool for the Alignment of Depth of Understanding

| Differences of Cognitive Demands | Middle or High Schools | | |
|---|---|---|---|
| | Public | Private | All |
| | (%) | (%) | (%) |
| Item ≠ Objective | | | |
| Item level = Objective level | | | |
| Item level 4, Objective level 4 | | | |
| Item level 3, Objective level 3 | | | |
| Item level 2, Objective level 2 | | | |
| Item is 2 level > Objective | | | |
| Item level 4, Objective level 2 | | | |
| Item is 1 level > Objective | | | |
| Item level 4, Objective level 3 | | | |
| Item level 3, Objective level 2 | | | |
| Item is 1 level < Objective | | | |
| Item level 3, Objective level 4 | | | |
| Item level 2, Objective level 3 | | | |
| Item level 1, Objective level 2 | | | |
| Item is 2 level < Objective | | | |
| Item level 2, Objective level 4 | | | |
| Item level 1, Objective level 3 | | | |
| Item is 3 level < Objective | | | |
| Item level 1, Objective level 4 | | | |

***Breadth (range) of knowledge*** was needed as a criterion in order for assessments to align with the curriculum. This criterion is a measure of 'coverage' judging whether assessments measure the span of knowledge that is representative of the content domains in the curriculum (Bhola, et al., 2003; Rothman, Slattery, & Vranek, 2002; Webb, 2001). In other words, this criterion evaluates what percent of objectives in a unit is covered by test items.

***Balance of representation*** evaluates the evenness of the distribution of knowledge between test items in assessment and targeted objectives in the curriculum (Bhola, et al., 2003; Webb, 2001). This criterion is consistent with the balance of representation component in the Webb (2001). The assumption behind this evaluation is that test items across the targeted objectives should be distributed equally. In addition, this criterion examined the degree to which test items assess a certain objective given in the curriculum more than another, which is consistent with the criterion of balance in the Achieve model (Rothman, Slattery, & Vranek, 2002). Thus, looking at the equal distribution of items across objectives, this criterion provides a comparison between the emphasis of content offered by the set of test items and the emphasis of content described by the curriculum.

Table 6: Analysis Tool for the Alignment of Breadth of Knowledge and Balance of Representation

| Objectives | Consistent | Broad | Narrow | Inconsistent |
|---|---|---|---|---|
| Objective 1<br>Objective 2<br>Objective 3<br>Objective 4<br>Objective 5<br>Overall | | | | |

To obtain the degree of the breadth of knowledge and the balance of representation, four categories were used: *consistent*; *broad* (a test item is not specific and broader than an objective);

*narrow* (a test item is somewhat consistent and narrower than the targeted objective); and *inconsistent*. The matrix of Table 6 presents two dimensions that analyze test items by objectives in a unit based on scale points. Using the scale points presented above, the consistency of the two criteria is described at the intersection between unit objectives and scale points in the table. The analysis of test items is illustrated by the numbers and percentages of test items across the objectives, differentiating the results of public and private schools. This analysis indicates the results of how much historical knowledge assessed covers the range of the knowledge stated in the objectives. Also, the analysis presented in the columns for *consistent* and *broad* describes how many test items are distributed across the objectives. Since the test items inconsistent with any objectives cannot be defined, only the overall number of items and percentages are provided.

### 3.3.2. Analysis of the Quality of Assessment

In order to evaluate the quality of multiple-choice test items and short answer questions based on the original taxonomy developed by Haladyna, et al., (1989, 2002) for multiple-choice item-writing, the rules were modified as illustrated in Table 7. The original taxonomy has been developed as a guideline for writing multiple choice items. This taxonomy was validated by the consensus appearing in the 27 textbooks he studied and by his research. The rules for this study include three themes: formatting concerns, writing stems, and writing choices. By using this taxonomy, the quality of test items was analyzed with tally marks.

Table 7: A Taxonomy of Multiple-Choice Item Writing Rules

(Haladyna, et al., 1989, p. 41; 2002, p. 312)

| Guidelines/Rules/Suggestions |
|---|
| *Formatting concerns* |
| Avoid complex MC format. |
| Keep vocabulary simple. |
| *Writing the stem* |
| State the stem in either question or completion form. |
| Minimize the amount of reading. |
| Clear directions. |
| Central idea in the stem. |
| Avoid window dressing (excessive verbiage). |
| Use positive terms. |
| *Writing the choices* |
| Write as many plausible distractors. |
| One right answer. |
| Logical/numerical order. |
| Keep choices not overlapping. |
| Provide choices homogenous. |
| Use equal length. |
| Use carefully 'none of the above.' |
| Avoid 'all of the above.' |
| Avoid the term "not" in choice. |
| Avoid giving clues to the right answer. |
| Use common errors of students. |

### 3.3.3. Comparison between Teaching Assignments

For one of the purposes of this study, demographic information, such as teachers' current teaching assignment (private or public school and middle or high school) was analyzed. Based on the analysis of assessments, this study indicates the similarities and differences in the degree of alignment between assessments and objectives and the quality of assessments based on types of schools and grade levels teachers assigned.

### 3.3.4. A Brief Survey of History Teachers

This study surveyed history teachers with a few open-ended questions regarding classroom assessment, including teacher preparation programs in their colleges, teacher professional development activities, and suggestions for their future professional development programs (Appendix B). With regard to teacher preparation course-work and teacher professional development programs, the survey provided seven questions for each topic about their learning such as theory of assessment, design of selected-response and constructed-response test items and performance assessments, and interpretation of assessment results. These questions were answered only by the teachers who received teacher preparation coursework or teacher professional development programs related to assessment.

## 3.4. Interrater Reliability of the Study

Sixty-six of 1,510 middle school test items (approximately 4%) and 71 of 1,315 high school test items (approximately 5%) were analyzed by two history teachers in order to examine interrater reliability. Middle school test items were from the four tests given in 1 year school, and high school test items were from the 2 tests given in one semester. For the analysis of test items, the two raters were trained about on how to evaluate the levels of historical understanding demanded by both test items and objectives and how to compare the content of test items to the content of objectives. Written criteria for the levels of historical understanding, a copy of the educational objectives in Korean history, and test items were provided. The coring was done independently by the 2 raters and the researcher.

With regard to the results for depth of understanding, the levels of historical understanding presented in each test item and objective judged by the raters was compared to the

researcher's rating. Twenty middle school items and 23 high school items whose content was not consistent with the content of objectives were excluded from the comparison. Forty-three out of 46 test items (93%) for the middle school were judged as consistent with each other, while 42 out of 48 test items (88%) for the high school were examined as consistent. It was found that high school test items required higher levels of historical understanding than the test items of the middle school, which might have resulted in a lower interrater reliability.

Regarding the alignment of content, the classification of each test item across objectives judged by the raters and the researcher was compared. Different from the comparison of the levels of historical understanding, inconsistent items with the content of objectives were included in the comparison. With respect to consistent, broad, narrow, and inconsistent, 94% of the classifications for 66 test items from the middle school by the raters and by the researcher was consistent, while 90% of the classifications for 71 test items from the high school by the raters and by the researchers were judged as consistent. Both raters were somewhat more generous than the researcher, indicating more consistent items with the content of objectives.

## 3.5. Significance of the Study

The value of this study is that the findings can be contributed to the field of history education in the context of understanding classroom assessment by offering important suggestions, alternatives, or remedies. First, this study will contribute by clarifying the practice of history assessment in the secondary school classroom. This study will illustrate the degree to which classroom assessments developed by history teachers reflect the educational goals of the 7$^{th}$ National Curriculum, which emphasizes the mastery of complex-thinking processes and the application of knowledge to the real world context. The results of this study will provide an

indicator both for history teachers and experts on national curriculum on whether to modify either classroom assessments or expected students' outcomes. Also, history teachers for their own professional development will learn from this study by considering the quality of their own test items, regarding content, coverage of historical knowledge, complex thinking process, and the educational targets embodied in the 7th National Curriculum. Through the feedback from this study, curriculum experts in history assessment will realize which educational objectives and content are emphasized more importantly at the grade level. Moreover, history teachers will be informed about the limitations of the test items and assessment tools and will be encouraged to consider the findings of the study for their future constructions of assessment items.

The findings will also contribute both to history teachers themselves and researchers in the field of history assessment by giving attention to the issues of learning and assessment in the school setting. Given that analyzing textbooks has dominated the field of research in history education in Korea, this study will enlarge the horizon of discussion on the classroom assessment of history by realizing history teachers' own practices of assessment.

Moreover, this study might be used to find the current frame of references 1) for assessment measuring reasoning processes influenced by various standards appearing in philosophical and psychological research; 2) for the nature of assessment regarding the definition, purposes, and tools considering the benefits for students' learning; 3) for historical reasoning skills proposed by National Standards for history in the United States and influenced by cognitive studies and discipline of history. Also, this study provides the current educational situations in Korea and history education under the 7th National Curriculum.

Lastly, this study will provide benefits for teacher education programs and teacher in-service programs that strive to enable pre-service teachers and practicing teachers to become active consumers in using history and assessment in effective and relevant ways.

# 4. RESULTS

## 4.1. Introduction

The results of this study for classroom history assessments are presented in six major sections: 1) nature of history assessments; 2) depth of historical understanding of history assessments; 3) breadth of knowledge and balance of representation of history assessments; 4) performance assessments; 5) quality of test items and alternatives; and 6) teachers' responses to a survey. The first section describes the primary sources from middle and high schools for the data analysis, such as the size and types of data. For the second and third sections, descriptions are provided regarding the extent to which the history assessments reflect the requirements of the 7$^{th}$ National Curriculum for Korean history in terms of historical knowledge and cognitive demands. Also, for these sections, the basic criteria for coding are illustrated by introducing examples of both test items and unit objectives. The fourth section depicts the topics and cognitive complexity of performance assessments. The fifth section presents the quality of test items in terms of their formatting and writing of the stem and choices. Finally, the sixth section illustrates the results of teachers' responses to a survey, describing pre-service coursework and professional activities with respect to history assessments. For this chapter, items and objectives were analyzed one by one based on the coding schemes that had already been developed for the analysis of the history classroom assessments.

## 4.2. Nature of History Assessments

The primary source of the data analyzed was from teacher-developed assessments for Korean history for middle school 9th grade and high school 10th grade. These assessments were from four tests during the 2004 school year and include multiple-choice and short answer questions and performance assessments. Because performance assessments did not assess the knowledge presented in the curriculum, and only a brief description of them was provided, the analyses of assessments focused on the multiple-choice and short answer questions only for the depth of understanding, breadth of knowledge, balance of representation, and the quality of test items.

### 4.2.1. Middle School Test Items

As Table 8 indicates, the total number of test items analyzed was 1,510 from 22 middle schools. The middle schools provided from 64 to 100 test items for the four tests in the 2004[17] with these schools, in general, there was 16 or 17 test items per test[18]. The test items were mainly multiple-choice items (98%) and a few short answer questions (2%). Private schools tended to offer more short answer questions than public schools; only one public middle school among 13 provided short answer questions. However, the sample size is too small to make a judgment for the differences between public and private schools. Moreover, none of these 22 middle schools offered any other questioning formats such as essay tests.

[17] One test from one school was excluded from this analysis because it measured part of the content for 8th grade, which is not the targeted knowledge. Thus, only 39 items from this school were analyzed.
[18] Since 2002 under the implementation of the 7th curriculum, as part of social studies, history assessments for 9th grade can be worth 50 point on a 100-point test, with the other 50 points allots to social studies questions.

Table 8: Numbers of Middle School Test Items

| *Public Middle School (n=13)* | | | *Private Middle School (n=9)* | | |
|---|---|---|---|---|---|
| School | Multiple-Choice | Short Answer | School | Multiple-Choice | Short Answer |
| | n=930 | | | n=580 | |
| A | 68 | - | N | 66 | - |
| B | 66 | - | O | 67 | 3 |
| C | 77 | - | P | 64 | - |
| D | 71 | - | Q | 66 | - |
| E | 66 | - | R | 39* | - |
| F | 68 | - | S | 48 | 16 |
| G | 71 | - | T | 66 | - |
| H | 67 | - | U | 80 | - |
| I | 68 | - | V | 62 | 3 |
| J | 66 | - | | | |
| K | 66 | - | | | |
| L | 71 | 5 | | | |
| M | 100 | - | | | |
| Total | 925 (99%) | 5 (<1%) | Total | 558 (96%) | 22 (4%) |
| **Overall** | 1,483 (98%) | 27 (2%) | | | |

* One test (n=20) was excluded because the test evaluated the knowledge for 8th grade.

### 4.2.2. High School Test Items

Ten high schools provided 1,315 test items, ranging from 111 to 137 test items for each of the four tests in 2004. These schools on average offered 33 test items per test. As Table 9 presents, as with the middle school test items, the high school test items were mainly multiple-choice test items (99%), including few short answer questions (1%). Only two private schools among 10 provided short answer questions while none of the public schools offered this type of question. However, it is hard to generalize this situation to all schools due to the small sample.

Table 9: Numbers of High School Test Items

| *Public High School (n=4)* | | | *Private High School (n=6)* | | |
|---|---|---|---|---|---|
| School | Multiple-Choice | Short Answer | School | Multiple-Choice | Short Answer |
| | n=529 | | | n=786 | |
| A | 133 | - | E | 121 | 13 |
| B | 132 | - | F | 140 | - |
| C | 132 | - | G | 105 | 6 |
| D | 132 | - | H | 132 | - |
| | | | I | 132 | - |
| | | | J | 137 | - |
| Total | 529 (100%) | | Total | 767 (98%) | 19 (2%) |
| **Overall** | 1,296 (99%) | 19 (1%) | | | |

### 4.3. Results for Depth of Historical Understanding

Judgments about the depth of historical understanding focus on alignment between the cognitive demand described by each targeted objective and the cognitive demands of each test item. In order to examine the relationships between test items and objectives in terms of their cognitive demands, this study considered both the verb and content of objectives. In general, educational objectives consisted of two parts—verb and content: the following is an example of an objective:

| Verb | Content |
|---|---|
| To analyze | the purposes and meanings of reform policy by Hung-son Tae-wōn'gun. |

With respect to this format of objectives, the method used to analyze the items from the 32 middle and high schools involved the following steps:

1. Compare the content of each test item and the content of each objective.
2. Identify the cognitive level of each test item and each objective.

3. Compare the cognitive level of each test item with the cognitive level of the verb used in the objective.

A test item whose content is unrelated to the content of a targeted objective was not considered for depth of understanding. However, a test item whose content is somewhat related to or broader than the content demanded by the targeted objective was analyzed in terms of its level of historical understanding to the level of historical understanding demanded in the objective. Test items and objectives were evaluated to determine whether they included either historical reasoning skills of students only the ability to recall knowledge. The next sections provide examples of test items and objectives based on the criteria for the levels of historical understanding.

### 4.3.1. Levels of Historical Understanding of Test Items

Each test item requires a certain levels of cognitive demand. If a test question fell into the category 'level 1 historical understanding,' it measured the ability to recall general historical knowledge and facts presented in materials or textbooks. The followings are the examples that were coded for level 1 historical understanding.

Q: What is NOT an explanation for Gwang-mu Reform[19]?

A. Reinforced military forces by reforming the army system
B. Strove to develop industry and improve education
C. Established a various companies and factories in order to develop commerce and industries
D. Established a various schools as well as vocational and medical schools in order to train technicians and administrators
E. Advanced a reform meeting the demands of Tōng-nip Hyōp-hoe[20]

[19] The modern reform in 1895 under the Tae-han Che-guk, including the improvement of education and industry.

This test item requires students to recall the information about Gwang-mu Reform that they learned in the class from their memory. For this question, students are not provided any text either to analyze or to infer information from. Students depend only on their ability to remember what they acquired from textbooks or class explanations to answer to the question. The next item was also treated as requiring level 1 understanding.

> Q: The following diagram is about the developmental process of a nation. What is an <u>incorrect</u> explanation for one of the processes?
>
> a=group society ⟶ b=tribal society ⟶ c=patriarchal[21] society ⟶
>
> d=federation state ⟶ e=centralized governing state
>
> A. a=A migratory life of a group of people, centering around a wise elder.
> B. b=Beginning farming, establishing tribes, and implementing Chokoehon.[22]
> C. c=Appearance of an individual's property, social stratus, and the idea of God's selection of a predominant tribe.
> D. d=Appearance of a king who led politics and integrated the heads of tribes into his own subjects.
> E. e=Preparation of national systems by promulgating laws and adopting Buddhism.

This question provides a diagram that helps students understand the process of how a group society became a centralized governing country. However, this question was categorized as level 1 historical understanding because it requires students to answer about general concepts of each different society in the past. Students have to bring the general knowledge that they already have obtained either from history class or from instructional materials and not make any conclusions of their own.

Historical understanding level 2 goes beyond general concepts. For example:

---

[20] The Independence Club that was led by So Chae-p'il (1896) acted to awaken the public to the needs of modernization and national sovereignty by asserting mass education, publishing Hangul news paper, and organizing meeting for the mass. It was crushed by the royalists in the Imperial Association.
[21] The first ruling class that appeared in the Bronze Age with a political and economic power.
[22] Social tradition that tribal people married ones from out of its tribe.

Q: The following is the fact that happened during the modernization of the Chosōn. Which answer occurred between (a) and (b) in time?

(a) A coalition cabinet by a pro-Russia faction was established due to Samguk (three Countries[23]) intervention.
(b) A pro-Russia cabinet was established when Kojong[24] took refuge in the Russian legation

A. Old military rose up causing a riot with discontentment.
B. Reformists, such as Kim Ok-kyun and Hong Yong-shik carried out a coup d'etat.
C. The first modern reform was implemented with the T'ongnigimuamun[25] as the center.
D. After Japan assassinated the Myōngsōng Empress[26], Elmi Reform[27] was implemented.
E. T'ong-nip Hyop-hoe was established and deployed the activities for public enlightenment.

This question measures level 2 historical understanding by asking students to identify a cause and effect relationship of events, determining the proper sequence of past events. This question is related to chronological thinking, which includes the ability to choose an event which occurred between events (a) and (b), placing a cause and result in order of occurrence.

Historical understanding level 3 requires students to reason for a deeper understanding of the past. The following is a question that was classified as historical understanding level 3.

Q: The following (1) is a phenomenon that was presented in the Bronze Age. What is the result of this social phenomenon that can fill in the blank (2)?

| (1) | | (2) |
|---|---|---|
| Use of a bronze sword<br>Increase of surplus products<br>Increment of private ownerships | → | ( ) |

A. Development of an equal society

[23] International pressure to Japan by three countries Russia, Germany, and France in 1895 to overt Japanese aggression as a counter to the Russian threat. After this political intervention, Japan returned Liaodong Peninsula to Qing China.
[24] The 26th king of the Chosōn who proclaimed the nation as the Great Han (The-han) Empire (Che-guk) in 1897.
[25] Office for the Management of State Affairs.
[26] Queen Min, the empress of the Tae-han Che-guk.
[27] A reform by a Pro-Japanese cabinet.

B. Emergence of federation states
C. Promotion of division of social stratus
D. Development of a product-monetary economy
E. Preparation for a centralized governing structure

This question measures how students form generalizations from historical evidence or cases given. Students have to integrate the social phenomena or details provided in order to reach conclusions about general situations in this era using the skill of inference. The following is also a question identified as level 3.

> Q: What is the correct explanation about features about the Three Kingdoms based on the following text?
>
> There was a rock called Jōngsa-am in Hoam temple in Paekche. When the country selected a prime minister (chae-sang), it put the names of three or four candidates in a box…. and determined the person as a prime minister who had a signature on his name.
>
> A. Royal authority was autocratic.
> B. Confucianism was adopted in political ideology.
> C. The government was king-centered aristocracy.
> D. Shamanism led the government.
> E. The tradition of tribal society was cut off.

In this question, students' skill to analyze the main idea of the given historical record is assessed. Students use inferential skills to identify the relationship between the social status of the prime minister (Chae-sang) and the power (king). Once this relationship is analyzed, students understand information related to Paekche, which had a centralized governing power and related to the heads of tribes who were placed as a subject of the king.

Level 4 historical understanding requires a complex historical reasoning process such as formulating historical questions, finding alternative actions to solve current problems, and supporting interpretations with relevant evidence. For example:

> Q: What is an <u>incorrect</u> evidence to support the conclusion presented in the following passage?

> The basic element of Chosōn society was a patriarchal family system. The society was established and run centering around it. All standards of life and ceremonies based on a patriarchal family system were strictly controlled under the Confucian order.
>
> A. It was widespread that the family that did not have its own son adopted a step-son.
> B. Because a kinship was established with the relatives of the paternal line as the center, the solidarity (bond) of the same family was promoted.
> C. According to the equal inheritance system of children, they were devoted to their parents in order regardless of their sex.
> D. This society prohibited the remarriage of widows and honored devoted sons and virtuous wives.
> E. A son from the second wife of Yangban could not apply to the examination for liberal arts and was excluded from ritual ceremony and inheritance.

In this question, students' ability to verify the accuracy of the evidence in order to support the ideas presented in the historical account given is assessed. To answer the question, students must identify the main ideas from the account that are Confucian order and patriarchal-social systems. These two elements are the criteria to prove the relevance of the evidence among choices. Students need to infer the interrelationship between the evidence and their conclusions.

#### 4.3.1.1. Levels of Historical Understanding of Middle School Test Items

According to the criteria explained above, the level of test items was determined depending on their levels of historical understanding. As Table 10 indicates, overall, 87% of the test items were the ones that measured the ability to recall knowledge such as general facts, situations, names of events or people acting in the past, and the ability to understand written or visual historical messages (74%, 13%, respectively). High level of historical reasoning skills was rarely assessed—Only less than 14% of the test items were at level 3 and 4. In fact, only three items among 1,510 items fell into the category of historical understanding level 4. Across schools, private schools tended to provide a slightly higher percentage of test items that were

classified as level 1 while public schools tended to provide a slightly higher percentage of test items that were treated as level 2 and 3. For level 4, the two types of school were determined to offer the same percentage of test items, both <1%.

Table 10: Numbers and Percentages of Middle School Test Items for Level of Historical Understanding (1)

| Level of Understanding | Middle School (n=22) | | | | | |
|---|---|---|---|---|---|---|
| | Public | | Private | | Overall | |
| | N | % | N | % | N | % |
| Level 1 | 670 | 72% | 450 | 78% | 1,120 | 74% |
| Level 2 | 121 | 13% | 68 | 12% | 189 | 13% |
| Level 3 | 137 | 15% | 61 | 10% | 198 | 13% |
| Level 4 | 2 | <1% | 1 | <1% | 3 | <1% |
| Total | 930 | 100% | 580 | 100% | 1,510 | 100% |

Across middle schools, as can be seen in Table 11, there were large differences for the level of historical understanding. For example, it ranged from 53% to 96% for level 1 and 0% to 27% for level 3. There were six schools among 22 that provided over 80% of the level 1 test items (two from public schools and four from private schools). However, there were eight schools that assessed historical understanding level 3 and 4 with less than 10% test items (two from public schools and six from private schools). One school never assessed historical understanding above level 2. In fact, there were only four schools that assessed historical understanding above level 2 with more than 20% of the test items (three from public schools and one from private schools). Public schools provided a slightly higher percentage of test items that assessed higher level of historical understanding than private schools.

Table 11: Numbers and Percentages of Middle School Test Items for Level of Historical Understanding (2)

| Public Middle School (n=13) | | | | | | Private Middle School (n=9) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Level of Understanding | | | | | | Level of Understanding | | | |
| School | | 1 | 2 | 3 | 4 | School | | 1 | 2 | 3 | 4 |
| | Total n=930 | | | | | | Total n=580 | | | | |
| A | n=68 | 79% | 3% | 18% | | N | n=66 | 92% | 5% | 3% | |
| B | n=66 | 74% | 21% | 5% | | O | n=70 | 81% | 13% | 6% | |
| C | n=77 | 77% | 12% | 11% | | P | n=64 | 63% | 22% | 15% | |
| D | n=71 | 70% | 11% | 19% | | Q | n=66 | 70% | 12% | 17% | 1% |
| E | n=66 | 68% | 14% | 17% | 1% | R | n=39 | 90% | 5% | 5% | |
| F | n=68 | 72% | 13% | 15% | | S | n=64 | 89% | 3% | 8% | |
| G | n=71 | 56% | 17% | 27% | | T | n=66 | 77% | 15% | 8% | |
| H | n=67 | 63% | 18% | 19% | | U | n=80 | 66% | 14% | 20% | |
| I | n=68 | 60% | 18% | 22% | | V | n=65 | 77% | 14% | 9% | |
| J | n=66 | 53% | 20% | 26% | 1% | | | | | | |
| K | n=66 | 74% | 15% | 11% | | | | | | | |
| L | n=76 | 96% | 4% | - | | | | | | | |
| M | n=100 | 84% | 8% | 8% | | | | | | | |

### 4.3.1.2. Levels of Historical Understanding of High School Test Items

Test items from high schools were also identified with respect to the level of historical understanding. Overall, high schools also had many low level test items with high levels of historical reasoning being assessed rarely. As Table 12 presents, approximately 73% of the

Table 12: Numbers and Percentages of High School Test Items for Level of Historical Understanding (1)

| Level of Understanding | High School (n=10) | | | | | |
|---|---|---|---|---|---|---|
| | Public | | Private | | Overall | |
| | N | % | N | % | N | % |
| Level 1 | 258 | 49% | 446 | 57% | 704 | 54% |
| Level 2 | 123 | 23% | 123 | 16% | 246 | 19% |
| Level 3 | 136 | 26% | 186 | 23% | 322 | 24% |
| Level 4 | 12 | 2% | 31 | 4% | 43 | 3% |
| Total | 529 | 100% | 786 | 100% | 1,315 | 100% |

items were at levels 1 and 2 while about 27% of the test items were classified at levels 3 and 4. Compared to public schools, private schools were more likely to have test items that assessed the ability to recall historical knowledge with fewer test items for level 2. However, private schools measured historical understanding level 4 with a slightly higher percentage of items.

Across high schools, the percentages of test items varied among the levels of historical understanding as shown in Table 13. For example, they ranged from 18% to 80% for level 1 and 9% to 39% for level 3. Among the ten high schools examined, three schools had high levels (level 3 or 4) of historical understanding for more than 40% of their test items. However, no schools measured historical understanding at level 4 with more than 10% of their test items. Two private schools among the six (33%) measured level 1 historical understanding with less than 35% of the test items while two public schools among the four (50%) assessed the same level of historical understanding with less than 35% of their test items. Also, two private schools (33%) had over 35% level 3 and 4 test items while two public schools (50%) had over 35% level 3 and 4 test items.

Table 13: Numbers and Percentages of High School Test Items for Level of Historical Understanding (2)

| | Public High School (n=4) Level of Understanding | | | | | Private High School (n=6) Level of Understanding | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| School | | 1 | 2 | 3 | 4 | School | | 1 | 2 | 3 | 4 |
| | Total=529 | | | | | | Total=786 | | | | |
| A | n=133 | 32% | 31% | 36% | 1% | E | n=134 | 80% | 10% | 9% | 1% |
| B | n=133 | 77% | 11% | 10% | 2% | F | n=140 | 28% | 31% | 35% | 6% |
| C | n=132 | 68% | 12% | 18% | 2% | G | n=111 | 33% | 23% | 40% | 4% |
| D | n=132 | 18% | 39% | 39% | 4% | H | n=132 | 61% | 17% | 19% | 3% |
| | | | | | | I | n=132 | 74% | 5% | 14% | 7% |
| | | | | | | J | n=137 | 62% | 9% | 28% | 1% |

### 4.3.2. Levels of Historical Understanding of Unit Objectives

The 7th National Curriculum for Korean history provides 50 objectives for 9th grade (middle school) and 66 objectives for 10th grade (high school), including general objectives and specific objectives for each unit[28]. 'General objective' refers to a lesson goal that covers the main historical knowledge and behaviors that students must accomplish for a certain unit. Each unit includes five or six general objectives requiring, in general, the cognitive process to use the skills 'comprehension,' 'comparison,' or 'inference.' 'Specific[29] objective' refers to a goal that allows students to use complex mental process and specific historical knowledge in order to achieve in-depth understanding of the content provided by a unit. These objectives require a high level of historical understanding, asking students to use and to analyze a variety of historical information and to form conclusions of past events. Each topic in a unit includes one or two specific objectives for deep understanding of the past. As presented in Table 14, the 7th National Curriculum for middle school Korean history provides a total of 32 general objectives including 18 specific objectives, and for high school it provides 30 general objective and 36 specific objectives. Compared to middle school, the curriculum includes more specific objectives for

Table 14: Numbers of Educational Objectives

| Unit Objectives | Middle School 9th | | High School 10th | |
|---|---|---|---|---|
| | N | % | N | % |
| General Objective | 32 | 64% | 30 | 45% |
| Specific Objective | 18 | 36% | 36 | 55% |
| Total | 50 | 100% | 66 | 100% |

[28] The national curriculum provides both 'unit objectives' and 'specific objectives' for each unit in order to provide differentiated education by adding high levels of mental activities. Here, 'unit objectives' are called as 'general objectives' differentiating them from 'specific objectives.'
[29] The curriculum calls 'specific objectives' as 'deepening learning process.'

high school history (36%, 55%, respectively).

The level of the objectives was coded in terms of whether the objectives focused on historical understanding including thinking skills as defined in the methodology section. Specifically, action verbs of the objectives were a clue to identifying the level of objectives. It should be noted that this curriculum does not include historical understanding level 1. For the category level 2, objectives use action verbs like 'to comprehend,' 'to explain,' and 'to understand.' For example, 'to understand the background and intention of Yōng-jo's implementation of the Tang-pyōng-chaek[30]' and 'to comprehend the establishment of Koryō marked the beginning of medieval society' were determined to be at level 2 of historical understanding. These objectives require students to acquire historical background, intention, and facts from instructional explanations and to produce their understanding by using different forms of communication.

Objectives that include 'to analyze,' 'to compare,' and 'to infer' were coded as level 3. Objectives classified into this level require students to use historical reasoning skills to understand the past and then form general ideas from historical data or evidence or compare various aspects between different periods. The following objectives were coded as level 3: 'to infer the significances and limitations of the Kabō[31] Reform by studying the process of reform and its content'; and 'to compare Eastern early modern society to Western early modern society.' Objectives that were evaluated as historical understanding level 4 use action verbs such as 'to explore' or 'to discuss.' These verbs require students to use heuristic methods by formulating questions from historical accounts, identifying problems of the past, or proposing alternatives to

[30] An appeasement policy in response to factional fights among political groups.
[31] A reform in 1894 by the Tōng-hak Peasant Uprising that was widely implemented through political, economic, and social areas.

solve the problems of the past. Examples of objectives that belong to level 4 are: 'to explore the significances of the Shil-ryok Yang-song Movement[32] after the Elsa Treaty[33]'; and 'to explore the efforts of the ruling class to rationalize the administrating order during the Chosōn era.'

With regard to the level of historical understanding presented in objectives, almost half of the objectives for both middle and high schools fell into the level 2 (49%, 48%, respectively) as Table 15 indicates, the percentage for level 3 for middle school is slightly higher than for high school (33%, 29%, respectively). However, 10th grade students were provided with more objectives at level 4. It is important to note that among the 50 objectives for 9th grade, one objective was excluded from the analysis because it sought to improve attitude. In fact, a total of three objectives were aimed at improving attitude by studying people's actions in the past. However, the other two[34] were included in the analysis and regarded as level 2 because of the important knowledge they required students to attain.

Table 15: Numbers and Percentages of Educational Objectives for Level of Historical Understanding

| Level of Understanding. | Middle School 9th | | | High School 10th | | |
|---|---|---|---|---|---|---|
| Unit Objective | General | Specific | All | General | Specific | All |
| Level 1 | - | | - | - | | - |
| Level 2 | 23 | 1 | 24(49%) | 27 | 5 | 32(48%) |
| Level 3 | 8 | 8 | 16(33%) | 2 | 17 | 19(29%) |
| Level 4 | - | 9 | 9(18%) | 1 | 14 | 15(23%) |
| Total | 31 | 18 | 49*(100%) | 30 | 36 | 66(100%) |

* One objective that requires improving attitude was excluded.

[32] Movement for the improvement of national capability focusing on 'Tongdo sogi (Eastern ethics and Western technology).

[33] A treaty in 1905 between the Tae-han Che-guk and Japan that the Tae-han Che-guk in which it abandoned its status as an independent state.

[34] 1) To follow the attitudes of overcoming national crisis by understanding the specific facts of a variety of movements for the national sovereignty safeguard deployed against the infringement of sovereignty by Japan. 2) To list the efforts made in order to establish a peaceful reunification after 7.4 South-North Joint Statement and to have an attitude that contributes to a peaceful reunification of our nation.

### 4.3.3. Depth of Historical Understanding

When matching each test item to a targeted objective, both the content and cognitive processes assessed in the item and in the objective were considered. If the content knowledge and cognitive processes in a test item did not match any unit objectives, it was evaluated as being *inconsistent* with objectives. It should be noted that the being *inconsistent* does not mean that the historical cognitive activity demanded in an item are not consistent with the cognitive processing in a targeted objective. Thus, the number of items that were coded as being *inconsistent* should be the same number of items that were evaluated as being *inconsistent* for the criteria 'breadth of knowledge' and 'balance of representation.' In order to explain the alignment of historical cognitive demands between test items and objectives, a coding system was developed. For example, if the cognitive demand of an item was *consistent* with the targeted objective, it was coded in the category of '*item=objective*.' Also, when the cognitive demand of an item was two levels lower or higher than the ones required in the targeted objective, it fell into the category of '*Item is 2 level<objective*' or '*Item is 2 level>objective*.' The rest of the categories for coding followed the same procedure.

#### 4.3.3.1. Depth of Historical Understanding of Middle School Test Items

1,510 questions for middle school were analyzed to determine the alignment between the historical cognitive demands of test items and of objectives, and the results are shown in Table 16. Overall, 29% of the items from the middle schools were evaluated as not aligned to any in objectives. Thus, the relationship between the historical cognitive demands of test items and objectives could not be judged. In general, the test items tended to measure a lower level of cognitive activities than the cognitive activities expressed in the objectives. Only 8% of the test

items and objectives were aligned and this occurred at levels 2 and 3. Also, there were few items that asked for a higher level of historical understanding than one required in the objectives. For example, less than 1% of the items at level 4 historical understanding assessed objectives at level 2, and 8% of the items at level 3 assessed level 2 historical understanding in objectives.

Table 16: Depth of Historical Understanding of Middle School Test Items

| Differences of Cognitive Demands | Middle School | | |
|---|---|---|---|
| | Public | Private | All |
| | n=930 | n=580 | n=1,510 |
| Item ≠ objective | 29% | 29% | 29% |
| Item level = objective level | | | |
| item level 4, objective level 4 | - | - | - |
| item level 3, objective level 3 | 3% | 2% | 2% |
| item level 2, objective level 2 | 7% | 6% | 6% |
| Item is 2 level > objective | | | |
| item level 4, objective level 2 | <1% | <1% | <1% |
| Item is 1 level > objective | | | |
| item level 4, objective level 3 | - | - | - |
| item level 3, objective level 2 | 8% | 8% | 8% |
| Item is 1 level < objective | | | |
| item level 3, objective level 4 | <1% | <1% | <1% |
| item level 2, objective level 3 | <1% | <1% | <1% |
| item level 1, objective level 2 | 39% | 45% | 42% |
| Item is 2 level < objective | | | |
| item level 2, objective level 4 | <1% | <1% | <1% |
| item level 1, objective level 3 | 10% | 8% | 9% |
| Item is 3 level < objective | | | |
| item level 1, objective level 4 | 3% | 2% | 2% |

In contrast, approximately 55% of the test items a lower level of historical understanding than was required in the objectives: approximately 43% of the test items were one level lower than the level in objectives. Among the items that were categorized one level lower than the level in objectives, it was rare when objectives were level 3 or 4, and test items were level 2 or 3,

both <1%, respectively. There were much higher percentages when objectives were level 3 and items were level 1 (9%) than when objectives were level 4 and items were level 2 (<1%). It should be noted that approximately 75% of the test items were coded as ones which assessed level 1 historical understanding while 49% of the objectives required level 2 historical understanding.

**a. Across schools**

With regard to public and private schools, the test items in both schools showed very similar results as shown in Tables 17 and 18. The proportion of the inconsistency between the level of historical understanding of test items and objectives is both 29%. Less than 10% of the test items were aligned with the level of historical thinking in targeted objectives at level 2 and 3 with a larger percentage of items at level 2 matching to objectives at level 2. Moreover, more than 50% of the items from both types of schools assessed lower levels of historical understanding than the levels of thinking in objectives. Most of these items required level 1 historical understanding regardless of whether objectives were at levels 2, 3, or 4.

There were minor differences found between the two types of schools. Public schools tended to offer a little more consistency with respect to items matching the level of historical understanding presented in objectives than private schools (approximately10% for public and 8% for private). Also, public schools provided slightly fewer test items whose historical understanding was level 1 when objectives require level 2 historical understanding (approximately 39% for public and 45% for private). Although test items were usually one level lower in historical understanding than the objectives were, the items for public schools varied across the categories more than the ones for private schools such as 'item level 2 and objective level 3' and 'item level 3 and objective level 4' rather than 'item level 1 and objective level 2.'

Table 17: Depth of Historical Understanding of Public Middle School Test Items

| Differences of Cognitive Demands | Public Middle School (n=13) | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | G | H | I | J | K | L | M | All |
| | n=68 | n=66 | n=77 | n=71 | n=66 | n=68 | n=71 | n=67 | n=68 | n=66 | n=66 | n=76 | n=100 | n=930 |
| Item ≠ objective | 28% | 27% | 43% | 31% | 30% | 18% | 17% | 25% | 21% | 17% | 26% | 34% | 45% | 29% |
| Item level = objective level | | | | | | | | | | | | | | |
| item level 4, objective level 4 | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| item level 3, objective level 3 | 4% | 3% | 1% | - | 3% | 6% | 4% | - | 3% | 12% | - | - | 1% | 3% |
| item level 2, objective level 2 | 2% | 12% | 4% | 4% | 6% | 12% | 11% | 9% | 12% | 12% | 11% | 1% | 1% | 7% |
| Item is 2 level > objective | | | | | | | | | | | | | | |
| item level 4, objective level 2 | - | - | - | - | - | - | - | - | - | 2% | - | - | - | <1% |
| Item is 1 level > objective | | | | | | | | | | | | | | |
| item level 4, objective level 3 | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| item level 3, objective level 2 | 7% | - | 3% | 13% | 9% | 7% | 20% | 15% | 13% | 12% | 6% | - | 3% | 8% |
| Item is 1 level < objective | | | | | | | | | | | | | | |
| item level 3, objective level 4 | - | - | 1% | 1% | - | - | 1% | 2% | - | - | - | - | - | <1% |
| item level 2, objective level 3 | - | 3% | 3% | 1% | 3% | - | 3% | 2% | 2% | 2% | - | - | 1% | 1% |
| item level 1, objective level 2 | 47% | 47% | 31% | 40% | 40% | 44% | 24% | 34% | 35% | 32% | 47% | 51% | 39% | 39% |
| Item is 2 level < objective | | | | | | | | | | | | | | |
| item level 2, objective level 4 | - | - | - | - | - | - | - | 1.5 | - | - | - | - | - | <1% |
| item level 1, objective level 3 | 7% | 8% | 14% | 9% | 8% | 12% | 13% | 8% | 12% | 11% | 8% | 9% | 9% | 10% |
| Item is 3 level < objective | | | | | | | | | | | | | | |
| item level 1, objective level 4 | 4% | - | - | 1% | 2% | 2% | 7% | 5% | 3% | 2% | 3% | 4% | 1% | 3% |

Table 18: Depth of Historical Understanding of Private Middle School Test Items

| Differences of Cognitive Demands | Private Middle School (n=9) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | N | O | P | Q | R | S | T | U | V | All |
| | n=66 | n=70 | n=64 | n=66 | n=39 | n=64 | n=66 | n=80 | n=65 | n=580 |
| Item ≠ objective | 24% | 31% | 30% | 20% | 15% | 42% | 29% | 31% | 31% | 29% |
| Item level = objective level | | | | | | | | | | |
| item level 4, objective level 4 | - | - | - | - | - | - | - | - | - | - |
| item level 3, objective level 3 | - | - | 3% | 2% | 3% | 3% | 2% | 5% | - | 2% |
| item level 2, objective level 2 | 2% | 6% | 13% | 12% | 3% | - | 5% | 4% | 8% | 6% |
| Item is 2 level > objective | | | | | | | | | | |
| item level 4, objective level 2 | - | | - | 2% | - | - | - | - | - | <1% |
| Item is 1 level > objective | | | | | | | | | | |
| item level 4, objective level 3 | - | - | - | - | - | - | - | - | - | - |
| item level 3, objective level 2 | 3% | 6% | 11% | 14% | 3% | 2% | 5% | 14% | 9% | 8% |
| Item is 1 level < objective | | | | | | | | | | |
| item level 3, objective level 4 | - | - | - | - | - | - | - | 1% | - | <1% |
| item level 2, objective level 3 | 2% | - | - | - | - | - | 2% | - | - | <1% |
| item level 1, objective level 2 | 59% | 47% | 34% | 39% | 64% | 39% | 49% | 38% | 46% | 45% |
| Item is 2 level < objective | | | | | | | | | | |
| item level 2, objective level 4 | 2% | - | - | - | - | - | - | - | - | <1% |
| item level 1, objective level 3 | 8% | 9% | 8% | 11% | 13% | 14% | 5% | 6% | 5% | 8% |
| Item is 3 level < objective | | | | | | | | | | |
| item level 1, objective level 4 | 2% | 1% | 2% | 2% | - | - | 6% | 1% | 2% | 2% |

Moreover, test items for public schools provided slightly higher percentages for the category, items level 1 and objective level 3' and 'items level 1 and objective level 4' than the items for private schools. Yet, these proportions between the two types of schools are two small to determine the quality of alignment between test items and objectives. Across schools, there were four schools (two from public and two from private) which had less than 5% of the items whose levels of historical understanding were consistent with or higher than the levels of historical understanding in the objectives.

#### 4.3.3.2. Depth of Historical Understanding of High School Test Items

For high school tests, 1,315 test items were coded to evaluate the relationship of historical understanding between test items and objectives. Overall, with the exception of those items inconsistent with objectives, a small relationship between the level of historical understanding of test items and that of objectives was found, as Table 19 indicates. The test items analyzed examined lower levels of historical understanding than those levels required by the objectives. Approximately thirty-eight percent of the test items were classified as inconsistent with the content required by the objectives, and thus were excluded from consideration. Only approximately 14% of the test items were consistent with the levels of historical understanding in the targeted objectives. Alignment between test items and objectives at level 2 and 3 are 10% and 4%, respectively, while the alignment between test items and objectives at level 4 is less than 1%. In addition, about 12% of the test items were at higher historical understanding levels than the objectives: most of the level 3 items were matched to an objective level 2 (11%). However, the percentages for the categories 'item level 4 and objective level 3' and 'item level 4 and objective level 2' were very low (<1%, 1%, respectively).

Moreover, approximately 35% of the test items measured a lower level of historical understanding than the objectives. Specifically, test items that assessed lower levels of historical understanding than the objectives focused on 'item level 1 and objective level 2.' However, the categories 'item level 3 and objective level 4' and 'item level 2 and objective level 3' revealed small proportions.

**a. Across schools**

Similar patterns of the relationship between items and objectives existing across public and private schools are shown in Table 19. For example, more than 35% of the items in both types of schools were determined to be inconsistent with the knowledge presented in the objectives. Also, approximately the same percentages of test items as indicated above were evaluated as ones that measured lower levels of historical understanding than the objectives. Specifically, the percentages of items coded for the 'item level 1 and objective level 2' showed the greatest proportion. Moreover, less than 16% of the items were evaluated as being consistent with the objectives at historical understanding levels 2 and 3, with the higher percentages at level 2. However, a small number of items did assess higher levels of historical understanding than the levels in the objectives (13% for both types of schools).

With regard to public and private schools, minor differences were found between the historical understanding levels of items and of objectives. The percentage of items consistent with the levels in the objectives is slightly higher for public schools than for private schools (16% and 11%, respectively). In addition, fewer test items for public schools were classified in the category 'item level 1 and objective level 2.' No items from public schools were consistent with the objectives at level 4. However, these differences were too small to make a valid comparison of the two schools. Finally, it should be noted that in four schools (two public and

Table 19: Depth of Historical Understanding of High School Test Items

| Differences of Understanding | Public High School (n=4) | | | | | Private High School (n=6) | | | | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | All | E | F | G | H | I | J | All | (n=10) |
| | n= 133 | n= 132 | n= 132 | n= 132 | n= 529 | n= 134 | n= 140 | n= 111 | n= 132 | n= 132 | n= 137 | n= 786 | n=1,315 |
| Item ≠ objective | 36% | 44% | 36% | 31% | 37% | 42% | 41% | 34% | 39% | 37% | 45% | 40% | 38% |
| Item level = objective level | | | | | | | | | | | | | |
| item level 4, objective level 4 | - | - | - | - | - | - | <1% | <1% | <1% | 2% | <1% | <1% | <1% |
| item level 3, objective level 3 | 7% | 3% | 6% | 5% | 5% | 2% | 3% | 5% | 3% | <1% | <1% | 2% | 4% |
| item level 2, objective level 2 | 14% | 5% | 6% | 17% | 11% | 5% | 16% | 12% | 8% | 2% | 7% | 9% | 10% |
| Item is 2 level > objective | | | | | | | | | | | | | |
| item level 4, objective level 2 | 2% | <1% | <1% | 2% | 1% | <1% | 2% | 2% | 2% | 3% | <1% | 2% | 1% |
| Item is 1 level > objective | | | | | | | | | | | | | |
| item level 4, objective level 3 | - | - | - | - | - | - | - | - | - | <1% | - | <1% | <1% |
| item level 3, objective level 2 | 15% | 4% | 8% | 21% | 12% | 4% | 13% | 23% | 11% | 5% | 12% | 11% | 11% |
| Item is 1 level < objective | | | | | | | | | | | | | |
| item level 3, objective level 4 | 4% | 2% | 2% | 4% | 3% | 2% | 6% | 4% | <1% | 2% | 4% | 3% | 3% |
| item level 2, objective level 3 | 2% | <1% | 2% | 2% | 2% | <1% | 6% | <1% | 2% | - | - | 2% | 2% |
| item level 1, objective level 2 | 12% | 22% | 31% | 10% | 19% | 34% | 14% | 12% | 24% | 36% | 22% | 24% | 21% |
| Item is 2 level < objective | | | | | | | | | | | | | |
| item level 2, objective level 4 | 5% | 2% | <1% | 6% | 3% | - | 1% | 4% | 3% | - | - | 1% | 2% |
| item level 1, objective level 3 | 2% | 7% | 2% | - | 3% | 5% | <1% | 2% | 2% | 7% | 4% | 3% | 3% |
| Item is 3 level < objective | | | | | | | | | | | | | |
| item level 1, objective level 4 | 2% | 11% | 5% | 2% | 5% | 5% | <1% | 2% | 4% | 5% | 4% | 3% | 4% |

private) more than 30% of the test items measured historical understanding levels consistent with or higher than the levels in the objectives (45%, 41%, 38%, and 34%, respectively).

#### 4.3.3.3. Differences of Depth of Historical Understanding between Middle and High School Test Items

In general, differences across grades were moderately small. As Table 20 presents, when considering alignment between the level of historical understanding and the content presented in test items and objectives, approximately 10% more were inconsistent for high schools than

Table 20: Comparison between the Depth of Understanding of Middle and High School Test Items

| Differences of Understanding | Middle School 9th | High School 10th |
|---|---|---|
| | n=1,510 | n=1,315 |
| Item ≠ objective | 29% | 39% |
| Item level = objective level | | |
| item level 4, objective level 4 | - | <1% |
| item level 3, objective level 3 | 2% | 4% |
| item level 2, objective level 2 | 6% | 10% |
| Item is 2 level > objective | | |
| item level 4, objective level 2 | <1% | 1% |
| Item is 1 level > objective | | |
| item level 4, objective level 3 | - | <1% |
| item level 3, objective level 2 | 8% | 11% |
| Item is 1 level < objective | | |
| item level 3, objective level 4 | <1% | 3% |
| item level 2, objective level 3 | <1% | 2% |
| item level 1, objective level 2 | 42% | 21% |
| Item is 2 level < objective | | |
| item level 2, objective level 4 | <1% | 2% |
| item level 1, objective level 3 | 9% | 3% |
| Item is 3 level < objective | | |
| item level 1, objective level 4 | 2% | 4% |

middle schools. However, these were a higher degree of alignment for high schools than middle schools. Moreover, the test items for high schools varied more across categories, indicating that high school tests tend to measure a deeper understanding of history than middle schools tests.

### 4.4. Results for Breadth of Knowledge and Balance of Representation

Test items were also analyzed to examine how well balanced historical knowledge measured by tests from both middle and high schools was across unit objectives, and to determine which type of knowledge received more emphasis. All items were classified one by one with respect to their alignment with the content of unit objectives. It should be noted that test items were not coded more than once if they were classified for a targeted category. In addition, objectives may have been divided into two or more categories, depending on their targeted categories. For example, the objective 'causes, processes, and effects of Oaeran,[35]' was divided into three categories–the 'causes,' 'processes,' and 'effects' of the war. In the case of this objective, one test item could not measure all the content of the war required by the unit objective.

Labels were assigned to test items for the objectives *consistent*, *inconsistent*, *broad*, and *narrow*. The term *consistent* means that the content of a test item satisfied the knowledge required in a unit objective. The following item is an example from this category:

> Q: The following table presents a comparison between the features of the Koryō and the Chosōn societies. Which fact is incorrect?
>
> | | Koryō[36] | Chosōn[37] |
> |---|---|---|
> | A | Munbol[38] aristocratic society | Yangban[39] bureaucratic society |

[35] The war between the Chosōn and Japan from 1592 to 1598.
[36] A Korean dynasty in the medieval era established in 918 and destroyed in 1392 by Yi Song-gye.
[37] A Korean dynasty established by Yi Song-gye in 1392 and destroyed by Japan in 1910.

| | | |
|---|---|---|
| B | Monopoly of power by the privileged | Wide participation in politics by bureaucrats |
| C | Centralized governing state | Localized governing state |
| D | Caste society | Merit-oriented society |
| E | Importance of Buddhism | Importance of Confucianism |

This item was coded as being *consistent* with the objective 'to explain the differences between Koryō and the Chosōn societies' on the unit "The Establishment and Development of the Chosōn Dynasty" for 9th grade middle school. This item assessed the differences between the two societies in terms of the central power of each society, the features of medieval or modern society, and an ideological background for the ruling class.

A test item labeled *broad* meant that the question covered more knowledge than the content emphasized in a targeted objective**.** The following item is an example of a broad test item.

> Q: Which answer illustrates the commonality between the Koryō Kwang-jong[40] and the Chosōn Tae-jong[41]?
>
> Koryō Kwang-jong
> - Institutionalizing the civil-service examination
> - Confiscating private slaves owned by the aristocracy
> - Eliminating meritorious retainers and local warlords
> - Replacing aristocracy with officials appointed by the throne
>
> Chosōn Tae-jong
> - Implementing Yukjo Gikgye system[42]
> - Strengthening relations with Ming China
> - Prohibiting a private army
> - Eliminating meritorious retainers at the founding of the dynasty
>
> A. Anchored the political order based on Confucianism
> B. Held fast to the independent stand toward China
> C. Established an appropriate balance between royal authority and cabinet power

[38] Hereditary literati class based on a strong family background.
[39] The ruling class of the Chosōn recruited by the civil-service examination known as the kwagō.
[40] The fourth king of the Koryō (925-975), a son of T'aejo who established the Koryō Dynasty.
[41] The third king of the Chosōn, a son of Yi Songgye who established the Chosōn Dynasty.
[42] A direct control system ruled by the Six Departments (Yukjo), independent from the Office of State Councilors (Uijōngbu, the supreme administrative structure).

D. Succeeded in integrating the resistant power in a political system
E. Tried to secure governing stability from a discontented power in a system

This item was coded as *broad* for the objective 'centralized policies of the Chosōn Dynasty' in the unit "Administrative Structure and Political Activities" for 10th grade high school because it assessed centralized policies not only for the Chosōn but also for the Koryō Dynasty. Kwang-jong, the king of the Koryō, was a son of Wang-gun, who founded the new dynasty: he tried to strengthen royal authority by implementing a new government employment system called Kwagō (civil-service examination). It was a means to centralize power by creating a new capital bureaucracy, and by recruiting officials based on a merit rather than on family background. In addition, by eliminating meritorious retainers, both kings intended to promote a strong monarchy. Because this unit does not require an understanding of the centralized political policies of the Koryō era, this item was classified as *broad* only for the Chosōn era.

A *narrow* categorization means that a test item assessed only some elements of the knowledge presented in the objective. The following item is an example:

> Q: The following is part of the content of 'Chosōn Chaek-ryak,[43]' a book brought from Japan. What was the direct effect of the spread of knowledge from this book?
>
> "In order to prevent an invasion by Russia, the Chosōn should strengthen itself by implementing a foreign policy of 'intimate relations with China, coordination with Japan, and alliance with America' and adopt Western culture and institutions."
>
> A. The Tianjin Treaty[44]
> B. The Jemulpo Treaty[45]

[43] A booklet entitled Chaoxian celue (A Strategy of Chosōn) that was delivered by a Chinese diplomat, Huang Junxien, in Tokyo in 1880.
[44] A treaty in 1885 between Ching (Qing) China and Japan after the intervention of the Kab-shin Chung-pyun (Coup d'etat) in 1884 carried by Korean reformists. The two countries agreed: a) to pull their expeditionary forces out of Korea simultaneously; b) not to send military instructors for the training of the Korea army; and c) notify the other side beforehand should one decide to send troops to Korea.
[45] A treaty between Chosōn and Japan as the results of the Imo-Military Riot in 1882. Through this treaty, Japan received indeminities from Chosōn and the right to station its troops in Chosōn.

C. The proliferation of Tōng-hak[46]
D. The aggravation of Wi-jōng Chōk-sa[47] Movements
E. The outbreak of the Imo Military Riot[48]

This item was classified as *narrow* for the objective 'differences and meanings of Kae-wha[49] and Wi-jōng Chōk-sa Movements' for the unit "The Enlightenment and Independence Movement" for 9th grade middle school. It focuses on the content of the Kae-wha movement, the main point of the book *Chosōn Chaek-ryak*, which influenced Kae-wha reformists. Students should choose D as an answer, but this item does not assess knowledge of the significant meanings of *Chosōn Chaek-ryak* as part of the Kae-wha movement or the meanings of Wi-jōng Chōk-sa Movements, nor does it require students to compare differences between the two.

A test item coded as *inconsistent* means that the item was not related to the knowledge required in the unit objectives. However, if the content assessed related to the knowledge of a unit but not the unit objectives, it was still coded as *inconsistent* with the knowledge for that unit. For example, an item that asked about 'a historical figure who participated in the movement for the revival of Paek-che[50]' was classified as being inconsistent with the objective for the unit "Administrative Structure and Political Activities" for 10th grade high school. Although this unit does not focus on the Paek-che revival movements against the unification of the Korean peninsula by Shilla in the 7th century, the item is related to this unit. Thus, test items that did not assess the content of any of the objectives were coded *inconsistent* for a certain unit.

[46] Eastern Learning established by Choe Che-u in 1860 through a mixture of traditional elements from Confucianism, Buddhism, and Son-gyo (teachings of Hwarang in Shilla).
[47] Conservative movement led by rustic literati (Neo-Confucians), protecting Confucian social rules and opposing foreign power (including Japanese economic invasion) and the proliferation of Christianity.
[48] An army rebellion in 1882 by the old army against a military reformation under Japanese training, matching the Japanese idea of 'rich nation and strong military.'
[49] Enlightenment/progressive movement that focused on 'tongdo sogi' (Eastern ethics and Western technology).
[50] One of three kingdoms that was prominent in the southwest of Korean peninsula and existed up to the 7th C.

### 4.4.1. Historical Knowledge of Test Items for 9th Grade

As mentioned earlier, 9th grade middle school Korean history consists of six units which focus on understanding the political activities of the Chosōn era as they relate to contemporary society. To clearly illustrate the results of test item coding, this section includes descriptions for each unit concerning the distribution of test items along unit objectives, the knowledge stressed by test items, and the difference in results between public and private schools.

#### 4.4.1.1. Test Items for the Unit "The Establishment of the Chosōn Dynasty and Its Development"

For the unit "The Establishment and Development of the Chosōn Dynasty," 339 test items were analyzed in order to determine to what degree the items tested the knowledge required in the unit objectives. Table 21 reflects the percentages and numbers of total items related to this unit, regardless of whether the test items were consistent or inconsistent with objectives. As indicated in Table 21**,** this unit provides six general objectives and three specific objectives. To provide greater clarity, the four objectives were divided into at least two, and up to six categories (two for general and two for specific). For example, the objective 'increase of the Sarim[51] power and political changes in the middle of the Chosōn dynasty' was divided into two categories. The objective 'causes, processes, and effects of Oaeran and Horan[52]' were treated as six categories--'causes', 'processes,' and 'effects' of each war (Oaeran and Horan). Thus, a test item that assessed only causes, processes, or effects was not coded as *narrow* because the knowledge required in the original category was too broad to assess it with one test item. Objectives with

[51] New scholar officials known as rustic literati that had emerged from the countryside. They were faced with a series of literati purges caused by the conservative capital yangban (ruling group) in 16th C, but finally intensified their political power in the middle of the Chosōn era.
[52] Two wars between the Chosōn and the Kem (1627) and the Chosōn and the Chin (1636).

more than one subject, such as 'differences between the Chosōn and the Koryō societies and Chosōn foreign policies toward Japan and Yōjin' and 'conversion of foreign policy to Chin-myōung Pae-kem[53] after In-jo[54]'s coup d'etat and politic situations in East-Asia after two wars' were each assigned two separate objectives. The objective 'differences between the Chosōn and the Koryō societies' was excluded because it contains the same content as the general objective.

As mentioned earlier, if an item measured part of the content or more than the content of the knowledge demanded in the unit, it was regarded as *narrow* or *broad*, respectively. With relation to the objective 'differences between the Koryō and the Chosōn societies,' a test item that assessed only the political and ideological characteristics of the Chosōn was coded as *narrow* for that objective. In addition, a test item that assessed the features of the transportation, communication, and taxation systems fell into the category of *narrow* for the objective 'central and local political and educational systems in the Chosōn society.' The systems addressed in each item were also used as a means by which the state tried to establish a centralized government along with the political and educational systems. If a test item asked only for the names of political struggles, it was treated as *narrow* for the objective 'political changes during the Chosōn period.' The coding assigned to each test item was based on the official descriptions of the history curriculum.

**a. Across all items**

Overall, as Table 21 indicates, approximately two-thirds of the test items assessed knowledge consistent or at least somewhat consistent with the content required in the unit objectives, while one-third of the items were considered to be inconsistent with the knowledge of the unit. However, with regard to the alignment between the content of items and objectives,

[53] A policy that favored Ming China and opposed Kem (Chin)
[54] The 12th King of the Chosōn who was pro-Ming China and con-Chin.

36% of the test items assessed the content emphasized in the unit, and 30% of the test items focused somewhat on knowledge included in the unit. In addition, the items for this unit focused heavily on assessing the knowledge of the general objectives: only nine of the 339 items were coded for specific objectives. Across the objectives, test items were distributed mostly along the objectives dealing with the content of the new governing structures, the new political power of the Chosōn, and the facts of the political struggles after Sarim power was established, 23%, 9%, and 7%, respectively. In addition, knowledge of the war, Oaeran, assessed by the test items received greater emphasis over other knowledge (9%).

**b. Across schools**

The test items from both public and private schools showed a very similar pattern. Regarding the distribution of items consistent with objectives, knowledge measured did not vary across the objectives. The test items of both public and private schools tended to measure the governing structures of new dynasties such as political, educational, or military systems (13% and 11%, respectively), the results of the two wars (8%=5%+3% and 9%=7%+2%, respectively), and the establishment of new Confucian power (5% for both types of schools). However, in both types of schools, these items lacked assessment of the understanding of the ideological features of the Chosōn society by asking for the comparison of it to the Koryō (both <1%); neither did they assess knowledge of the establishment of a national identity (both <1%), such as through the creation of the Korean alphabet, Hangul. Moreover, these items did not seem to take into consideration the causes of the two wars: only three items for each school focused on causes. In particular, the test items did not assess the Chosōn society and its dynamic international relationships with Japan and Chin (Kem) in the late 16$^{th}$ C and 17$^{th}$ C. In addition, the items did not assess the relationship between political changes due to establishment of Confucian powers

Table 21 Frequency of Test Items for the Unit "Establishment and Development of the Chosōn Dynasty

| Knowledge in Objectives | Consistent | | Broad | | Narrow | | All** |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| All items | Public n=215 | | | Private n=124 | | | 339(100%) |
| Inconsistent items | Public n=64(30%) | | | Private n=44(36%) | | | 108(32%) |
| ***General Objectives*** | | | | | | | |
| Differences between the Koryō and the Chosōn societies | 1(<1%) | 1(<1%) | - | - | 5(2%) | 5(4%) | 12(4%) |
| Central and local political and educational systems in the Chosōn society | 27(13%) | 14(11%) | 3(1%) | 1(<1%) | 23(11%) | 9(7%) | 77(23%) |
| Growth of national identity in the early Chosōn period | 2(1<%) | - | - | - | - | - | 2(<1%) |
| Increase of Sarim power* | 10(5%) | 6(5%) | 1(<1%) | - | 7(3%) | 7(6%) | 31(9%) |
| Political changes caused by Sarim power* | 6(3%) | 2(2%) | - | - | 12(6%) | 3(2%) | 23(7%) |
| Contents and meanings of Chosōn foreign policy toward the neighboring countries of Ming China, Yōjin[55], and Japan | 5(2%) | 1(<1%) | - | - | 4(2%) | 2(2%) | 12(4%) |
| Causes of Oaeran* | 2(<1%) | 2(2%) | - | - | - | - | 4(1%) |
| Processes of Oaeran* | 3(1%) | 2(2%) | - | - | - | - | 5(2%) |
| Effects of Oaeran* | 11(5%) | 9(7%) | - | - | - | - | 20(6%) |
| Causes of Horan* | 1(<1%) | 1(<1%) | - | - | 13(6%) | 8(7%) | 23(6%) |
| Processes of Horan* | 1(<1%) | - | - | 1(<1%) | - | - | 2(<1%) |
| Effects of Horan* | 6(3%) | 3(2%) | - | - | 2(<1%) | - | 11(3%) |
| ***Specific Objectives*** | | | | | | | |
| Chosōn foreing policies toward Japan and Yōjin* | - | - | - | - | - | - | - |
| Political meaning of publishing Kyong-guk Tae-jeon | 1(<1%) | 1(<1%) | - | - | 1(<1%) | - | 3(<1%) |
| Conversion of foreign policy to Chin-myōng Pae-kem after Injo's coup d'etat* | 1(<1%) | 1(<1%) | 1(<1%) | - | - | - | 3(<1%) |
| Political situations in East-Asia after the two wars* | 2(<1%) | 1(<1%) | - | - | - | - | 3(<1%) |
| Total | 79(29%) | 44(36%) | 5(2%) | 2(2%) | 67(31%) | 34(27%) | |
| Public and Private Total | 123(36%) | | 7(2%) | | 101(30%) | | 231(68%) |

*: Divided objectives

**: All reflects *consistent, broad*, and *narrow* for both public and private.

[55] The tribal state that existed in the north-east side of Manchuria. The Jurchen Manchu created a Chin Empire lasted until 1234. It rose again and strengthened its hegemony in Manchuria and northern China (Kem, 1627) and Chin (1636) in Chinese territory. The establishment of Manchu kingdom in China was a big threat to the Korean state.

and the publishing of Kyoung-guk Tae-jeon[56] as the definer of a completive Confucian state. It should be noted, however, that in public schools the relatively high percentages for the *narrow* category were assigned to several objectives, including 'the systems for the Chosōn society', 'the causes of Horan,' 'political changes of the Chosōn', and 'increase of the Sarim power' because the items for these categories assessed knowledge only partly related to the content in the objectives.

#### 4.4.1.2. Test Items for the Unit "The Changes in the Chosōn Society"

Regarding the unit "The Changes in the Chosōn Society," 241 test items were coded by their content emphasis. As shown in Table 22, the 7th National Curriculum provided the content that has to be emphasized as well as the unit objectives, which consisted of four general objectives and two specific ones. Of the six objectives, the two were each divided into two: 1) 'features of factional politics' and 'its positive and negative influences'; and 2) 'social disorder and taxation corruptions in the Sedo government[57] and 'various peasant resistances' and 'the impact of Tōng-hak[58] and Catholicism on the peasant society under the Sedo government.' These two objectives were too difficult to analyze because the test items did not fit into these broad content assessments.

For the objective 'backgrounds and intentions of Yōng-jo implementing the Tang-

[56] Grand Code for State Administration: A written form of constitutional law of the Chosōn Dynasty in order to actualize the monarchial system based on the Confucian-ideal government.
[57] The administration that families of in-laws to the throne, part of the Patriarch literati dominated the Chosōn court from 1800 to 1863, especially a family known as the Andong Kims.
[58] Means Eastern Learning established by Choe Che-u in 1860 through a mixture of traditional elements from Confucianism, Buddhism, and Son-gyo (teachings of Hwarang in Shilla)

pyōng[59] (impartiality) policy,' a test item that assessed either background or intention was treated as *narrow*. In particular, a test item that assessed only the content of the scholarly features of Shil-hak[60] and did not relate them to the reformative ideas proposed by Shil-hak scholars was coded as *narrow*. In addition, a test item that measured the understanding of Shil-hak as either physiocrats or mercantilists without comparison between the two factions of Shil-hak scholars was also classified as *narrow* for the objective 'comparison between the two factions of Shil-hak scholars.' If a test item asked for fairly detailed information, and directly related to the content presented in an objective, the test item was regarded as *consistent* with the objective. Again, once a category was chosen for a test item, the test item was not assigned to any other category or content area.

**a. Across all items**

The overall percentages and numbers for the unit "The Changes in the Chosōn Society," shown in Table 22 reflect the percentages and numbers of total items related to this unit. In general, of the 241 test items, 187 items (78%) were coded as being at least somewhat aligned with objectives, and 54 (22%) were coded as *inconsistent* with objectives. Of the 187 test items for this unit, almost half of them were treated as *narrow* (49%) and only one fourth of the items were classified as testing the content required by the unit objectives.

As Table 22 illustrates, the highest percentage of test items was related to the knowledge of 'social situations and scholarly dispositions of Shil-hak' (24%). However, a relatively large percents were considered to be *narrow*. The percentage for 'social situations under the Sedo government and peasants resistances' was higher than for the remaining categories (19%).

[59] An impartial policy implemented in 18th C by two kings, Yōng-jo and Chōng-jo for resolving the factional fights among political groups, but without much success.
[60] Pragmatic studies developed by the off-court scholars since the 17th C. These scholars urged practical reforms of Confucian state-craft and the established politics based on the orthodox Neo-Confucianism.

Table 22: Frequency of Test Items for the Unit "The Changes in the Chosōn Society"

| Knowledge in Objectives | Consistent | | Broad | | Narrow | | All** |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| All items | Public n=170 | | | Private n=71 | | | 241(100%) |
| Inconsistent items | Public n=40 (24%) | | | Private n=14 (20%) | | | 54(22%) |
| ***General Objectives*** | | | | | | | |
| Features of factional politics[61]* | 9(5%) | 6(9%) | 2(1%) | 1(1%) | 1(<1%) | 3(4%) | 22(9%) |
| Positive and negative influences of factional politics* | 5(3%) | 1(1%) | - | - | 2(1%) | 1(1%) | 9(4%) |
| Background and intentions of Yōng-jo implementing the Tang-pyōng (impartiality) policy. | 4(2%) | 1(1%) | 1(<1%) | - | 8(5%) | 3(4%) | 17(7%) |
| Social situations and the scholarly dispositions of Shil-hak | 1(<1%) | 1(1%) | - | - | 39(23%) | 17(24%) | 58(24%) |
| Social disorder and taxation corruptions in the Sedo government and various peasants resistances* | 19(11%) | 8(11%) | 1(<1%) | 4(6%) | 8(5%) | 5(7%) | 45(19%) |
| Impact of Tōng-hak and Catholicism on peasant society under the Sedo government* | - | 1(1%) | - | - | 2(1%) | - | 3(1%) |
| ***Specific Objectives*** | | | | | | | |
| Comparison between physiocrats[62] and mercantilists | 1(<1%) | - | - | - | 16(9%) | 3(4%) | 20(8%) |
| Political and social background of the diffusion of Catholicism and Tōng-hak | 2(1%) | - | - | - | 9(5%) | 2(3%) | 13(5%) |
| Total | 41(24%) | 18(25%) | 4(2%) | 5(7%) | 85(50%) | 34(48%) | |
| Public and Private Total | 59(25%) | | 9(4%) | | 119(49%) | | 187(78%) |

*: Divided objectives

**: All reflects *consistent, broad*, and *narrow* for public and private.

[61] Factional fights among political groups: Sarim power from a younger and an elder group of Confucian scholars called Tong-in (East Faction), So-in (West Faction), Nam-in (South Faction), and Pug-in (North Faction).

[62] Shil-hak scholars who attempted agricultural reforms in the 18th C.

However, very few test items were classified as ones that assessed the understanding of the impact of Tōng-hak and Catholicism on a peasant society under the Sedo government.

**b. Across schools**

Table 22 also shows public and private school test items that emphasized content in the unit objectives. With regard to the frequency of those test items consistent with objectives, both school types showed very similar results, 24% for public schools and 25% for private. Across objectives, test items were unevenly distributed, and focused heavily on a single objective such as 'the Sedo government and peasants resistances.' In contrast, only a few items from both school types were classified for 'impact of Tōng-hak and Catholicism,' 'comparison between physiocrats and mercantilists,' and 'backgrounds of the diffusion of Catholicism and Tōng-hak.' Similar percentages across schools existed for test items that assessed content narrower than objectives (50% for public and 48% for private schools). Test item percentages from both schools for 'social situations and scholarly dispositions of Shil-hak' were the highest across content, 23% and 24% respectively. It should be noted that a number of items for 'social situations and scholarly dispositions of Shil-hak' were classified as *narrow* because they assessed the names of scholars or their products, without relating them to political or social problems or to why they tried to reform society. In this unit, test items were likely to assess simple facts of the meanings or dispositions of Shil-hak without asking about either the connection between the current political and social circumstances of the time or the reformative studies suggested by Shil-hak scholars and the diffusion of new religions through society.

#### 4.4.1.3. Test Items for the Unit "The Enlightenment and Independence Movement"

Test items were analyzed for the unit "The Enlightenment and Independence Movement." To accomplish the goals for this unit, the 7th National Curriculum provides six general objectives and three specific ones, and emphasizes the understanding of how the Chosōn Dynasty tried to overcome both its national and international crises from its position at the end of the 19th century (MOE, 1999). Table 23 shows test items analyzed for those objectives emphasized for this unit. One broad objective 'background and significance of the Tōng-hak Peasants Movement[63]' was divided into two: 'background' and 'significance' of the Tōng-hak Peasants Movement. However, the objective was also maintained in its original form because some of the test items assessed the two areas simultaneously within one test item. In this case, test items were coded no more than once for the targeted objectives. If a test item required an answer only about the content of reform policies by Hung-sōn Tae-won'gun[64] or the content of the Kae-wha[65] and Wi-jōng Chōk-sa[66] Movements, it was classified as *narrow* for a targeted objective. These items did not cover all of the content required in the objectives, such as the differences in or significance of the movements or the historical meanings of the reform polices. In addition, when an item tested only knowledge of Kabō Reform, it was categorized as *narrow* for a targeted objective. It did not assess the meanings and limitations of the reform.

**a. Across all items**

As can be seen in Table 23, of the 300 test items, a little more than half were coded as *consistent* or somewhat consistent; the remaining items were classified as *inconsistent* with

[63] A revolt of armed peasants in 1894 against corrupt magistrates and a foreign economic power, Japan.
[64] Father of King Ko-jong, and a regent who ruled the Chosōn Dynasty from1863 to 1873.
[65] Enlightenment/progressive movement that focused on 'tongdo sogi' (Eastern ethics and Western technology).
[66] Conservative movement led by rustic literati (Neo-Confucians), protecting Confucian social rules and opposing the Japanese economic invasion and the proliferation of Christianity.

objectives, 54% and 46%, respectively. Overall, the percentage for 'reform policies of Hung-sōn Tae-won'gun' (16%) was the highest, if we include those items whose contents were somewhat related to the objective. In addition, the percentage of items for the 'Kabō[67] Reform' is relatively higher than those for other objectives. However, the overall percentages of the above items included content they were somewhat related to the targeted objectives. By contrast, only a few items were coded for 'significance of Tōng-hak Peasant Movement' and 'characteristics of the Tōng-hak Peasant Movement and the Kabō Reform.'

Across the categories *consistent, broad,* and *narrow*, the percentage of items treated as *narrow* is higher than the items treated as *consistent.* Of the 137 items, 66 items (22%) were classified as being *consistent* with objectives, two (<1%) were classified as being *broad*(er) than objectives, and 95 items (32%) were classified as being *narrow*(er) than objectives.

**b. Across schools**

Table 23 also illustrates the items that were coded for targeted objectives, for both public and private school. The analysis for both types of schools shows somewhat similar results: as mentioned above, the percentages of items that were coded as *consistent* are lower than the ones that were coded as *narrow*. Specifically, the percentages of items that were consistent with 'purposes and meanings of reform policies by Hung-sōn Tae-won'gun' were much lower than those labeled as narrower than the objective. Most of the items for this objective assessed the understanding of the reformative contents rather than its meanings and purposes. The objective 'differences and meanings of Kae-wha and Wi-jōng Chōk-sa Movements' was not assessed by either of the school types, but some elements of the objective were assessed as *narrow*. With regard to the differences, in both types of schools, between the percentages for being *consistent*,

[67] A modern reform in 1894 after the Tōng-hak peasant uprising, which was widely implemented through political, economic, and social areas.

Table 23: Frequency of Test Items for the Unit "The Enlightenment and Independence Movement"

| Knowledge in Objectives | Consistent | | Broad | | Narrow | | All** |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| All items | Public n=187 | | | Private n=113 | | | 300(100%) |
| Inconsistent items | Public n=81(43%) | | | Private n=56(50%) | | | 137(46%) |
| ***General Objectives*** | | | | | | | |
| Purposes and meanings of reform policies by Hung-sōn Tae-won'gun | 9(5%) | 3(3%) | - | - | 26(14%) | 10(9%) | 48(16%) |
| Characteristics and meanings of the Kanghwa-do Treaty[68] | 7(4%) | 6(5%) | - | - | 1(<1%) | 1(<1%) | 15(5%) |
| Differences between and meanings of Kae-wha and Wi-jōng Chōk-sa Movements | - | - | - | - | 12(6%) | 5(4%) | 17(6%) |
| Characteristics of Kae-wha faction and the reformative purposes of the Kab-shin Chōng-pyōn[69] | 1(<1%) | 2(2%) | - | - | 4(2%) | 3(3%) | 10(3%) |
| Background and significance of Tōng-hak Peasants Movement | 2(1%) | 3(3%) | 2(1%) | - | 6(3%) | 7(6%) | 20(7%) |
| Background of Tōng-hak Peasants Movement* | 5(3%) | 3(3%) | - | - | - | - | 8(3%) |
| Significance of Tōng-hak Peasants Movement* | 2(1%) | - | - | - | - | - | 2(<1%) |
| Significance and limitations of the Kabō Reform | 6(3%) | 3(3%) | - | - | 10(5%) | 7(6%) | 26(9%) |
| ***Specific Objectives*** | | | | | | | |
| People's reactions to the Isolationist foreign policy (Shōae-guk Chōung-chaek) by Hung-sōn Tae-won'gun | - | - | - | - | - | - | - |
| Content of the Kanghwa-do Treaty | 10(5%) | 1(<1%) | - | - | - | 1(<1%) | 12(4%) |
| Characteristics of the Tōng-hak Peasant Movement and the Kabō Reform | 3(2%) | - | - | - | - | 2(2%) | 5(2%) |
| Total | 45(24%) | 21(19%) | 2(1%) | - | 59(32%) | 36(32%) | |
| Public and Private Total | 66(22%) | | 2(<1%) | | 95(32%) | | 163(54%) |

*: Divided objectives

**: All reflects *consistent*, *broad,* and *narrow* for public and private.

[68] Korea's first modern treaty with a foreign country (Japan) in 1876, much to Korea's disadvantage. After that, Japan monopolized the Korean market.

[69] Coup d'etat in 1884 carried out by reformists Kim Ok-kyun, Hong Yong-shik, and So Chae-p'il.

those for private schools are somewhat balanced across the general objectives, but only one test item the content in specific objectives. Test items from public schools focused more on 'the Kanghwa-do Treaty' and 'reform policies by Hung-sōn Tae-won'gun.' Compared to private schools, test items from public schools assessed more of the content emphasized in specific objectives. However, overall, the distribution of knowledge assessed did not vary across the objectives.

#### 4.4.1.4. Test Items for the Unit "The Deployment of Movement for National Sovereignty Safeguard"

For the unit "The Deployment of Movement for Sovereignty Safeguard," 307 test items were coded according to their assessment of the unit objectives and the degree to which content was assessed. The eight objectives for this category consist of five general objectives and three specific ones. Of these five general objectives, one objective 'the establishment of the Tae-han Che-guk[70] and the purposes and results of the Gwang-mu Reform'was divided into two: 'the establishment of the Tae-han Che-guk[71]' and 'the purposes and results of the Gwang-mu Reform.' For most test items, this objective would be too broad to include all of the required content.

As with the other units, the test items for this unit were classified according to whether they: included all of the content that the objectives require; assessed one or two more elements than the objectives; or were related only to the content area or assessed fewer elements than the content in objectives. For example, if an item did not connect the purposes of Tōng-nip Hyōp-

[70] The Great Han (Tae-han) Empire (Che-guk) proclaimed by King Kojong in 1897 to the nation and the world, the establishment of an independent nation.
[71] The Great Han (Tae-han) Empire (Che-guk) proclaimed by King Kojong in 1897 to the nation and the world, the establishment of an independent nation.

hoe to the activities of Man-min Kong- dong-hoe,[72] assessing only a part of content of the objective, the item would be treated as *narrow*. If an item asked about both the Tae-han Che-guk and the Gwang-mu Reform, it would be treated as *broad*. In this case, the item should satisfy the content required by the targeted objective.

**a. Across all items**

As Table 24 indicates, approximately two-thirds of the test items were consistent or somewhat related to the objectives of this unit, and one-third of the test items were not related to any objectives in the unit (70% and 30%, respectively). In addition, the majority of the test items fell into the category of satisfying the content emphasized in unit objectives only when we add together the percentage for both consistent and broad items (50%=46%+4%). In general, the distribution of the test items for this unit was severely unbalanced among the objectives, and most of the test items concentrated on measuring knowledge of unit objectives but not specific objectives. Across the objectives, 41% (n=125) of the items were related to knowledge of the movement for national sovereignty safeguard against Japan at the end of the 19th century. However, only one item out of 307 assessed an understanding of the objective of the 'Shil-ryok Yang-song Movement.'

**b. Across schools**

Table 24 also provides the frequency of test items for both public and private schools in terms of the targeted objectives and knowledge emphasized. Across schools, the percentages of test items across objectives are very similar, with the distribution of knowledge assessed being severely unbalanced across objectives. As an example, with respect to the items that assessed

[72] A mass assembly known as the Joint Meeting of Government and People in 1898 led by the Independence Club, attended by a variety of people--students, women, Buddhist monks, high-level of ministers, literati, and so on, to propose the transformation of the government to a modern legislative body.

the content in objectives, both types of schools focused more on general objectives, and on three in particular, than on specific objectives. These two types of schools showed a similarity in much higher percentages of items consistent with objectives than those percentages narrower than objectives. In considering the label *consistent*, the test items for this unit tended to assess the names of people or the activities of people and organizations involved in the national independence from foreign power. However, the items lacked assessment of the background, intention, or limitations of the political organizations and ruling class for national independence from foreign powers. For example, percentages for both public and private schools were the highest with respect to items assessing the understanding of the movement as a national sovereignty safeguard against infringement by Japan. The items for this objective involved a number of names, places, and activities for national independence. In addition, the percentages of items for 'international situations that the Chosōn faced after the war between Russia and Japan' and 'the purposes of Tōng-nip Hyōp-hoe towards the activities of Man-min Kong- dong-hoe' were relatively higher than others. In contrast, only a few items were related to knowledge of the establishment of the Tae-han Che-guk and the similar characteristics of Ui-byong[73] movement and Yaeguk-gyemong[74] movement. Moreover, the background and intention of the enlightenment movement by Tong-nip Hyop-hoe and the limitations of Confucian Ui-byong were not assessed.

[73] The righteous Army that was organized during the Japanese political intervention of Korea and intensified in 1908 after Japan (p.122).
[74] Patriotic enlightenment movements led by the intelligentsia to enlighten the public through education.

Table 24: Frequency of Test Items for the Unit "The Deployment of Movement for National Sovereignty Safeguard"

| Knowledge in Objectives | Consistent | | Broad | | Narrow | | All** |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| All items | Public n=179 | | | Private n=128 | | | 307(100%) |
| Inconsistent items | Public n=54 (30%) | | | Private n=39 (30%) | | | 93(30%) |
| ***General Objectives*** | | | | | | | |
| Purposes of Tōng-nip Hyōp-hoe and the activities of Man-min Kong-dong-hoe members | 10(6%) | 5(4%) | - | - | 7(4%) | 8(6%) | 30(10%) |
| Establishment of the Tae-han Che-guk and its significance* | - | - | 1(<1%) | 2(2%) | - | 1(<1%) | 3(1%) |
| Purposes and results of the Gwang-mu Reform* | 1(<1%) | 1(<1%) | 1(<1%) | 1(<1%) | 8(5%) | 3(2%) | 15(5%) |
| International situations that the Chosōn faced after the war between Russia and Japan[75] | 10(6%) | 6(5%) | - | - | 9(5%) | 9(7%) | 34(11%) |
| A variety of movements for the national sovereignty safeguard deployed against the infringement of sovereignty by Japan | 61(34%) | 45(35%) | 6(3%) | 2(2%) | 10(6%) | 1(<1%) | 125(41%) |
| Similar characteristics between Ui-byong and Yaeguk-gyemong Movements | - | 2(2%) | - | - | 1(<1%) | 2(2%) | 5(2%) |
| ***Specific Objectives*** | | | | | | | |
| Background and intention behind Tōng-nip Hyop-hoe promoting the enlightenment movement to the people | - | - | - | - | - | - | - |
| Limitations of Confucians who led Ui-byong against Japan at the end of the Chosōn Dynasty | - | - | - | - | - | - | - |
| Significance of the Shil-ryok Yang-song Movement[76] after the Elsa Treaty[77] | - | 1(<1%) | - | - | - | - | 1(<1%) |
| Total | 82(46%) | 60(47%) | 8(5%) | 5(4%) | 35(20%) | 24(19%) | |
| Public and Private Total | 142(46%) | | 13(4%) | | 59(19%) | | 214(70%) |

*: Divided objectives

**: All reflects *consistent, broad*, and *narrow* for public and private.

[75] A war between the two nations from 1904 to 1905. Russia and Japan, among other foreign powers, were the most aggressive in expanding their economic interests in Korea.

[76] Movement for the improvement of national capability over foreign powers.

[77] A 1905 treaty between the Tae-han Che-guk and Japan in which the Dae-han Che-guk abandoned its status as an independent state.

#### 4.4.1.5. Test Items for the Unit "The National Independence Movement"

The unit "The National Independence Movement" requires students to understand the deployment of colonial policies by Japan, Korea's struggles for national sovereignty, and the relationship between politics and those struggles. For this unit, the National Curriculum provided six general objectives and four specific objectives, as shown in Table 25. One of the general objectives was excluded because it was aimed to improving patriotic attitude by studying people who fought for national sovereignty during the Japanese colonization (MOEHRD, 1997, p. 317). Two objectives were divided into separate parts: 'Japanese colonial policies during each period and background and reason for changes in the colonial policies' was divided into two; and 'background, processes, effects, and significance of the 3.1 Movement[78]' was divided into four. However, considering the 3.1 Movement, there still remained the broad objective as a category due to test items covering the content of the objective.

As in other units, the test items were coded into the categories *consistent, inconsistent, broad,* and *narrow* depending on the type of content tested. For example, an item that required an answer only about the situation of the national movements after the 3.1 Movement was treated as *narrow* because it did not assess any of the aspects and characteristics regarding nation movements after the 3.1 movement. In addition, if an item did not test the content as it related to either the background of or reason for changes in Japanese colonial policies, it was coded as *narrow* for the targeted objective.

**a. Across all items**

Two hundred ninety nine test items from both public and private schools were analyzed for this unit. Overall, 269 of the 299 test items (90%) were coded as ones aligning with or at

[78] A nation-wide mass protest in 1919 against Japanese colonization.

least somewhat assessing the knowledge in the objectives, and 30 of the 299 items (10%) were coded as items not matching any objectives of the unit. The majority of items dealing with focusing on general objectives included 69% test items, while 20% test questions were regarded as ones assessing some part of the content required in the objectives. In looking at content alignment across the categories, over two-thirds of the test items for the unit aligned with the content required in the objectives. Across the objectives, test items for this unit assessed primarily two objectives (one general and one specific objective): 'Japanese colonial policies' and 'situations national independence movements against Japan' (30% and 24%, respectively).

**b. Across schools**

With regard to the distribution of test items between the two types of schools as shown in Table 25, the assessment of knowledge emphasized in the curriculum is severely unbalanced across objectives, with a similar pattern appearing for each type of schools. In other words, as mentioned above, while over two-thirds of the items were regarded as being *consistent* with objectives, only two objectives were heavily emphasized, and none of the test items assessed the background of or reason for the change in colonial policies. Of the ten objectives, only four were frequently assessed by teachers: 'Japanese colonial policies,' '3.1 Movement,' 'national independent movements,' and 'protection movements for Korean culture.' For the category '3.1 Movement,' 29 test items from both types were classified as *consistent* for this objective (n=15 (approximately 9%) in public and n=14 (approximately 11%) in private schools). Those test items generally asked what happened, who was involved in an event, or when it took place during the colonial period. In contrast, only a few items focused on how the struggles for national independence were related to changes of Japanese colonial policies (1%). Moreover, none of the items assessed an understanding of the way in which independence movements

Table 25: Frequency of Test Items for the Unit "The National Independent Movement"

| Knowledge in Objectives | | Consistent Public | Consistent Private | Broad Public | Broad Private | Narrow Public | Narrow Private | All** |
|---|---|---|---|---|---|---|---|---|
| | All items | Public n=168 | | | Private n=131 | | | 299(100%) |
| ***General Objectives*** | Inconsistent items | Public n=23 (14%) | | | Private n=7 (5%) | | | 30(10%) |
| Japanese colonial policies during each period* | | 48(29%) | 32(24%) | 1(<1%) | - | 3(2%) | 6(5%) | 90(30%) |
| Background of and reasons for changes in the colonial policies* | | - | - | - | - | 8(5%) | 3(2%) | 11(4%) |
| Background, processes, effects, and significances of the 3.1 Movement | | 2(1%) | - | - | - | 1(<1%) | - | 3(1%) |
| Background of 3.1 Movement* | | 5(3%) | 6(5%) | - | - | - | - | 11(4%) |
| Processes of 3.1 Movement* | | 2(1%) | - | - | - | - | - | 2(<1%) |
| Effects of 3.1 Movement* | | 2(1%) | 1(<1%) | - | - | - | - | 3(1%) |
| Significance of 3.1 Movement* | | 4(2%) | 7(5%) | | | | | 11(4%) |
| Organizations, places, and times in which armed independence resistances acted | | 4(2%) | 4(3%) | - | - | 5(3%) | 3(2%) | 16(5%) |
| Aspects and features of various national movements after the 3.1 Movement | | 3(2%) | 2(2%) | - | - | 6(4%) | 15(12%) | 26(9%) |
| Relationship between the changes in Japanese colonial policies and national independence struggles | | 2(1%) | - | - | - | - | - | 2(<1%) |
| ***Specific Objectives*** | | | | | | | | |
| Ultimate purposes of the economic policies of the Japanese Colonization | | - | - | - | - | - | - | - |
| Differences in independence movements before and after the 3.1 Movement | | - | - | - | - | - | - | - |
| Situations that national independence movements faced during the struggles against Japan after the 3.1 Movement | | 30(18%) | 36(28%) | - | - | 5(3%) | 2(2%) | 73(24%) |
| Actual circumstances and significance of the protection movements for Korean culture during the colonial period | | 11(7%) | 7(5%) | - | - | 3(2%) | - | 21(7%) |
| Total | | 113(67%) | 95(73%) | 1(<1%) | - | 31(19%) | 29(22%) | |
| Public and Private Total | | 208(70%) | | 1(<1%) | | 60(20%) | | 269(90%) |

*: Divided objectives

**: All reflects *consistent, broad*, and *narrow* for public and private.

changed after the 3.1 Movement, compared to either before the movement or the ultimate purpose of the colonial economic policies. Overall, there were no striking differences between the two schools. One small difference to note is that private schools seemed to focus more on national independence movements after 3.1 Movement (28%: approximately 10% higher than public schools).

#### 4.4.1.6. Test Items for the Unit "The Development of the Tae-han Min-guk"

As presented in Table 26, this unit emphasizes the understanding of the ideological conflicts after the 8.15 Liberation, the process of establishment of the Tae-han Min-guk[79] government, the Korean War, the growth of the economy, the movements for democratization of Korea, and efforts for reunification. The National Curriculum provides five general and three specific objectives for this unit. As in other units, the objective 'process of the establishment of the Tae-han Min-guk and background, processes, effects, and influences of the 6.25 War[80] (the Korean War)' was divided into two: 'the Tae-han Min-guk' and 'the 6.25 War.' Because there were no items that measured this objective, the objective for the 6.25 War was not divided further, despite its involving several different aspects. Since only five schools among the 22 assessed the knowledge of this unit, which asks about history after the colonization by Japan, only 24 test items focused on this unit (1.6% of 1,510 items). However, two schools used only one item to assess this unit, therefore it should be noted that only three schools actually attempted to assess the content of this unit. Of the 24 test items, half of them were labeled *inconsistent* with the content presented in the objectives, as shown in Table 26. Overall, test

[79] The Republic of Korea, established on August 15, 1948.
[80] The Korean War on June 25 , 1950.

Table 26: Frequency of Test Items for the Unit "The Development of the Tae-han Min-guk"

| Knowledge in Objectives | Consistent | | Broad | | Narrow | | All** |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| All items | Public n=11 | | | Private n=13 | | | 24(100%) |
| In consistent items | Public n=5 (46%) | | | Private n=7 (54%) | | | 12(50%) |
| ***General Objectives*** | | | | | | | |
| Process of the establishment of the Tae-han Min-guk and disorder following the 8.15 Liberation[81]* | 3(27%) | - | - | - | - | 1(8%) | 4(17%) |
| Background, processes, effects, and influences of the 6.25 War* | - | - | - | - | - | 1(8%) | 1(4%) |
| Graft and corruption under the Rhee Syngman Administration and the process of the 4.19 Revolution | 1(9%) | 1(8%) | - | - | - | 1(8%) | 3(13%) |
| Economic growth after the 5.16 Military Coup d'etat[82] and causes (motives) of economic growth | - | - | - | - | - | - | - |
| Specific situations and significance of democratic movements under the Yushin System[83], the 5.18 Democratic Movement[84], and the June Democratic Resistance[85] | 1(9%) | 1(8%) | 1(9%) | - | - | - | 3(13%) |
| Efforts made in order to establish a peaceful reunification after 7.4 South-North Joint Statement | - | 1(8%) | - | - | - | - | 1(4%) |
| ***Specific Objectives*** | | | | | | | |
| Our nation's confrontation concerning the proposal of Trusteeship and the movement of negotiation between the South and the North | - | - | - | - | - | - | - |
| Results of the Rhee Syngman administration's maneuvers to grasp political power for the long term | - | - | - | - | - | - | - |
| People's awareness of national problems for the last 30 years | - | - | - | - | - | - | - |
| Total | 5(46%) | 3(23%) | 1(9%) | - | - | 3(23%) | |
| Public and Private Total | 8(33%) | | 1(4%) | | 3(13%) | | 12(50%) |

*: Divided objectives

**: All reflects *consistent*, *broad,* and *narrow* for public and private.

[81] Korean liberation from the Japanese colony on August 15, 1945.

[82] Military coup d'etate by Pak Jong-hee on May 16, 1961.

[83] Revitalizing Reforms: an oppressed political system under the military regime.

[84] The Gwangju Democratic Movement on May 18, 1980 against Chon Tu-hwan who lead a military rebellion on December 12, 1979.

[85] A mass protest against Chon Tu-hwan, a military dictator, in June, 1987.

items from both types of schools focused on the three objectives: 1) the process of establishment of the Tae-han Min-guk government, 2) the Rhee Syngman[86] Administration and the 4.19 Revolution[87], and 3) the movements for Korean democratization. Across schools, over half the items from public schools were treated as *consistent* with and *broader* than the objectives (total 55%: 46%+9%). In contrast, 23% of the items from private schools were coded as *consistent*. However, for the purpose of this study, this generalization does not seem to be appropriate because test items analyzed for this unit were proved too insignificant to allow for accurate judgment. The information in this unit not emphasized in the assessments of both types of schools is valuable to know.

### 4.4.2. Historical Knowledge of Test Items for 10th Grade

Tenth grade high school Korean history spans history from the ancient era to the middle of the 19th century. Compared to middle school Korean history emphasizing political development, 10th grade high school are required to history from its political, social, economic, and cultural perspectives. In particular, the unit "Administrative Structures and Political Activities" contains relatively more pages[88] than any other unit, which is why approximately 40% of test items (n=519 of 1,315) concentrated on knowledge of political history.

As was performed for middle school test item analysis, high school history test items were coded according to degree to which the knowledge being tested was consistent with objectives. This analysis determines the distribution of test items across the objectives and those

[86] The first president of the Republic of Korea elected in 1948.
[87] Elite protests on April 19, 1960 by students, intellectuals, and remaining aristocrats against the corrupt Rhee Syngman administration.
[88] The high school textbook includes 101 pages of political history, with six units in 329pages. The section on 11th grade modern history was omitted from this count.

items emphasized by history teachers, and also compares the testing in public and private schools. Test items were labeled as *consistent, inconsistent, broad,* and *narrow*. Once an item was coded in one category, it was excluded from other categories. Objectives broad to assign items to were divided into two or more categories as appropriate. It is important to mention here that detailed contents were added in order to help code test items having objectives that were too general for 10th grade Korean history. For example, 'land system (Chōnsikwa)' was added to 'the development of agriculture during the Koryō era' in the unit of "Economic Structure and Life" as part of the agriculture-first policy for the Koryō. Several other categories were added to general objectives according to the knowledge emphasized in the curriculum. For example, 'to explore facts of early countries (in the Iron Age) in remaining records' was added to "The Culture of Pre-history Era and the Establishment of Nation." This knowledge is emphasized in both the curriculum (MOEHRD, 2001, p. 73) and in test items developed by history teachers.

#### 4.4.2.1. Test Items for the Unit "An Understanding of Korean History"

As Table 27 illustrates, "An Understanding of Korean History" emphasizes an understanding of the nature of history based on understanding of a variety of historical perspectives and the relations between the particularities and commonalities of Korean history. Originally, four general and two specific objectives were provided for this unit; one specific objective 'historical awareness in the East and the West and judgment on the value of historical materials' was divided into 'historical awareness in the East and the West' and 'judgment on the value of historical materials.' Twenty-six total test items measured the knowledge related to this

unit[89]. An item assessing both the purpose of studying history and the commonalities and particularities of history was treated as being broader than the objective. An item that assessed the understanding of either a commonality or a particularity of Korean history was coded as *narrow*.

Table 27: Frequency of Test Items for the Unit "An Understanding of Korean History"

| Knowledge in Objectives | Consistent | | Broad | | Narrow | | All** |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| All items | Public n=12 | | | Private n=14 | | | 26(100%) |
| Inconsistent items | Public n=0 (0%) | | | Private n=1 (7%) | | | 1(4%) |
| ***General Objectives*** | | | | | | | |
| Meaning of history in various ways | 5(42%) | 7(50%) | - | - | - | - | 12(46%) |
| Various perspectives of understanding of history and its characteristics | 2(17%) | 1(7%) | - | - | - | - | 3(12%) |
| Commonality and particularity of national tradition and culture | 1(8%) | 2(14%) | 1(8%) | | - | 1(7%) | 5(19%) |
| Purpose of studying history | - | 1(7%) | 1(8%) | 1(7%) | - | - | 3(12%) |
| ***Specific Objectives*** | | | | | | | |
| Historical awareness in the East and the West* | - | - | - | - | 2(17%) | - | 2(8%) |
| Judgment on the value of historical materials* | - | - | - | - | - | - | - |
| Commonalities of Korean history to world history and the particularities of Korean history | - | - | - | - | - | - | - |
| Total | 8(67%) | 11(79%) | 2(17%) | 1(7%) | 2(17%) | 1(7%) | |
| Public and Private Total | 19(73%) | | 3(12%) | | 3(12%) | | 25(96%) |

*: Divided objectives
**: All reflects *consistent*, *broad,* and *narrow* for public and private.

**a. Across all items**

Overall, most of the items (96%) fell into the category of *consistent*, *broad,* or at least partly consistent (*narrow*) as shown in Table 27. Only one item was not included in either of the categories. A number of items were classified as being consistent with objectives (73%),

[89] Only 8 pages of the textbook contain knowledge of the unit "Understanding of Korean History."

although the majority of items (n=23) concentrated on general objectives. Only two items assessed knowledge related to specific objectives. With regard to objectives, 'the meaning of history in various aspects' received the most emphasis. However, none of the items assessed the understanding of the evaluation of historical materials or the relationship between the particularity of Korean history as local history and the commonality of Korean history as a part of world history.

**b. Across schools**

Across schools, the distribution of items shows results similar to their item analysis. For example, a large portion of the test items from both types of schools measured the meaning of history, a general objective, from a variety of perspectives; none or few items measured the knowledge required by specific objectives. In fact, all of the items treated as *broad* (n=3) were related to the knowledge required in the general objectives. These items touched on all the aspects of this knowledge in their content.

#### 4.4.2.2. Test Items for the Unit "Culture of the Prehistoric Era and the Establishment of a Nation"

This unit emphasizes an understanding of the social changes in each period during the pre-history era in connection with the development of tools and the improvement of products. It also focuses on the establishment of the Korean nation and the origin of a national culture (MOEHRD, 2001). With respect to the emphasis of content to be mastered, this unit provides four general and six specific objectives. One category was added to the general objective 'facts

about early countries[90] in remaining records' because of its importance to the content for the Iron Age emphasized by the curriculum (MOEHRD, 2001)[91]. For coding, if a test item measured just the knowledge of living features or facts regarding pre-history or early countries without offering the explanation of historical artifacts or records, it was classified as *narrow.* This was done because this unit emphasizes understanding of pre-historical periods by studying historical remains and of the early historical era by studying historical records with the comprehension of the basis of archaeological perspectives.

**a. Across all items**

A total of 136 test items were analyzed for this unit according to their content. Table 28 reflects the percentages and number of test items related to the knowledge for this unit. Overall, a number of test items were coded as *consistent* with or at least somewhat similar to the objectives in the unit (85%); only 21 items did not measure the knowledge required in the objectives. In general, the frequency of the test items coded for this unit were distributed unevenly across the objectives, emphasizing the knowledge in the general objectives (75%) or focusing on two content categories 'living features of the prehistoric era' and 'facts of early countries' among the 11 objectives. In addition, 47% of the test items were considered to be aligned with the content in the objectives, and 36% of the test items were coded as *narrow*.

**b. Across schools**

As Table 28 indicates, test items from public and private schools are distributed very similarly, focusing on two content categories 'living features of the prehistoric era' and 'facts of

[90] Puyo, Koguryo, Okcho, Dong-ye, Paekche, Silla established in the Iron era. Puyo and Koguryo were placed in the central Machuria, Okcho and Dong-ye were located in the northeast in the Korean peninsula, and Paekche and Silla were placed in the south of Korea.

[91] The national curriculum requires students "to explore the facts of political systems, social traditions, and economic activities in early countries based on remaining records" (MOEHRD, 2001, p. 73).

Table 28: Frequency of Test Items for the Unit "Culture of the Prehistoric Era and the Establishment of a Nation"

| Knowledge in Objectives | Consistent | | Broad | | Narrow | | All** |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| All items | Public n=61 | | | Private n=75 | | | 136(100%) |
| Inconsistent items | Public n=8 (13%) | | | Private n=13 (17%) | | | 21(15%) |
| ***General Objectives*** | | | | | | | |
| Our national living places and the features of our national race and language in the prehistoric era | 2(3%) | - | - | - | 1(2%) | 1(1%) | 4(3%) |
| Living features of the prehistoric era by understanding remains and artifacts from that period | 18(30%) | 22(29%) | - | - | 10(17%) | 11(15%) | 61(45%) |
| Relationship of the developmental conditions of the culture and the changes of society to the establishment of the nation | 2(3%) | - | - | 1(1%) | - | - | 3(2%) |
| Background and process of the establishment of the Ko-Chosōn. | - | - | - | - | - | 1(1%) | 1(<1%) |
| Facts of early countries in remaining records* | 2(3%) | 7(9%) | - | 1(1%) | 12(20%) | 11(15%) | 33(24%) |
| ***Specific Objectives*** | | | | | | | |
| Process of human development in the prehistoric era | - | 1(1%) | - | - | - | - | 1(<1%) |
| Background of the change from the Paleolithic Age to Neolithic Age | - | 1(1%) | - | - | - | - | 1(<1%) |
| Social phenomena of the Neolithic Age | 1(2%) | - | - | - | - | - | 1(<1%) |
| Relationship between the growth of patriarchal power and social change | 1(2%) | 1(1%) | - | - | - | - | 2(2%) |
| Significance of the establishment of a nation by Tan-gun[92] by studying its mythological record | 1(2%) | 2(3%) | - | - | - | - | 3(2%) |
| Relationship between the culture of the Iron Age and social changes | 1(2%) | 2(3%) | - | - | 2(3%) | - | 5(4%) |
| Total | 28(46%) | 36(48%) | - | 2(3%) | 25(41%) | 24(32%) | |
| Public and Private Total | 64(47%) | | 2(2%) | | 49(36%) | | 115(85%) |

*: An included objective according to the curriculum

**: All reflects *consistent*, *broad,* and *narrow* for public and private.

[92] Priest-King. Tan-gun is assumed to be the first shaman-king of Ko-Chosōn. The myth of Tan-gun has been treated as the root of Korean identity; people regard him as the founder of Ko-Chosōn.

early countries.' However, the number of test items concentrating on these two was also labeled *narrow* because the items did not assess knowledge of the time presented in historical artifacts or records, which the unit emphasizes. Specifically, for 'the facts of early countries,' test items (as in the case of public schools) were labeled *narrow* more often than they were labeled *consistent*. In fact, these items tended to focus on the knowledge of under which social or cultural circumstances ancient people lived. In contrast, test items did not emphasize how people progressed throughout each period, the ways in which the pre-historical era advanced to the historical era, with what processes a political state could be established in the Bronze Age, or how a cultural civilization could effect social changes. For example, in the category of *consistent*, none of the items were found to be consistent with the objective 'the establishment of the Ko-Chosōn[93].' Both schools also placed less emphasis on 'the relationship between the culture of the Iron Age and social changes' and 'the relationship between the growth of patriarchal power and social changes.'

#### 4.4.2.3. Test Items for the Unit "Administrative Structures and Political Activities"

Of the 1,315 test items examined from both types of high schools, approximately 40% of the items (n=520) assessed content related to the unit "Administrative Structures and Political Activities". This unit stressed the understanding of political activities and changes in each historical era as a process of social development and efforts to solve current social problems (MOEHRD, 2001, p. 73). Thus, as can be seen in Table 29, the content emphasized by the objectives for this unit focuses on the process of national history by emphasizing knowledge of

[93] The oldest kingdom of Korea, known as the Chosōn and established between 2000 BC and 1000 BC. It reached the Iron civilization by the 3rd century. Later on, compared to the Chosōn, founded by Lee Song-gye in 1392, this state has been known as Old (Ko) Chosōn.

political changes from an ancient society to modern society. This unit features seventeen objectives: six general and eleven specific. For clarity, 'the reorganization of territory and the administrative structure in the 7th century' was divided into two categories: the reorganization of 'territory' and 'administrative structure' in the 7th century. Two specific objectives were excluded: 1) 'the political characteristics of the Koryō dynasty as a medieval nation'; and 2) 'the features of the Chosōn dynasty as an early modern society.' The content they required was similar to that in the general objectives. Three general objectives were enhanced: 'the establishment of the Barhae and its development' was added to the objective 'the reorganization of territory in the 7th century.' The original objective did not cover all of the knowledge of the Barhae kingdom, which recovered most of the old Koguryo territories; 'establishment of aristocratic Koryō society and its disturbances' was added to the Koryō era because of the importance of content to the curriculum (MOEHRD, 2001)[94]; 'the positive and negative aspects of factional politics' as part of the features of an early modern country was added to the objective 'features of an early modern nation in the political changes of Chosōn.' All of these combined contents are stressed in the curriculum (MOEHRD, 2001, pp. 75-76).

Test items were coded according to the knowledge presented in the curriculum. For example, in order to be consistent with the objective 'the development of the Koguryō, Paekche, and Shilla as the establishment of ancient nations,' test items should assess the comprehension of the fact that, while they established their centralized governing systems, these three kingdoms had characteristics similar to one another, such as expanding their territories, adopting Buddhism, and integrating local powers into central aristocrats. If an item measured only the political situations in the three kingdoms and included the features mentioned above but did not

[94] The curriculum requires students "to understand the establishment of centralized aristocratic politics, its disturbances and the changes in the Koryō society by a military coup" (MOEHRD, 2001, p. 75).

connect them to the establishment of an ancient nations, the item was classified as *narrow*. This principle was applied to the objectives for the political characteristics of the Koryō as a medieval society and of the Chosōn as an early modern society.

**a. Across all items**

In Table 29, the percentages of test items for each category reflect the proportions of total items from either public or private school (n=209, 311, respectively). Overall, 54% of the test items fell into the categories that assessed knowledge satisfied (31%), knowledge which included more than (3%), or knowledge which was somewhat related to the unit objectives (20%) while the remaining items (46%) did not measure knowledge that satisfied the objectives. Seventy-four percent of the items assessed knowledge stressed in the general objectives, while 26% of the items measured content from specific objectives. In both types of schools, some knowledge received greater emphasis, while other knowledge received less across objectives. For example, approximately 30% of the test items measured knowledge of the content related to four objectives: 'the development of three kingdoms as ancient nations,' 'the Koryō as a medieval society,' 'the Chosōn as an early modern society,' and 'features and problems of politics at the end of the Chosōn' (9%, 8%, 7%, 7%, respectively). However, they also included a number of test items identified as not sufficiently matching the content of the objectives. In contrast, none of the test items assessed knowledge of 'the differences of ancient countries, of medieval societies, and of early modern countries between the East and the West' or 'the backgrounds of Shilla's unification of three kingdoms.'

**b. Across schools**

Across schools, both public and private, the distribution of test items was very similar. For test items falling into the *broad* category, approximately 34% for public schools and 29% for

private schools were treated as measuring satisfactorily the knowledge emphasized in unit objectives, whereas approximately 20% of the test items did not satisfactorily assess unit objectives. Most test items assessed general objectives, specifically these fitting into the categories *consistent* and *narrow*. Regarding the number of items consistent with objectives, the percentage for 'features and problems of politics at the end of the Chosōn society' was the highest for public school assessments, while private schools placed greater emphasis on 'features of an early modern nation in the political changes of Chosōn.' The percentage for the objectives 'centralized politics of the Chosōn society' was relatively higher in public schools and 'development of three kingdoms as ancient nations' was relatively higher in private schools. However, because many test items were also categorized as *narrow*, it is hard to ascertain whether these objectives were highly emphasized by both types of schools. As mentioned earlier, the different characteristics of ancient, medieval, and modern societies between East and West were never assessed in either type of school. In addition, knowledge of territory reorganization, administrative structure in the 7$^{th}$ century as the process of national development, and modern elements in the 18$^{th}$ century were hardly assessed. For this unit, the majority of assessments in both types of schools tended to measure either knowledge that is not emphasized in the curriculum or that is only somewhat related to the unit.

Table 29: Frequency of Test Items for the Unit "Administrative Structure and Political Activities"

| Knowledge in Objectives | Consistent | | Broad | | Narrow | | All*** |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| All items | Public n=209 | | | Private n=311 | | | 520(100%) |
| Inconsistent items | Public n=92 (44%) | | | Private n=148 (48%) | | | 240(46%) |
| ***General Objectives*** | | | | | | | |
| Dominant power of the kings in ancient society | 3(1%) | 9(3%) | 1(<1%) | 1(<1%) | 2(<1%) | 3(<1%) | 19(4%) |
| Development of the Koguryō, Paekche, and Shilla[95] as the establishment of ancient nations | 5(2%) | 14(5%) | - | 2(<1%) | 11(5%) | 12(4%) | 44(9%) |
| Reorganization of territory in the 7th C* | 2(<1%) | - | - | - | 2(<1%) | 3(<1%) | 7(1%) |
| Reorganization of administrative structure in the 7th C* | - | - | - | - | 6(3%) | 10(3%) | 16(3%) |
| Establishment of Barhae and its development** | 5(2%) | 6(2%) | - | - | - | 1(<1%) | 12(2%) |
| Political characteristics meaning that the establishment of the Koryō[96] turned its society into a medieval nation | 7(3%) | 9(3%) | - | 3(<1%) | 10(5%) | 11(4%) | 40(8%) |
| Establishment of the aristocratic Koryō society** | 3(1%) | 2(<1%) | - | - | 2(<1%) | 3(<1%) | 10(2%) |
| Disturbances in the aristocratic Koryō society ** | 4(2%) | 3(<1%) | | | | 1(<1%) | 8(2%) |
| Features of an early modern nation in the political changes of Chosōn[97] | 7(3%) | 18(6%) | 1(<1%) | 3(<1%) | 4(2%) | 4(1%) | 37(7%) |
| Positive and negative aspects of factional politics** | 2(<1%) | 1(<1%) | - | - | 4(2%) | 5(2%) | 12(2%) |
| Modern elements present in the Chosōn society in the late 18th C | 1(<1%) | - | - | - | - | - | 1(<1%) |
| ***Specific Objectives*** | | | | | | | |
| Characteristics and differences between ancient nations in East and West | - | - | - | - | - | - | - |
| Development of the three kingdoms in relation to the changes in Chinese societies and the activities of northern nations | 4(2%) | 3(<1%) | - | - | - | - | 7(1%) |
| Background of Shilla's unification of the three kingdoms | - | - | - | - | - | - | - |
| Influence of the Kolp'um[98] System (the Bone-Rank System) on The political and social problems in Shilla society | 1(<1%) | 2(<1%) | - | 1(<1%) | - | - | 4(<1%) |

[95] In the last stages of the bronze culture, three kingdoms were established. Paekche and Shilla were prominent in the south, Koguryō in the north.
[96] One of the Korean dynasties established in the medieval era in 918 and destroyed in 1392 by Chosōn Dynasty.
[97] One of the Korean dynasties that appeared after the Koryō was founded by Yi Sōng-gye in 1392.
[98] A system in Shilla that differentiated social stratum according to the hereditary of bone linkage.

| ***Specific Objectives*** | Consistent | | Broad | | Narrow | | All*** |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| Particularities and differences of medieval Societies in East and West | - | - | - | - | - | - | - |
| Reformative administration of King Kong-min[99] in relation to national and international political situations | 2(<1%) | 2(<1%) | - | 1(<1%) | 1(<1%) | 2(<1%) | 8(2%) |
| Comparison between the early modern societies of East and West | - | - | - | - | - | - | - |
| Centralized policies of the Chosōn Dynasty | 9(4%) | 7(2%) | - | 1(<1%) | - | 1(<1%) | 18(4%) |
| Features of and problems in politics at the end of the Chosōn Society | 16(8%) | 14(5%) | - | - | 2(<1%) | 5(2%) | 37(7%) |
| Total | 71(34%) | 90(29%) | 2(<1%) | 12(4%) | 44(21%) | 61(20%) | |
| Public and private Total | 161(31%) | | 14(3%) | | 105(20%) | | 280(54%) |

*: Divided objectives

**: Added content according to the curriculum

***: All reflects *consistent*, *broad*, and *narrow* for public and private.

[99] A king at the end of the Koryō era who implemented a reform in order to control ruling classes.

#### 4.4.2.4. Test Items for the Unit "Economic Structure and Life"

This unit offers five general and six specific objectives in order to accomplish the goal of the unit. For coding, one objective 'the development of agriculture and international trade during the Koryō era' was divided into: 'the development of agriculture' and 'the development of international trade' during the Koryō era. It is important to note that four categories were added for three general objectives in order to code test items more precisely. These contents are emphasized in the curriculum for the purpose of learning this unit (MOEHRD, 2001, pp76-78). For example, as shown in Table 30, the agriculture of the Koryō was based on the agriculture-first policy, which included the system of land distribution and taxation. The taxation reformation system of the Chosōn also played a role in activating and improving the economy in the 18th century. In coding test items for this unit, when an item assessed knowledge of the land system (Chōnsikwa[100]) in the Koyrō society not as a private consideration but as a salary, it was labeled *narrow* and treated as assessing part of the targeted objective. One item that assessed the understanding of the hard life of peasants in an ancient society was classified as *narrow* for the objective 'institutionalized management of labor force and productive resources in ancient periods.' That is, this test item failed to measure the understanding of the institutionalized economic system although it did assess the public's lives under this system. As before, the coding of test items consistent with objectives was done according to the content presented in the curriculum.

**a. Across all items**

Table 30 indicates the percentages and number of test items that measured knowledge of

[100] The Field and Woodland Rank System: the salary system for the Koyrō aristocrats based on their rank within the bureaucracy.

this unit. For this unit, 245 test items were analyzed (104 from public and 141 from private schools). Overall, a total of 63% test items fell into the categories *consistent*, *broad*, or *narrow* while 37% were classified as not assessing the knowledge emphasized in the unit. Across objectives, over 80% of the items concentrated on the knowledge in general objectives. These items are fairly distributed across the general objectives: only 'agriculture-based economic life from early times' and 'international trade during the Koyrō era' were less assessed. In addition, 40% of the items were treated as measuring the content emphasized in the unit, while 6% were labeled as *broad*, and 17% were labeled as *narrow*.

**b. Across schools**

In Table 30, the results for both types of schools indicate no significant differences across objectives: public schools had a few more test items that measured the knowledge in the objectives (approximately 6%). Otherwise, the results from both types of schools show similar patterns. Across objectives, the distribution of test items is fairly even, with a few exceptions. In the category of being *consistent*, none of the schools assessed the understanding of 'the growth of sea power,' 'the background of Koyrō monasteries participating in industries and commerce,' or 'influence of Neo-Confucianism on the industrial politics in the Chosōn period.' In addition, test items for 'agriculture-based economic life' and 'the relationship between the division of peasant class and the conversion of the economy to capitalism' were less emphasized. Moreover, knowledge about 'agriculture-based economic life of our nation since early years' was ignored. According to the results of this unit from both types of school, an assessment of the understanding of the causes and background of the economic situations in each period was absent.

Table 30: Frequency of Test Items for the Unit "Economic Structure and Life"

| Knowledge in Objectives | Consistent Public | Consistent Private | Broad Public | Broad Private | Narrow Public | Narrow Private | All |
|---|---|---|---|---|---|---|---|
| All items | Public n=104 | | | Private n=141 | | | 245(100%) |
| Inconsistent items | Public n=34 (33%) | | | Private n=56 (40%) | | | 90(37%) |
| ***General Objectives*** | | | | | | | |
| Agriculture-based economic life since early times | - | 1(<1%) | - | - | - | - | 1(<1%) |
| Institutionalized management of labor force and productive resources in ancient periods | 3(3%) | 9(6%) | - | 4(3%) | 2(2%) | 4(3%) | 22(9%) |
| Ruling class-centered ancient economy** | - | - | 1(1%) | - | 3(3%) | 5(4%) | 16(7%) |
| Development of international trade during the Koryō era* | 1(1%) | 1(<1%) | 1(1%) | - | 1(1%) | - | 4(2%) |
| Development of agriculture during the Koryō era* | 4(4%) | 6(4%) | 1(1%) | | 2(2%) | - | 13(5%) |
| Land system (Chōnsikwa) of the Koryō society** | 7(7%) | 3(2%) | - | 1(<1%) | 2(2%) | 6(4%) | 19(8%) |
| Taxation system of the Koryō society** | 3(3%) | 6(4%) | - | 1(<1%) | 1(1%) | 1(<1%) | 12(5%) |
| Circumstances of an agriculture-first policy grounded in the Confucian ideology reinforced by the Chosōn Dynasty | 5(5%) | 6(4%) | - | 1(<1%) | - | 1(<1%) | 13(5%) |
| In the last 18th C, economy activation that was improved through an increased productive capacity and a brisk market | 7(7%) | 6(4%) | - | - | 1(1%) | 2(1%) | 16(7%) |
| Reformation of taxation system of the Chosōn* | 3(3%) | 5(4%) | - | - | 4(4%) | 1(<1%) | 13(5%) |
| ***Specific Objectives*** | | | | | | | |
| Process of growth where a sea power became a political power at the end of the Shilla Kingdom | - | - | - | - | - | 1(<1%) | 1(<1%) |
| Background of a monastery participating in manual industries and commerce during the Koryō era | - | - | - | - | 1(1%) | - | 1(<1%) |
| Problems with the Kwa-jeon[101] system | 1(1%) | 2(1%) | 4(4%) | 2(1%) | 1(1%) | 1(<1%) | 11(5%) |
| Influence of Neo-Confucianism as an administrative ideology in the industrial policies of the Chosōn society | - | - | - | - | - | - | - |
| Germination of Capitalism introduced to each industry | 6(6%) | 3(2%) | - | - | - | - | 9(4%) |
| Relationship between the division of the peasant class and the conversion of the economy to capitalism | 2(2%) | 1(<1%) | - | - | - | 1(<1%) | 4(2%) |
| Total | 45(43%) | 53(38%) | 7(7%) | 9(6%) | 18(17%) | 23(16%) | |
| Public and Private Total | 98(40%) | | 16(6%) | | 41(17%) | | 155(63%) |

[101] Land distribution system during the Chosōn Dynasty and the financial background of the ruling class.

#### 4.4.2.5. Test Items for the Unit "Social Structure and Life"

As can be seen in Table 31, this unit includes five general objectives and six specific objectives for understanding the various aspects of the social lives in each historical era. If an item dealt with the facts (features) of social structures of either the Koryō period or the Chosōn period, it was labeled *narrow* for the targeted objective. That is, the item failed to assess an understanding of those social structures in connection with the social order or ideology of each period, as emphasized by the curriculum. If an item questioned only the structure of the Kolp'um System (the Bone-Rank System)[102], it was coded *narrow* for the objective 'establishment of the Kolp'um System in the process of Shilla growing as a centralized ancient nation.' This item also failed to assess the understanding of the relationships between establishing this system and a centralized ancient nation in Shilla.

**a. Across all items**

For this unit, as Table 31 indicates, 162 test items were coded to determine whether their content was consistent with the knowledge in the unit objectives. Sixty-three percent of the items (n=102) were coded as *consistent, broad*, or *narrow* (36%, 4%, 23%, respectively) while 37% of the items were treated as not matching the content in unit objectives. The items were fairly well distributed across general objectives, with more items assessing knowledge in general objectives than that in specific objectives. Of the six specific objectives, one featured a relatively high number of items: 'efforts to rationalize the administrating order of the ruling class during the Chosōn era.' However, none of the test items assessed an understanding of 'the background of the Kolp'um System (the Bone-Rank System)' or 'relationship between the Sarim

[102] A caste system that the Shilla aristocracy, one of three kingdoms, tried to continue with their prestige and privilege, differentiating the inherent blood lines. The highest class was called Sōnggol (the Holy Bone) and the next was called Chin'gol (the True Bone).

Table 31: Frequency of Test Items for the Unit "Social Structure and Life"

| Knowledge in Objectives | Consistent | | Broad | | Narrow | | All* |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| All items | Public n=74 | | | Private n=88 | | | 162(100%) |
| Inconsistent items | Public n=29 (35%) | | | Private n=31 (33%) | | | 60(37%) |
| ***General Objectives*** | | | | | | | |
| Social strata formed in an ancient society and greater importance of familial social status than individual merit | 5(7%) | 4(5%) | 1(1%) | - | 4(5%) | - | 14(9%) |
| Establishment of the Kolp'um System in the process of Shilla's growth as a centralized ancient nation | - | 1(1%) | 1(1%) | 1(1%) | 1(1%) | 7(8%) | 11(7%) |
| Great importance of the Munbol aristocracy[103] during the Koryō era in accordance with the consolidation of social stratification | - | - | - | - | 6(8%) | 6(7%) | 12(7%) |
| Social structure of Chosōn in connection with the order of Confucianism | - | - | - | - | 6(8%) | 6(7%) | 12(7%) |
| Disturbance of the social status order and the development of active movements to raise social status in the late 18th C | 4(5%) | 10(11%) | 2(3%) | - | - | - | 16(10%) |
| ***Specific Objectives*** | | | | | | | |
| Background of the Kolp'um system that could be maintained in ancient society | - | - | - | - | - | - | - |
| Open society of the Koryō period | 4(5%) | 6(7%) | - | - | - | - | 10(6%) |
| Efforts to rationalize the administration order of the ruling class during the Chosōn era | 6(8%) | 12(14%) | 1(1%) | - | - | - | 19(12%) |
| Relationship between Sarim[104] power and the Confucian clan rules | - | - | - | - | - | - | - |
| Relationship between social constitutions and social changes | 2(3%) | 3(3%) | - | - | - | - | 5(3%) |
| Thoughts that influenced social changes | 2(3%) | - | - | - | - | 1(1%) | 3(2%) |
| Total | 23(31%) | 36(41%) | 5(7%) | 1(1%) | 17(30%) | 20(23%) | |
| Public and Private Total | 59(36%) | | 6(4%) | | 37(23%) | | 102(63%) |

*: All reflects *consistent*, *broad*, and *narrow* for public and private.

[103] A noble lineage.
[104] Group of Confucian elites during the Chosōn Dynasty.

power and the Confucian clan.'

**b. Across schools**

As can be seen in Table 31, with respect to the consistency of test items with the curriculum, only minor differences exist between public and private schools. Items from public schools were less consistent with the objectives than those from private schools, and items from private schools were less inconsistent than those from public schools. However, with respect to objectives, similar patterns across schools were found. The test items of both schools tended to measure part of the knowledge in two objectives, 'social stratification of the Koryō' and 'social structure of the Chosōn.' These items failed to assess the social structure of each society in connection with the Koyrō as the Munbol aristocratic society and the Chosōn as the Confucian society. The results for both schools show that for those items consistent with the objective 'efforts to rationalize the administrating order of the ruling class during the Chosōn era' had the highest number. However, in the category of being *consistent*, items measuring 'social structures of the Koryō and the Chosōn' regarding their ruling principles or ideology, 'the background of Kolp'um System,' or 'Confucian clan rules' based on Confucian ideology did not exist. Test items for this unit failed to assess basic knowledge about social structures throughout different eras presented in the general objectives.

### 4.4.2.6. Test Items for the Unit "The Development of a National Culture"

For the unit "Development of a National Culture," a total of 226 test items were analyzed. This unit includes eleven total objectives, six general and five specific, which reflect the requirements of the curriculum. As with other units, two objectives were divided into two or three more specific ones, depending on their content: 'the influence of Confucianism and

Buddhism and the development of an ancient society,' and 'high level of the Koryō culture influenced by Confucianism, Buddhism, Daoism, and the theory of geomancy' were divided into three as Table 32 shows. A test item that asked about one of these cultural developments was not coded as *narrow*; instead, it was coded as *consistent* with one of these objectives.

When considering the degree to which the content of items matched unit requirements, if an item assessed only the name of a Buddhist pagoda that was established during the ancient or the Koryō period, it was not treated as satisfying the category 'the cultural development influenced by Buddhism.' Regarding alignment with objectives, content on 'the background of composing the Taejangkyōng[105]' and 'cultural significances of development in the art of printing during the Koryō period' were regarded as content for 'influence of Buddhism.' In addition, the relationship between the development of ancient culture and ruling class was also regarded as matching the content of 'the influences of Buddhism on an ancient culture.'

**a. Across all items**

Table 32 reflects the percentages and numbers of test items that measured the knowledge emphasized in or related to this unit. As in other sections, the percentages in the table indicate the proportions of items from each types of school. Overall, more than half the items were regarded as assessing knowledge at least somewhat related to the objectives; a large proportion of items assessed knowledge not emphasized in this unit (57% and 43%, respectively). In addition, the frequency of the items was distributed more across the general objectives than across the specific: approximately 8% (n=19) of the items fell into the categories for specific objectives. With respect to the objectives, the percentages of items related to 'the cultural

[105] Wooden carved Tripitaka Koreana. The second Tripitaka Koreana, made during the war with Mongolia, consisted of over 80,000 wooden blocks, inscribed on both sides, and is now stored at Haein Temple. When the Mongolians invaded Korea in the early 13th century, the Koryō court performed this task in order to instill patriotism to secure the protection of Buddhism against the Mongols.

development influenced by Buddhism and Confucianism' in both the ancient and the Koryō periods and 'modern, national, and popular cultures in the modern society' were relatively higher than those for other objectives. Items that addressed 'the positive and negative aspects of the Sarim culture' were absent, and few test items about the creation of Hangul (Korean alphabet), the influences of Buddhism on the Koryō periods, and the relationships between the ruling class and the development of culture were in evidence. In addition, taking into consideration those items broader than the targeted objectives, half of the test items assessed knowledge emphasized by the unit (n=113, 50%=48%+2%).

**b. Across schools**

As shown in Table 32, for this unit, there are fewer similarities across objectives between the distribution of items in public schools and private schools. More of the test items of private high schools had content that matched knowledge of the objectives, and fewer test items were coded as *narrow* (52%, 5%, respectively). In addition, more of the test items of private schools were distributed across the intensive objectives. However, test questions of private schools tended to focus on the cultures of the modern society (approximately 16%), including modern, national, and popular cultures. In the case of public schools, most test items fell into the category of general objectives, especially for 'the features of culture development' but not for the connections of ruling class to the development of culture in each historical period. In addition, both types of schools lacked measurement of 'the particularity and commonality of national culture': none of the test items related to this knowledge. Of the items analyzed, regardless of the type of school, many assessed the general dispositions of cultures influenced by Confucian learning, the development of Buddhism, and cultural circumstances in the modern era. However, none of the items assessed the general concept of how national culture was established or how

Table 32: Frequency of Test Items for the Unit "The Development of a National Culture"

| Knowledge in Objectives | Consistent | | Broad | | Narrow | | All** |
|---|---|---|---|---|---|---|---|
| | Public | Private | Public | Private | Public | Private | |
| All items | Public n=69 | | | Private n=157 | | | 226 |
| Inconsistent items | Public n=32 (46%) | | | Private n=65 (41%) | | | 97(43%) |
| ***General Objectives*** | | | | | | | |
| Establishment of process and features of national culture | - | - | - | - | - | - | - |
| Influence of Confucianism and the development of an ancient culture* | 6(9%) | 4(3%) | 1(1%) | - | - | - | 11(5%) |
| Influence of Buddhism and the development of an ancient culture* | 5(7%) | 11(7%) | 1(1%) | | 5(7%) | 1(<1%) | 23(10%) |
| High level of the Koryō culture influenced by Confucianism* | 7(10%) | 10(6%) | - | 1(<1%) | 1(1%) | 2(1%) | 21(9%) |
| High level of the Koryō culture influenced by Buddhism* | 4(6%) | 9(6%) | | | 2(3%) | 1(<1%) | 16(7%) |
| High level of the Koryō culture influenced by Daoism and Pungsujiri theory[106] (the theory of geomancy)* | 2(3%) | 4(3%) | - | - | - | - | 6(3%) |
| Significance of the creation of Hangul[107] in terms of the development of national culture | - | 2(1%) | - | - | - | - | 2(<1%) |
| Learning and arts of the Chosōn era in relation to a governing order | - | 9(6%) | - | - | - | 2(1%) | 11(5%) |
| Elements of modern, national, and popular cultures presented during the quickening period of modern society | - | 18(12%) | - | - | - | 2(1%) | 20(9%) |
| ***Specific Objectives*** | | | | | | | |
| Influences of Buddhism on ancient society and culture | - | 6(4%) | 1(1%) | 1(<1%) | - | - | 8(4%) |
| Political, social, and cultural influences of Buddhism during the Koryō | - | 2(1%) | - | - | - | - | 2(<1%) |
| Positive and negative aspects of the Sarim culture | - | - | - | - | - | - | - |
| Relationship between the awareness of the ruling class on current situations and the development of culture | 1(1%) | 1(<1%) | - | - | - | - | 2(<1%) |
| Elements of popular and Korean culture during the quickening period of modern society | 1(1%) | 6(4%) | - | - | - | - | 7(3%) |
| Total | 26(38%) | 82(52%) | 3(4%) | 2(1%) | 8(12%) | 8(5%) | |
| Public and Private Total | 108(48%) | | 5(2%) | | 16(7%) | | 129(57%) |

*: Divided objectives

**: All reflects *consistent*, *broad,* and *narrow* for public and private.

[106] During the Koryō period, the sites on which all temples and monasteries of Buddhism, court buildings, and even aristocratic houses were to be built depended heavily on geomantic process.

[107] Korean alphabet, created in 1441 by King Se-jong and his scholars, during the Chosōn Dynasty

Buddhism and Confucianism became integrated, establishing the Koryō culture. Both types of schools also featured few test items that questioned why science and technology remained underdeveloped under the Confucians who governed during the Chosōn period. Specifically, few items assessed how the Chosōn ruling class, which was based on Confucianism, tried to establish a national identity

### 4.4.3. Differences between the Measurement of Historical Knowledge in Middle School and High School

**a. Across units for middle school**

Overall, those test items measuring the knowledge consistent with the content in the curriculum was at a low level, and high percentages of the items assessing the content not emphasized in the curriculum can be seen in Table 33. The percentages reflect those test items related to the targeted unit only. In only two units did more than 50% of the test items assess the knowledge in the unit objectives: "The Deployment of Movement for National Sovereignty Safeguard" and "The National Independence Movement" (the percentages for *consistent* + *broad*). More than 30% of the items for four units among the six assessed content not emphasized in the unit objectives. Regarding *narrow*, almost half the items for the unit "The Changes in the Chosōn Society" assessed knowledge somewhat similar to the content in the objectives.

The percentages of the test items across units varied with respect to measuring knowledge presented in the objectives. For example, the knowledge that was expected in test items for the unit "The National Independence Movement" showed the greatest alignment with the content emphasized in the curriculum (70%), while knowledge assessed for the unit 'The Enlightenment and Independence Movement' showed the least alignment with the objectives

(22%). With respect to *inconsistent*, test items for the unit "The Development of the Tae-han Min-guk" primarily assessed knowledge not presented in the objectives (50%), while test items for the unit "The National Independence Movement" least measured knowledge not emphasized in the objectives (10%).

Table 33: Summary of the Measurement of Historical Knowledge in Middle School Test Items

| | | Measurement of Knowledge | | | |
|---|---|---|---|---|---|
| Units | N | Consistent | Broad | Narrow | Inconsistent |
| The establishment of the Chosōn Dynasty and its development | 339 | 36% | 2% | 30% | 32% |
| The changes in the Chosōn society | 241 | 25% | 4% | 49% | 22% |
| The enlightenment and independence Movement | 300 | 22% | <1% | 32% | 46% |
| The deployment of movement for national sovereignty safeguard | 307 | 46% | 4% | 19% | 30% |
| The national independence movement | 299 | 70% | <1% | 20% | 10% |
| The development of the Tae-han Min-guk | 24 | 33% | 4% | 13% | 50% |

**b. Across units for high school**

Overall, for middle schools, test items that assessed knowledge of the objectives were at low level, and a high number of items assessed content were not related to the objectives. As can be seen in Table 34, over 70% of the targeted items that measured the content emphasized in the objectives addressed only one unit "Understanding of Korean History." However, only 26 items among 1,315 assessed this unit (2%), therefore, this result should be regarded as an exception. In three units, nearly 50% of the items that assessed content aligned with the objectives; in two units, less than 40% of the items measured knowledge consistent with the objectives. In particular, the knowledge emphasized in the objectives for the unit "Administrative Structures and Political Activities" was assessed least. On the other hand, there were four units in which more than 35% of their knowledge did not relate to the objectives. With regard to *inconsistent*,

the percentage for the unit "Understanding of Korean History" was the lowest, and the percentage for the unit "The Culture of Pre-history Era and the Establishment of A Nation" was relatively low. Again, the content for the unit "Administrative Structures and Political Activities" was least related to the knowledge in the objectives.

Table 34: Summary of the Measurement of Historical Knowledge in High School Test Items

| | | Measurement of Knowledge | | | |
|---|---|---|---|---|---|
| Units | N | Consistent | Broad | Narrow | Inconsistent |
| Understanding of Korean history | 26 | 73% | 12% | 12% | 4% |
| The culture of pre-history era and the establishment of a nation | 136 | 47% | 2% | 36% | 15% |
| Administrative structures and political Activities | 520 | 31% | 3% | 20% | 46% |
| Economic structure and life | 245 | 40% | 6% | 17% | 37% |
| Social structure and life | 162 | 36% | 4% | 23% | 37% |
| The development of a national culture | 226 | 48% | 2% | 7% | 43% |

**c. Differences between middle and high schools**

Across middle and high schools, no significant differences exist between school types in finding the items categorized as *consistent* or *broad*: test items from both schools were found to be less likely to assess the content in the unit objectives. For example, approximately 40% of the items from both school types measured knowledge presented in the objectives as Table 35 indicates. However, a large proportion of items from both school types did assess knowledge somewhat similar to or never emphasized in the objectives (more than 50%). With regard to *narrow* and *inconsistent*, minor differences were found between both schools. Items from middle school tended more to assess the content labeled as *narrow* than high school (29%, 19%, respectively) while items from high school assessed more content not expected in objectives (39%, 29%, respectively).

Table 35: Measurement of Historical Knowledge in Middle and High Schools

| | | | Measurement of Knowledge | | |
|---|---|---|---|---|---|
| Schools | N | Consistent | Broad | Narrow | Inconsistent |
| Middle School | n=1,510 | 40% | 2% | 29% | 29% |
| High School | n=1,315 | 39% | 3% | 19% | 39% |

## 4.5. Results for Performance Assessments

For this study, data from the performance assessments of 29 of 32 schools was collected (19 middle schools and 10 high schools). The schools provided the annual plan for assessments for the 2004 school year, including the methods, topics, and criteria of assessments: eight schools provided only topics for performance assessments. In each school, performance assessments consisted of 30% to 40% of all of the history assessments. In these proportions, schools included assessing students' attitudes, or the organization of their class materials. This each school actually allotted 10% to 15% of the assessments for performance assessments. Because of the difficulty in obtaining the products of students for the results of performance of assessments, this study classified and described only those topics of performance assessments as they related to the level of historical understanding.

### 4.5.1. Topics of Performance Assessments

A total of 44 performance assessments were collected from both middle and high schools, and were then sorted within 17 topics, as seen in Table 36. Several schools required students to complete more than one performance assessment, thus, the total number of assessments is more than the total number of schools. Among the schools, where 'testing historical knowledge' was used as a performance assessment from seven schools (16%), it was excluded from this study

because it is unrelated to performance assessments. Overall, 59% of the assessments required students to complete tasks that assess a low level of historical understanding: this percentage was

Table 36: Topics of Performance Assessments

| Topics | Middle School (n=19) | High School (n=10) | Level of Understanding |
|---|---|---|---|
| Criticizing/interviewing a historical figure | 4 | 1 | 3 & 4 |
| Publishing a historical newspaper or magazine | 3 | - | 3 & 4 |
| Evaluating/discussing a historical event or society | 3 | - | 3 & 4 |
| Creating a cartoon of a past event or figure | 2 | 1 | 3 & 4 |
| Writing a historical account | 2 | - | 3 & 4 |
| Planning a reform for the past society | 2 | - | 3 & 4 |
| Writing a historical diary or drawing a past event | 1 | - | 2 |
| Remaking historical lyrics | 1 | - | 2 |
| Reporting on a historic site | 4 | 4 | 2 |
| Reporting on a historical figure | - | 1 | 2 |
| Reporting on a family history through genealogy | - | 1 | 2 |
| Describing the past life presented on a TV documentary | 1 | - | 2 |
| Writing reflections on a historical novel | - | 1 | 2 |
| Describing patriotic organizations in the end of Chosōn | 1 | - | 2 |
| Summarizing the main points of the textbook | 1 | 1 | 2 |
| Drawing historical maps in the textbook | 2 | - | 1 |
| Testing historical knowledge | 3 | 4 | NA |
| Total | 30 | 14 | |

included the category 'testing historical knowledge.' The most common topics from those schools were 'reporting about a historic site' (18%) and 'criticizing or interviewing a historical figure' (11%). In addition, seven topics occurred only once among the schools, including 'writing a historical diary or drawing a past event,' 'remaking historical lyrics,' 'reporting on a historical figure,' 'reporting on a family history through genealogy,' 'describing the past life

presented on a TV documentary,' 'writing reflections on a historical novel,' and 'describing patriotic organizations at the end of Chosōn.'

### 4.5.2. Levels of Historical Understanding on Performance Assessments

**a. Across all tasks**

*Historical Understanding Level 1*: 'Drawing a historical map in the textbook' was classified as historical understanding level 1 because this task requires students to reproduce the structure of the maps included in the textbooks. It does not require students to employ any complex mental activities, rather, it allows students to master where and when a historical event took place, the names of places where a historical agent acted, or the scale of these historical activities. Of 44, only two assessments (5%) required students to draw maps where the places of armies fought for national independence from Japan in early the 20th century.

*Historical Understanding Level 2*: Of the 17 topics, nine were categorized as level 2: 'writing a historical diary or drawing a historical event,' 'remaking a historical lyrics,' 'reporting on a historic site,' 'reporting on a historical figure,' 'reporting on family history through genealogy,' 'writing reflections on a historical novel,' 'summarizing the main points of the textbook,' 'describing the past life presented on a TV documentary,' and 'describing patriotic organizations in the end of Chosōn' (39%). Each of these topics requires students to describe the information they obtained from historical places, reading materials, or historical records. For example, the assessment 'describing about the patriotic organizations in the 19th C' required students to explain about the political activities, historical agents, or the background of Yaeguk-gyemong (patriotic enlightenment) organizations. In order to produce their final reports, students have to grasp the meaning of the materials or information about the organizations, and then, put

that knowledge into a new context. Students employ simple mental activities such as interpreting, translating, and organizing to obtain the information. The task 'writing a historical diary specifically encourages students to use their imaginative understanding about the past by putting themselves into the activities of historical agents or situations.

*Historical Understanding Levels 3 and 4*: Of these 17 topics, six were determined to be at levels 3 and 4: 'criticizing/interviewing a historical figure,' 'publishing a historical newspaper or magazine,' 'evaluating/discussing a historical event or society,' 'creating a cartoon of a past event or figure,' 'writing a historical account,' and 'planning a reform bill for the past society.' These two levels were classified in the same category because, in general, the tasks of each included two levels in their assessments. A total of 41% of the performance assessments were classified at these levels because they require students to use a variety of historical reasoning abilities. For example, in order to complete the assessment 'planning a reform for the past society,' students have to compare a variety of reforms in order to discover their differences and commonalities (level 3). They then have to determine the problems of the targeted society (level 4). After that, the students have to provide alternative solutions to the problems of that society by proposing a reform bill that includes reasonable evidence (level 4). Three schools in particular required students to perform the task 'publishing a historical newspaper or magazine,' allowing students to use multiple mental abilities. These schools asked students to include political, economic, cultural, and international pages; editorials; cartoons; and interviews of historical figures; among other things. In order to complete each section for the newspaper, students have to obtain historical data, formulate historical questions, identify issues and problems of the past, and evaluate the implementation of decisions or actions as well as use their imaginative understanding of those periods (level 4).

**b. Across schools**

Across middle and high schools, it was found that middle schools tended to offer tasks that demanded higher levels of historical understanding than did high schools. Only 14% of the performance assessments (2 tasks) from high schools were classified as levels 3 and 4, such as 'criticizing or interviewing a historical figure' and 'creating a cartoon of a past event or figure,' while 53% of the assessments (16 tasks) from middle schools were classified as levels 3 and 4. Moreover, when 10% of the tasks from middle schools replaced testing as performance assessments, 29% of the tasks from high schools used testing as performance assessments. In terms of a high level of understanding, the tasks from middle schools provided a greater variety of topics that required students to use various skills such as gathering, analyzing, corroborating, judging, or contextualizing materials in their presentations.

## 4.6. Results for the Quality of Test Items and Alternatives

Test items were also analyzed to determine whether the items and alternatives were well developed with regard to formatting and writing the stem and choices. For this section, only multiple-choice items were analyzed. In terms of constructing test items and alternatives, 46 short answer questions were excluded (27 from middle school and 19 from high school). Table 37 reflects the percentages of the items not meeting Haladyna et al.'s (2002) criteria for well developed items. Overall, the results of the analysis for constructing multiple-choice (MC) item writing were very similar across types of school (public or private) and levels of school (middle or high). The percentages for the category 'use negative terms' for the stem were highest among all others, and the percentages for the categories 'not use length equal' for the choices were much higher than other categories.

Table 37: Percent of Errors in Wiring Multiple-Choice Items

| Errors for Guidelines of MC Items | Middle School (9th) | | High School (10th ) | |
|---|---|---|---|---|
| | Public | Private | Public | Private |
| | n=925 | n=558 | n=529 | n=767 |
| *Formatting concerns* | | | | |
| Using complex multiple-choice format | 4% | 6% | 16% | 9% |
| Not keeping vocabulary simple | <1% | - | <1% | <1% |
| *Writing the stem* | | | | |
| Not stating the stem in either question or completion form | - | - | - | - |
| Not minimizing the amount of reading | 1% | <1% | 2% | 4% |
| Unclear directions | 2% | <1% | 5% | - |
| No central idea in the stem | 2% | 2% | 1% | 2% |
| Window dressing (excessive verbiage) | 2% | 0% | 2% | 3% |
| Using negative terms | 32% | 30% | 33% | 37% |
| *Writing the choices* | | | | |
| Distractors are not plausible | - | - | - | - |
| Not one right answer | - | - | - | - |
| No logical/numerical order | - | - | - | - |
| Choices are overlapping | - | - | - | - |
| Choices are not homogenous | - | - | - | - |
| Unequal length | 22% | 25% | 25% | 20% |
| None of the above used | NA | NA | NA | NA |
| All of the Above used | - | - | - | - |
| The term "not" in choice | - | - | - | - |
| Gives clues to the right answer | 2% | <1% | <1% | 2% |
| Distractors are not plausible | - | - | - | - |
| Not use common errors of students | - | - | - | - |

### 4.6.1. Formatting Concerns

With respect to multiple-choice (MC) test item formats, overall, most of the items were well developed. In terms of using vocabulary, most of the items provided simple and easy words for students to understand. Only a few test items used difficult words, such as classic Chinese, in the context. In this case, the items provided explanations to help students understand both the questions and the materials given. Primary sources for history offered as a text in a number of

the items included a Korean phonic writing system for Chinese, with explanations rather than graphic letters. Except for a few questions (<1% or 0% for all schools) that did not provide the explanations for Chinese, the context of most of the test items could be fully understood by students. In terms of using a complex MC format, a number of test items, choices were grouped into sets for students to choose. For example:

> Q: There are two meanings for history. What are the appropriate explanations in terms of that statement?
>
> History refers to events in the past, that is, about the whole past that human beings have experienced. History refers to a historical science or a historical account that reconstructs or explores the past experienced by human beings.
>
> 1) a) explains history as a record, and b) explains history as a fact
> 2) a) means history is objective, and b) means history is subjective
> 3) When people learn history, they, in general, learn the meaning of a)
> 4) In the case of b), history means recorded materials or historical accounts
>
> A. 1 & 2
> B. 1 & 3
> C. 2 & 3
> D. 2 & 4
> E. 3 & 4

This item is more complicated than the item that provides choices directly after the text. In order to answer this question, students first have to define the meaning of each text given, and then identify the interpretations for the text. They must choose a grouped answer. This process is more difficult for students and less efficient than multiple-choice items that provide options right after questions or text. As Table 37 indicates, approximately five percent of the items from middle schools and 12% of the items from high schools had complex MC format. In addition, of 2,779 items, 54% (n=793 from middle school) and 39% (n=511 from high school) of the items did not provide a context in their stems. Most of these items may encourage rote memory for students. For example, “Which explanation is an appropriate answer for the growth of Sarim?”,

or “What were the main features of Catholicism in the beginning in Korea?” In order to answer these questions about history, students have to remember what they learned in class and where to locate the answer in textbooks or other resources. Without context, these items do not allow students to use higher levels of mental activity in answering questions about history.

### 4.6.2. Writing the Stem

In general, most of the test questions were well constructed, as indicated in Table 37. None of the items were written in either an uncompleted format. All were presented with full questions in their stems. Among the categories for writing the stem, the category ‘using negative terms’ had the highest percentage: approximately one-third of the test items from both middle and high schools used negative questioning forms. These items used the negative words with cautions such as underlining or boldface in order to ask what was not true. However, the difficulty with these items is that they were not associated with a context, and, instead, encouraged students to use their abilities merely to remember and recall historical knowledge. For example, “what is an *inappropriate* explanation of the foreign policy in the early Chosōn?,” or “Among the following, what is NOT an explanation of the factional politics in the Chosōn?” Such items do not allow students to use their abilities to analyze or evaluate certain historical events or actions of people in the past. Twenty-five percent (25%, n=373) of the items from middle schools and 21% (n=278) of the items from high schools had negative questions without context.

A few items did not present central ideas in their questions. For example, questions such as “what is a correct explanation?,” or “what is NOT an appropriate explanation among the following?” are too general. With regard to ‘window dressing’ and ‘minimizing the amount of

reading,' a few items had unnecessary pieces of information or were too long. These items tended to make it more difficult for students to focus on the questions. Although the percentages of these types of items were quite low (1%, <1%, 2%, 4% for middle and high schools), they can distract students from answering correctly.

### 4.6.3. Writing the Choices

With regard to writing the choices, in general, the test items were well presented, used the options logical, and were functional, plausible, and appropriate to the stem with true statements and one right answer. All items had five distractors in, did not have choices with negative words such as NOT, and used common errors of students. However, many items had options of an unequal length: over 20% of the test questions from both middle and high schools had options whose length was not unequal. Specifically, 8% of the test questions from the schools had one long option, and almost half of them were correct answers. For example,

> Q: What is an historical description that differs from the perspective in the following text?
>
> A historian must reveal the past in its original circumstances, and not include his own perspective. Then, history itself should communicate about the past with only its facts.
>
> A. In the Paleolithic Age, people used a stone ax as a hunting device for the first time.
> B. Maga, Uga, Jōga, and Guga in the Puyō ruled Sachuldo.
> C. Ko-Chosōn was destroyed by the invasion of the Han China.
> D. In the 5$^{th}$ C, the Silla adopted Buddhism through the Koguryō.
> E. The scholarly world in the North Korea does not accept the unification of the Shilla because of the use of foreign power and territorial incompleteness.

In order to answer this item, students must determine which historical account is factual and which one includes a historian's interpretation. Option E is the answer; it is the lenthieset

choice. Option E presents an extreme case, indicating that it may be the correct answer because answers that tend to be long explain their correctness with more words. Thus, the correct answer is easier to distinguish from the other options. In this case, students who don't know the correct answer can choose the one option that has a relatively long answer.

A small number of the test items had clues to the answer within the test context or for other questions. The following question asks about Yōng-jo's Tang-pyōng (impartiality) policy.

> Q: What is the purpose of Yōng-jo's Tang-pyōng (impartiality) policy in terms of the context below?
>
> There have been no more times when *factions* on these days have been influenced so severely than on any other days. At first, the dispute occurred between different perspectives of Confucianism, but now one side has put a charge on another side… Now, facing the age of reform, adopt the spirit of the Tang-pyōng.
>
> A. Encouraging Shil-hak
> B. Breaking down Sedo administration
> C. Pursuing righteous government
> D. Settling down peasants' lives
> E. Reorganizing the balance of *factions*

The answer is E 'reorganizing the balance of factions.' This item provides students with a direct hint to the correct answer by offering the word 'factions' in the reading text. In this case, even without the first sentence of the text, this item could function well for students to answer. Although few items had clues to the answers (20 items from middle school and 21 items from high school), it is not an appropriate way in which to assess students' ability to understand a given text.

### 4.6.4. Across schools

The schools, regardless of type or level of schools indicated very similar results. As mentioned earlier, the category 'use negative, no positive' had the highest percentages for all

schools, and the percentages for the category 'not use length equal' were the next highest. All items classified in Table 37 appeared in the same categories across the schools. There were few sizable differences in the categories across these schools and only one minor difference with regard to using complex multiple-choice (MC) format. High schools, both public and private (16%, 9%, respectively), tended to offer more complex MC items with grouped answers, especially in public schools, than did middle schools. High school students learn a vast amount of historical knowledge, and high school tests tended to provide complex MC questions that attempted to assess a lot of knowledge within one item. Many of the items had no specific subject. For example, the item states "What is the correct answer about village lives in the Chosōn society?" Then provides five explanations to choose from. Based on this context, the options are: 1) a & b; 2) a, b, & c; 3) b & c; 4) a, b, c, & e; 5) b, c, d, & e.

## 4.7. Results of the Survey about Teacher Training on Assessment

With regard to assessment, history teachers who provided their test items for this study were asked about academic courses taken during their teacher preparation and about their professional activities. Of all the teachers in 32 middle and high schools, 28 history teachers responded to the survey questions. Teachers responded to seven questions that asked to what extent they had learned classroom assessment before and during their professional lives.

### 4.7.1. Teacher Responses to Teacher Preparation Courses related to Assessments

As shown in Table 38, of the 28 respondents, 24 history teachers (86%) answered that they took courses related to assessment in their teacher preparation programs. The remainder of the questions in the table were answered by teachers who had been trained in assessment during

Table 38: Summary of Teacher Responses to Teacher Preparation Courses on Assessment

| Questions | Category | # and % of teachers (n=28) | |
|---|---|---|---|
| Did you have a class/training on assessment in your college coursework on teacher preparation? | Yes | 24 | 86% |
| | No | 3 | 11% |
| | No answer | 1 | 4% |
| - How many credits was the assessment class? | 4 credits | 1 | 4% |
| | 3 credits | 10 | 36% |
| | 2 credits | 6 | 21% |
| | Don't remember | 7 | 25% |
| - What percent of the class was spent on the theory of assessment (e.g., validity, reliability)? | 60% | 1 | 4% |
| | 50% | 2 | 7% |
| | 40% | 1 | 4% |
| | 30% | 4 | 14% |
| | 20% | 2 | 7% |
| | 15% | 1 | 4% |
| | Don't remember | 13 | 46% |
| - What percent of the class was spent on how to design multiple-choice items? | 50% | 1 | 4% |
| | 10% | 5 | 18% |
| | Didn't learn | 5 | 18% |
| | Don't remember | 13 | 46% |
| - What percent of the class was spent on how to design constructed response items and/or performance assessments? | 20% | 1 | 4% |
| | 10% | 5 | 18% |
| | Didn't learn | 5 | 18% |
| | Don't remember | 13 | 46% |
| - What percent of the class was spent on how to interpret the results and use them for instructional planning? | 20% | 2 | 7% |
| | 10% | 3 | 11% |
| | Didn't learn | 3 | 11% |
| | Don't remember | 16 | 57% |

their college experience. More than half of these teachers did not remember the amount of time spent on learning the theory of assessment; the design of assessments, including multiple-choice items; constructed-response items and performance assessments; and the interpretation of test results for their instructional planning. With regard to the theory of assessment, 11 of the 28 history teachers (39%) responded that they had learned theory of assessments such as validity and reliability, ranging from 60% to 15% of the coursework on assessment. However, for the questions asking about the practice of assessment, only six of the teachers (21%) answered that

they had learned how to construct multiple-choice ones (five teachers with 10%) and constructed response items and performance assessments (one teacher with 20%, five of the 24 teachers with 10%). Only five of the teachers (18%) answered that they had learned how to interpret results of the tests for their instructional planning. Several teachers answered that they had never learned how to design multiple-choice and constructed response items, including performance assessments (five teachers) or how to interpret test results (three teachers, 11%). For an additional request, 'Please indicate below what other topics were covered in the class,' only two teachers (7%) responded that the class covered, in general, theories of assessment, such as the purposes or the forms of assessment.

### 4.7.2. Teacher Responses to Professional Development Activities related to Assessment

Table 39: Summary of Teacher Responses to Professional Development Activities on Assessment

| Questions | Category | # and % of teachers (n=28) | |
|---|---|---|---|
| Since you have been a teacher, have you had any professional development activities related to assessment during the past year? | Yes | 3 | 11% |
| | No | 25 | 89% |
| - How many hours did you spend on professional development activities related to assessment during the past year? | 20 hours | 2 | 7% |
| | No answer | 1 | 4% |
| - What percent of the activities was spent on the theory of assessment (e.g., validity, reliability)? | 20% | 1 | 4% |
| | 10% | 1 | 4% |
| | No answer | 1 | 4% |
| - What percent of the activities was spent on how to design multiple choice items? | 10% | 1 | 4% |
| | No answer | 2 | 7% |
| - What percent of the activities was spent on how to design constructed response items and/or performance assessments? | 10% | 2 | 7% |
| | No answer | 1 | 4% |
| - What percent of the activities was spent on how to interpret the results and use them for instructional planning? | 10% | 1 | 4% |
| | No answer | 2 | 7% |

With respect to professional development, teachers were also asked about training activities related to assessment. Overall, the results of this survey showed that very few history teachers had taken part in teacher training activities over the past year, as seen in Table 39. Of the 28 teachers, only three teachers (11%) responded that they had taken part in teacher training activities regarding assessment in the past year; two teachers (7%) responded that they had had 20 hours of teacher training. In addition, for each category, one or two teachers did not specify the type of assessment activities covered.

### 4.7.3. Teacher Responses to Future Professional Development Activities related to Assessment

Teachers were asked to describe what professional development activities related to assessment would help them in their teaching. Of those 28 teachers, 14 teachers described more than one professional development activity (50%). Overall, three common themes were classified for these responses: 'methods of constructing test items,' 'theories of classroom assessment,' and 'item bank.' Of those teachers who responded, 13 (46%) described taking part in training programs related to the methods of constructing test items, such as designing multiple-choice items and essay tests (6 teachers, 21%); constructing test items that align with educational objectives (3 teachers, 11%); developing various performance assessments (3 teachers, 11%)); building test items that measure higher order thinking skills and heuristic abilities (2 teachers, 7%); developing test items with respect to item difficulty and discrimination (2 teachers, 7%); and designing assessments that help measure students' ability in a large class.

This last description is one of the biggest concerns related to assessments in schools in Korea[108]. Three teachers described the activity related to the theory of classroom assessment such as the concepts, purposes, or forms of assessment. Among the three, one teacher emphasized the theories of assessment that can be applied to real classroom activities. Three teachers mentioned the use of item banks that offer a variety of test items. This description also seems to reflect the idea that a large class size results in a low quality of assessment. If history teachers could use item banks that provide a high quality of assessments, they could save administrational time and offer students better assessments.

[108] Middle and high school classrooms in Korea consist of about 35 to 40 students. A history teacher teaches the subject about 20 periods a week, which means that he or she has 10 classrooms of students (2 periods per classroom per week), or who must be taught and assessed.

# 5. SUMMARY, CONCLUSIONS, AND DISCUSSION

## 5.1. Summary

History assessments for both 9th grade middle schools and 10th grade high schools in Korea do not provide a variety of assessment methods, and instead, depending heavily on multiple-choice tests. The assessments also tended to measure lower levels of historical understanding than those required by the objectives, resulting in small proportions of test items whose levels of historical understanding were consistent with those demanded by the objectives. With respect to assessing historical knowledge, the distribution of the test items was not balanced across the content of the objectives, emphasizing factual knowledge instead of cause-effect relationships, background, differences, or significance of past events. Also, the assessments did not thoroughly cover the span of knowledge representative of the curriculum: a number of test items assessed knowledge not emphasized in educational objectives, focusing on trivial names, places, products of historical figures, or courses of events. In addition, the majority of performance assessments failed to assess historical reasoning skills, and, instead, focused on a simple summary of a historic site or historical figures, or on the descriptions of a family history and a TV historical documentary.

Minor differences were found between middle schools and high schools regarding the alignment between historical understanding levels and historical knowledge demanded both in test items and objectives and performance assessments. Test items from high schools were more likely to assess higher levels of historical understanding than those from middle schools. They

featured more items that assessed student abilities to analyze, infer, or evaluate historical materials. However, the tasks for performance assessments for middle schools gave students more opportunities to use a variety of historical information by analyzing, criticizing, or judging. In terms of assessing knowledge, test items from high schools primarily assessed knowledge not emphasized in objectives. With regard to the quality of test items and choices, both types of schools showed similar results.

Most of the items were well developed in terms of formatting and writing test item stems and alternatives. However, several things should be considered: a number of items used negative words in their questions did not provide context, thereby encouraging students to answer by rote memorization of historical knowledge. In addition, a number of questions that used unequal length of choices had clues to the right answer, in terms of the longest choice in particular.

Finally, the results of teacher preparation coursework and teacher professional development activities related to assessments indicated little training in assessment. The data from the survey showed that teachers had little professional development related to theories, methods, or designs of assessments in training courses or activities before and during their professional work. In addition, a number of teachers said that they wanted to learn how to build and use assessments for their lesson plans.

## 5.2. Conclusions and discussion

The results of this study indicating an emphasis on rote memorization and the lack of assessing high levels of historical understanding in middle ad high schools are not surprising; they coincide with the lack of pre-service courses and the absence of teacher professional development activities for classroom assessment. Only approximately 40% of the test items

measured historical knowledge emphasized in the objectives of the curriculum: about one-fourth of the test items assessed the knowledge only somewhat related to the objectives, and abut one-third of the test items assessed the knowledge not required by the objectives. Moreover, the majority of the test items did not allow students to use high level of historical reasoning skills for deep understanding of the past. The results of the current study revealed that some of the middle and high school teachers attempted to develop some assessment items that measure high levels of historical reasoning skills. The results also provided meaningful suggestions and alternatives about the way in which ideas and practices, in relationship to classroom history assessments, could be conceptualized within history education in Korea. Teachers and administrators of teacher education programs will be able to use the findings in planning future teacher preparation and teachers professional development. The following provides a discussion based on the findings of this study.

***Breadth of assessment tools***: The assessments collected from schools were heavily dependent on only one type of item format for assessing historical knowledge: multiple-choice tests, which covered more than 98% of all test items. As some researchers stated, using multiple-choice tests to assess knowledge has many strengths over other assessment methods: covering greater breadth of learning content within a limited time frame; dealing with students' learning outcomes effectively and objectively; measuring a wide range of high level reasoning skills, and consistency in computing test results (Cunningham, 1998; McTighe & Ferrara, 1998; Nitko, 1996; Smith et al., 2001; Wood, 1977). However, these schools featured a number of test items that predominantly assessed the ability to recall knowledge with rote memorization, middle schools in particular (75% for middle schools and 53% for high schools). Using multiple-choice items that assess factual information did not provide a context and appears to be the most

common way for teachers to assess student learning outcomes. This assessment environment may have led students to believe that the most important aspect of history is to memorize historical facts such as names, dates, places, or the courses of actions in the past—information that may not have significant meaning for their own lives.

This type of assessment tool may also lead students to form negative opinions of the purpose of education, providing 'only one correct answer' or 'an already prepared answer,' which may foster a use of absolutistic thinking (Paul, 1991). Although the test items assess historical knowledge that demands a high level of understanding, this type of assessment tool does not allow students to develop their own points of view. Rather, it encourages students to establish fixed judgments about both themselves and others. Historical knowledge is interpretive, uncertain, ambiguous, and even tentative. History class assessments should give students an opportunity to cultivate their own capacity for reasoned judgment about uncertain historical knowledge by offering them a variety of assessment tools. For example, through constructed-response tests, students would be able to show their thinking processes, provide different perspectives, or make conclusions about past events.

With respect to reasoning skills demanded in the performance assessments, some of the assessments appeared to require that students use complex mental abilities in completing given tasks (53% for middle schools and 14% for high schools), but did not cover the broad knowledge required by the curriculum. Considering the weaknesses and strengths of each assessment tool, history teachers should develop and use a variety of assessment tools that would allow for an accurate assessment of students' learning outcomes and thinking processes.

***Assessment of deep understanding of history***: The test items analyzed assessed predominantly the lower levels of historical understanding rather than the levels demanded in the

curriculum. Reasoning skills outlined in the curriculum range from low to high levels. However, actual test items focused heavily on assessing knowledge presented in those objectives with relatively low levels of historical understanding. In most cases, when targeted objectives required the comprehension of historical information from written or visual materials (level 2), most of the items assessed the ability to recall factual knowledge (level 1). This was the case for middle schools, in particular. Overall, a majority of test items from the schools lacked the ability to assess high levels of historical reasoning, including the meaning of historical agents or actions of the past in relation to the social and political background in order to analyze or judge them with reasoned evidence. Instead, the items focused on assessing knowledge of the past usually derived from students' rote memorization skills.

As many practicing scholars report, a deep understanding of history goes beyond a simple understanding of the past that many include merely comprehending the literal meaning of historical accounts and a historical passage as a certainty (Bain, 2000; Leinhardt, et al., 1994a; Perfitti, et al., 1994 & 1995; Rodrigo, 1994; Seixas, 2000; Voss & Wiley, 2000). Deep understanding of history can be described as learning about the past by involving students in gathering, analyzing, and inferring from the materials of the past so that they can enhance their abilities to develop alternative solutions for the problems of the past and construct meaning-making processes for their contemporary lives. A number of studies provide evidence that children gain more knowledge by contextualizing information in real situations rather than through rote memorization (Piaget, 1958; Resnick & Klopfer, 1989; Shepard, 2000). As the results of the current study show, history teachers, in particular those at the middle school level, may have the impression that requiring students to compare, to infer, or to judge historical materials may be far above the capacity of their middle school students. However, evidence was

found by the current study that some history teachers, even in middle schools, tried to develop multiple-choice test items and performance assessments that would assess a students' ability to sustain a line of reasoning by linking multiple sets of historical information in order to draw conclusions. As Schama wrote, "history was not a remote and funereal place. It was a world that spoke loudly and urgently to our concerns" (cited in Leinhardt, 1994c, p. 209). Students should be given more opportunities to learn how to interpret and analyze historical materials, how to hypothesize and question about a past event, or how to evaluate certain historical actions and events. Such activities are the means for students to build real sense and meaning of the past for their own lives and for their own communities.

***Assessment of the "why" or "how" of historical knowledge***: The results of the study suggest that history assessments in Korea prominently emphasize measuring historical knowledge about "what" (more than 70% for both middle and high schools). The content of political, social or economic systems, the definitions or principles of the systems, the course of events, or the activities of politics and military were common test topics. Rare were assessments of the differences in or relationships between A and B, the causes of an event, or aspects of Korea after a certain political event had taken place. Although several test items did assess high levels of reasoning skills, they also focused on the "what" of the past. For example, many test items assessed the ability to infer the way in which social life of people in the Iron Age was presented in a historical passage, but few test items assessed knowledge of "why" or "how" centralized governments could have been established in the Iron Age culture under which background. It appears that the ability to contextualize a variety of historical facts through a broad political or social background that reveals the purpose of an action, the reason for cultural

development, or the cause of political or social change in a society is not assessed by the teachers.

As Shemilt (2000) argues, historical narratives consist not only of 'what happened' but also of 'what was going on' (p. 95). That is, each bit of fact from the past (what happened) is interwoven into a social or political contexts by the explanation of a historian, revealing cause-effect relationships, intentions of people, importance of individuals or groups, or social trends and turning points (what was going on). The historical knowledge constructed by a historian includes in its accounts structures or elements, such as continuity and change, progress and decline, significance for the present, judgment, or power relationships (Seixas, 1996). Thus, in order to comprehend historical knowledge, history teachers themselves must be aware of the basic elements and nature of historical knowledge and be able to develop their abilities to analyze, infer, or form conclusions about the actions of the past. For example, the current study uncovered a failure to assess the causes of two wars in the Chosōn society (Oaeran and Horan) in conjunction with the relationships among three nations (the Chosōn, Kem (Chin), and Japan). Regarding these two wars, specific facts are 'Japan and Kem (Chin) invaded the Chosōn in 1592, 1627, and 1636,' 'to factional politics, the Chosōn court was dominated by two literati, the Easterners and the Westerners, in the late 16$^{th}$ century, ' 'Toyotomi Hideyoshi subdued most of the warlords in Japan in the 16$^{th}$ century,' and 'Jurchen Manchu (Chin) under the Nurhachi strengthened its hegemony in Manchuria and north China in the early 17$^{th}$ century.' In order to assess such knowledge, teachers must know how these facts are interwoven by analyzing the political and social background of each country, making inferences about the political relationships between the countries, and forming conclusions about the multiple causes of the wars.

In addition, as mentioned earlier, questions about the 'why' and 'how' of history (what is going on) are directly related to a historian's interpretation of the events, while 'what' is related to the facts of the past. A historian reconstructs the past through his or her own coherent accounts that impute narrative significance to chronicle events. Thus, 'history' involves interpretation by a historian, including taking different perspectives on a single event, selecting a certain historical date, person, or place to interpret, and placing more emphasis on one aspect rather than another. With respect to such issues, history teachers should offer students a variety of assessment tools that will stimulate and engage students in developing their own points of view by comparing and contrasting different sets of ideas about what occurred during a single event or across events.

***Absence of relevance in assessing content and reasoning skills***: When assessment takes place in education, the content of the assessment should be valid in terms of its relevant, representative, and meaningful. That is, in order for classroom assessment to provide valid score inferences, the test items should be aligned with the instructional goals or curriculum objectives. The 7$^{th}$ National Curriculum for Korean history provides standards of historical knowledge and thinking skills that facilitate meaningful learning outcomes for students. However, results of the current study indicate that only approximately 40% of the test items were aligned with the content of unit objectives: not only did many test items assess content that were only partly related to the objectives (approximately 30% for middle schools and 20% for high schools), but a large proportion of the test items also assessed the content not presented in unit objectives (approximately 30% for middle schools and 40% for high schools). Moreover, a large proportion of the test items (approximately 87% for middle schools and 70% for high schools) did not encourage enough students to use complex reasoning skills for their meaningful

understanding of history. The results of the study provide evidence that most history assessments were neither representative cognitively complex nor aligned with the objectives.

Assessment should reflect instructional goals and activities, but the results of this study indicating that this not a common practice. As Smith et al. (2001) emphasize, without instructional goals and the assessments to reflect them, students merely desire to obtain good grades. Teachers must consider which instructional goals to set and what requirements students will need to achieve them. Then, before administering a test, each test item or assessment task and the test as a whole should be reviewed in order to determine whether it is aligned with learning targets and significant content.

***Teacher preparation programs and professional development activities that focus on assessment*:** The evidence gathered in this study suggests that Korean history teachers are poorly prepared in the area of assessments, assessment being one of the most prevalent instructional activities implemented. According to Stiggins (1991), teachers spend up to 50% of their instructional time in assessment-related activities. And although assessment constitutes such a high proportion of their professional practice, history teachers in this study received only little or no formal assessment training in their professional development activities. A number of teachers (from 46% to 93%, depending on questions) did not remember what they had learned in their assessment classes for teacher preparation in college, classes that may have taken place over 15 or 20 years ago in the experiences of some of the teachers. All of the facts mentioned above suggest that activities regarding assessments in school settings are being ignored. The current trends in classroom assessment in Korea have involved performance assessments, portfolios, and other types of assessment strategies meant to measure students' learning processes, including reasoning and problem-solving skills. These assessment strategies are complex on the part of the

teacher to administer and to score and need teacher training related to both performance assessment and multiple-choice test items. The results of the study indicate that these teachers were not well prepared for developing multiple-choice test items and performance assessments that measure high levels of historical understanding.

As classroom assessment was defined earlier, the purpose of assessment is not just to test what students have learned in class. Assessment should also be the means for providing information about students' learning and for instructional planning. Considering the results of this study, teachers, as educational experts, should be able to develop and use appropriate assessment tools that measure students' learning outcomes. Teachers should also be well versed in grading procedures and in communicating the results of assessments. Assessment should serve as one of the main instructional activities for teachers in helping students' academic development. To provide for this professional need, both formal and informal in-service programs related to assessments should be offered systemically for teachers at both school types and district levels, and include theory and hands-on practice in designing assessment methods and interpreting and using the results.

***Relationship between classroom assessments and College Entrance Examinations***: As presented in the results of assessment analyses, differences were found between middle schools and high schools: high schools provided more multiple-choice test items (13% for middle schools and 28% for high schools) and less performance assessment tasks (53% for middle schools and 14% for high schools) assessing high levels of historical understanding than those provided by middle schools. These results suggest that high school assessments are more affected by the College Entrance Examination than middle school assessments, which consists of only multiple-choice items. However, high school assessments are not affected by the

examination that require high levels of reasoning skills and mastery of broader knowledge: less than 27% of the test items were consistent with or higher than the level of historical understanding required by the objectives, 39% of the test items were not consistent with the content of the objectives. Under the influence of the entrance examination, high school teachers were less likely to require students to complete actual performance assessments, which should make up at least 30% of all assessments. To some extent multiple-choice tests in several high schools replaced performance assessments. In a high-stakes situation, they have shaped education in Korea in undesirable ways—even though the examination assesses students' scholastic aptitude abilities that emphasize the mastery of broad skills. A Korean education scholar, Chung Bom Mo (1994), who adopted Bloom's taxonomy for the first time in the middle of 1950s for Korea argues that high levels of reasoning skills can be practiced when tests offer only multiple-choice tests. Teachers tend to focus on preparing students only for the entrance examination by basing instruction and classroom assessment on the type of content and skills assessed by the test. This situation has discouraged the development of analytical reasoning skills, open-minded thinking, and creativity, and to some extent has lead Korean education to be focused on test taking. The stakeholders—parents, teachers, school administrators, and policy makers—should consider seriously what meaningful learning for students' intellectual growth is and how educational systems could strike a balance between the benefits of education and the desire to be accepted into college via the entrance examination.

***Alignment between curriculum and practices of classroom assessment*:** This study uncovered that the expectations of curriculum, educational objectives in particular, were far above from the practices of classroom assessment. The results of this study are not surprising under the top-down educational system and limited teacher professional development activities

during the implementation the 7th National Curriculum. In general, as mentioned earlier, the Korean education system is highly centralized, which is run by the Ministry of Education and Human Resources Development (MOEHRD). The MOEHRD has the overall responsibility of improving the quality of Korean education, controlling the national curriculum for elementary and secondary schools, and determining the level of content for each subject area or characteristics (An, et al., 1995; Hwang, 1998; Lee, 1993; Shin, & Huh, 1991; Seth, 2002; Yun, et al., 1991). Under the circumstance of the power of central government over the determination of what to teach and which instructional strategies to be used, individual history teachers do not have enough autonomy to decide which historical knowledge is assessed and which assessment tools are used. Once the national curriculum was developed by the national institute (Korea Institute of Curriculum and Evaluation), textbooks, instructional manuals, and the standards for student learning outcomes related to new curriculum were distributed to regional school districts without consideration of input at each school level. Teachers, facing the new national curriculum, in general, had poor information about the curriculum and the intention of the authors of textbooks, so they could not implement it effectively.

Attempts to bridge the gap between the national curriculum and classroom practice are vital in history education. New solutions are necessary for integrating scholarly work (curriculum) and the practice of classroom assessments by collaborative actions between scholars in the discipline of history and practicing history teachers. One study done by McDiarmid and Vinten-Johansen (2000) provide an explicit example of a history teaching methods course taught by a teacher educator and a historian at Michigan State University, attempting the connection between the two. Successful joint inquiry work of scholars and teachers requires broad and multifaceted strategies and the mutual acceptance of each other's

points of reference of what each side brings to the historical inquiry. Each regional school board must be committed to inviting historians, history educators, curriculum developers, and specialists in classroom assessments, so that they can try together for teacher professional development, instructional strategies, and classroom assessment. History teachers need to incorporate current views of student learning and assessment strategies in their instructional practices, sharing professional ideas with colleagues in formal or informal meetings, and participating in in-service programs for improving their own professional development. In order to accomplish these goals, the government must provide each school level and teacher more autonomy to determine the content to teach and to develop his or her own classroom assessments to measure what students know and can do.

# 6. SUGGESTIONS FOR FURTHER RESEARCH

With respect to pedagogical issues related to classroom history assessments, the current study suggests the need for further research related to teachers' competence on history assessments in secondary history classrooms in Korea. The importance of the role of the teacher and the teacher's responsibilities in creating and using assessments of student understanding of historical knowledge and reasoning skills will be facilitated by further advanced research.

The findings of this study imply further research associated with what student teachers learn about classroom assessments in their teacher preparation programs and what history teachers actually know about classroom assessments. This study revealed evidence that some history teachers had taken assessment courses ranging from 2 to 4 credits, and others had never had an assessment course in their teacher preparation programs. Classroom assessment is integral to students' meaningful learning. In order to provide valid classroom assessments that accurately assess the improvement of students' learning, researchers must be able to determine to what extent, during the pre-service period, should student teachers learn about the assessment process in their assessment courses and how their concepts and knowledge of classroom assessment affect their classroom assessment practice in various classroom settings. Researchers must also provide information about the extent to which practicing history teachers know how to design and interpret classroom assessments. In particular, research should be conducted that examines how teachers set up learning targets, how they build assessment methods appropriate for instructional decisions, and how they use the results of their assessments to improve instruction and student learning. These investigations will offer opportunities for university

faculty members, educational policy makers, educators, and school administrators in preparing advanced teacher preparation programs or professional development programs, allowing both student and practicing teachers to develop their own assessment methods.

The practice of history assessments must be influenced by how history teachers conceive the discipline of history. Therefore, this study suggests further research related to history teachers' and student teachers' knowledge of subject matter and reasoning skills and in impact on classroom practices. In fact, because history in secondary schools in Korea is one of social studies, history teachers' majors vary from history, geography, economics, political science, and sociology. Some teachers may never have taken any course related to history in college–other than an introduction to Korean history. Another avenue of investigation should include to what extent do student teachers in social studies methodology courses learn to understand the discipline of history. The investigation into student teachers' learning about the nature of historical inquiry will be one of the important indicators for history teachers' understanding of subject matter and their competence in assessments in their classroom practices. Yeager, Elizabeth, Wilson, and Elizabeth (1997) report that pre-service teachers who had had an historical inquiry approach showed favorable attitudes and reflected that approach in their teaching. In addition, emerging research in the area of history teachers' knowledge of subject matter relates to teachers' competence in the classroom (Downey and Levstik, 1991; Stanley, 1991; Wilson and Wineburg, 1993).

Depending on teachers' perspectives, which may have been influenced by their teacher preparation programs, some teachers may design assessment tools that measure students' understanding of history. For example, a teacher who regards history as a discrete set of facts may develop tests that assess only factual historical knowledge, while a teacher who regards

history as interpretive, multi-causal, or power relationships may use methods that assess the ability to analyze, interpret, or infer historical materials. Factors that other than teachers' perspectives on the discipline of history may also be stronger influences on classroom history assessments, such as educational policies, College Entrance Examination, or a big classroom size. Therefore, it is recommended that research examines what typically occurs in secondary history classrooms; how teachers' historical knowledge relates to the matter of pedagogy; what history teachers actually learn in their professional development activities; and, in particular, how teachers' own concepts of history relate to their history assessments. These investigations will provide important directions, regarding the pedagogical and cognitive implications of the subject matter, history teachers may take, to achieve more active connections with their professional development. History teachers with explicit information about their teaching must have opportunities to attend in-service programs related to the issues and topics of historical reasoning, creating inquiry-based programs, and developing meaningful assessments. Teachers' heightened awareness of the discipline of history and historical reasoning skills will enable them to be more sensitive to the issues of multiple-causations and to the importance of understanding events in a broader context. Further research on history assessments must continue to include a discussion of how history teachers and student teachers in history education are aware of history, how they approach a variety of historical documents and materials, and how they develop history assessments as part of inquiry-based approaches.

# APPENDIX A:

## Consent Letters

Dear History Teacher,

I was a history teacher in Yangsan Middle School. I am now working on my dissertation in Social Studies Education (with a major in history education) at the University of Pittsburgh in the United States.

I would like to invite you to participate in my dissertation study. My study focuses on how Korean history teachers assess historical knowledge in their classrooms. I would like to obtain a copy of all of your multiple-choice tests and performance assessments that you have developed for the 2004 school year. I would also like you to answer a brief survey about your training in assessment. It should take no longer than 15 minutes. At the completion of this study, I will provide you with a summary of the results across the 25 teachers in the study.

Your participation is voluntary. You do not have to take part in this research study and can withdraw from the study at any time if you change your mind. Although there are no payments associated with your participation, your participation will help both you and me better understand the nature of classroom history assessments used in Korea. As a history teacher, I thought it would benefit my students if I had more experience and training in designing classroom assessments. After I obtain my degree and return to Korea, I would like to serve for the Gwangju Board of Education and share my learning and experiences in Korea with teacher colleagues.

All the information obtained for this study will be kept confidential (private). The assessments and responses to the survey will be used for research purposes only. No names will be associated with the data collected. Also, no individual or school names will be used in any publication of the research results. All your rights as a participant of my study will be protected by the Human Subject Protection Advocate of the IRB Office, University of Pittsburgh (001-1-866-212-2668).

Please feel free to contact me if you have any questions or comments. Thank you for your valuable contribution to this research!

Sincerely,

Mi-Sun Kim
University of Pittsburgh
001-1-412-683-5318

Dear Principal,

I was a history teacher in Yangsan Middle School. I am now working on my dissertation in Social Studies Education (with a major in history education) at the University of Pittsburgh in the United States.

I would like to invite you and your 9th (or 10th) grade history teachers to participate in my dissertation study. My study focuses on how Korean history teachers assess historical knowledge in their classrooms. I would like you to collect copies of your 9th (or 10th) grade history teachers' multiple-choice tests and performance assessments that they have developed for the 2004 school year. I would also like the teachers to answer a brief survey about their training in assessment. It should take no longer than 15 minutes. At the completion of this study, I will provide you and the teachers with a summary of the results across the 25 teachers in the study.

Teacher participation is voluntary. They do not have to take part in this research study and can withdraw from the study at any time. Although there are no payments associated with participation, participation will help us better understand the nature of classroom history assessments used in Korea. As a history teacher, I thought it would benefit my students if I had more experience and training in designing classroom assessments. After I obtain my degree and return to Korea, I would like to serve for the Gwangju Board of Education and share my learning and experiences in Korea with teacher colleagues.

All the information obtained for this study will be kept confidential (private). The assessments and responses to the survey will be used for research purposes only. No names will be associated with the data collected. Also, no individual or school names will be used in any publication of the research results. All rights as a participant of my study will be protected by the Human Subject Protection Advocate of the IRB Office, University of Pittsburgh (001-1-866-212-2668).

I have also enclosed a copy of the letter that was sent to the history teacher who works in your school. Please feel free to contact me if you have any questions or comments. Thanks for your school's valuable contribution to this research!

Sincerely,

Mi-Sun Kim
University of Pittsburgh
001-1-412-683-5318

To Whom It May Concern:

It is with pleasure that I write on behalf of Mi-Sun Kim, a native of South Korea and PhD degree candidate in the Social Studies Education program (with a major in history education) at the University of Pittsburgh. Mi-Sun is planning to complete all requirements for the degree and to graduate on August 2005.

Mi-Sun has been working on her dissertation in the field of history education, especially in the area of classroom history assessment. She is interested in studying the nature of history assessments used by teachers in South Korea. I hope you agree to participate in her dissertation study. As a history educator, I believe that Mi-Sun's research will contribute to advancing the state of classroom history assessments in Korea.

I hope you agree with me and support Mi-Sun by participating in her research study. Thank you for your cooperation.

Sincerely,

Suzanne Lane,
Professor of Research Methodology
School of Education
University of Pittsburgh

# APPENDIX B:

## Survey Questions

Thank you for your time in completing this survey. The primary purpose of this survey is to gather information about the assessment activities that are provided to teachers. All data will be kept strictly confidential and will be reported only in aggregate from. No information will be reported at the teacher or school level. Participating schools will have access to the study results. Please return the completed survey in the envelope (sealed) along with your assessments to the principal.

1. Did you have a class/training on assessment in your college coursework on teacher preparation?
   _____Yes _____No

2. If you answered Yes to the above question, please answer the following questions:

   a. How many credits was the assessment class? _____
   b. What percent of the class was spent on the theory of assessment (e.g., validity, reliability)? _____
   c. What percent of the class was spent on how to design multiple choice items? _____
   d. What percent of the class was spent on how to design constructed response items and/or performance assessments? _____
   e. What percent of the class was spent on how to interpret the results and use them for instructional planning? _____
   f. Please indicate below what other topics were covered in the class.

3. While you were a teacher, have you had any professional development activities related to assessment during the <u>past year</u>?
   _____Yes _____No

4. If you answered Yes to the above question, please answer the following questions:

   a. How many hours did you spend on professional development activities related to assessment during the <u>past year</u>? _____
   b. What percent of the activities was spent on the theory of assessment (e.g., validity, reliability)? _____

c. What percent of the activities was spent on how to design multiple choice items? _____
d. What percent of the activities was spent on how to design constructed response items and/or performance assessments? _____
e. What percent of the activities was spent on how to interpret the results and use them for instructional planning? _____
f. Please indicate below what other topics were covered in the activities.

5. Please describe what professional development activities related to assessment would help you in your teaching.

# APPENDIX C:

## The 7th National Curriculum of Social Studies

## THE 7TH NATIONAL CURRICULUM OF SOCIAL STUDIES[109]

### A. Characteristics of Social Studies

Social studies is a subject that helps students develop the qualities of a democratic citizen by being aware of social phenomena with a correct perception, acquire the knowledge of society and skills necessary for social lives, and have values and attitudes required for the members of a democratic society. A democratic citizen that social studies orients will be a person who: 1) has the knowledge necessary to carry on a social life; 2) cultivates democratic values and attitudes, such as the respect of human rights, the mentality of tolerance and compromise, the actualization of social justice, the awareness of community, and the consciousness of participation and responsibility; 3) acquires the capacity to contribute to the development of nation, society, and human race as well as individuals by developing the ability to solve individual and social problems rationally.

Social studies helps students understand and explore social phenomena synthetically by selecting and organizing the followings: the concepts and principles of geography, history, and social sciences; the systems and functions of society, and problems and values of society; and the elements of research methodology and process. Specifically, social studies emphasizes synthetic comprehension of history and activities of our nation based on the understanding of national territory, the historical understanding of current situations, national identity as a Korean and the values and attitudes as a global citizen.

Social studies emphasizes discovering the knowledge of social phenomena by using a

[109] The 7th National Curriculum of Social Studies is translated from *Sahoekwa kyoyuk kwajong: Che chimcha kyoyuk kwajong Kyoyukbu gosi che 1997-15 ho* [Social studies curriculum: The 7th National Curriculum (# 1997-15)] (MOE, 1998, pp. 28-30), *Chunghakgyo kyoyuk kwajong haesōl: Kukō, yunri, sahoe* [Middle school curriculum commentary: Korean, ethics, and social studies] (MOEHRD, 1999, pp. 249-256), and *Kodūnghakgyokyoyuk kwajong haesōl: Sahoe* [High school curriculum commentary: Social studies] (MOEHRD, 2001, pp. 15-21)

variety of information and developing the abilities of critical and creative thinking, problem-solving, and judgment and decision-making. For the development of these abilities, social studies offers the learner opportunities to learn by himself and is oriented toward effective instructional and learning strategies that provide educational experiences appropriate to the individual's learning level, considering his interests and concerns. In addition, social studies considers the characteristics of each region and current events depending on school situations.

In middle school, social studies focuses on discovering and applying knowledge based on the scientific process that is important for each domain, and on demonstrating citizenship by having students acquire the ability to solve individual and social problems so that they (will) voluntarily participate in community lives. Korean history in middle school emphasizes understanding the events or specific activities of the past based on the history of figures and lives that are learned in elementary school.

In high school, based on the knowledge and skills learned in elementary and middle schools, the learner acquires citizenship that actively participate in social problems understanding social phenomena and demonstrating the ability of critical thinking and decision-making.

**B. Objectives of Social Studies**

To understand the features of society and the various situations of the world by acquiring the basic knowledge of social phenomena and skills, and the ability to explore the basic concepts and principles of geography, history, and social sciences. To develop the ability to solve current social problems creatively and rationally by using a variety of information and to participate in community life voluntarily. And after all this, to enhance the quality of a democratic citizen who

is able to contribute to the development of a nation, society, and the human race as well as the individual.

The objectives of each domain

The objectives of each domain consist of six categories.

| Objectives | Domains | Essential elements |
|---|---|---|
| 1 | Unity | ▪ To understand a variety of social phenomena and features synthetically and systematically |
| 2 | Knowledge (‘human and space’) | ▪ To understand reciprocal actions between human and nature<br>▪ To understand the variety of human lives in different residential sites<br>▪ To understand the geographical specificities of regions |
| 3 | Knowledge (‘human and time’) | ▪ To comprehend our historical traditions and the particularity of culture<br>▪ To understand the development of national culture and national history<br>▪ To comprehend the process of development of human lives and the cultural features of each era |
| 4 | Knowledge (‘human and society’) | ▪ To understand the basic knowledge of social lives<br>▪ To understand the basic principles of political, economic, social, and cultural phenomena<br>▪ To comprehend the characteristics of contemporary society and of social problems |
| 5 | Skills | ▪ To acquire, construct, and use the knowledge and information<br>▪ To develop the ability of decision-making, social participation, and rational problem-solving |
| 6 | Value-attitude | ▪ To develop the attitude of democratic life<br>▪ To show the concern for current social problems<br>▪ To acquire the attitude to contribute to the development of national culture and democratic nation |

## C. Assessment of Social Studies[110]

1. Social studies assessment should align with the educational objectives, instructional content, and instructional and learning methods presented in the curriculum.
2. Considering assessment as one part of the educational processes, social studies assessment should be implemented to help individuals' the learning process and their achievement levels and should orient the reciprocal comparison and classification of assessment results.
3. Considering the different achievement level of each learner based on a differentiated curriculum, the learning process of each learner and the changes in his development should be assessed.
4. Social studies assessment should measure learning processes and performances in order to obtain useful information about the learning process to help students develop thinking skills, inquiry-oriented learning, and their outcomes.
5. Assessment methods should include a variety of methods, such as paper-pencil tests, interviews, check-lists, observations, and portfolios.
6. When using multiple-choice tests, the assessment should measure the understanding of basic concepts and principles, the process of acquiring knowledge and information, and the ability of using the knowledge and information, rather than assess the acquisition of knowledge established.
7. Assessment should use both quantitative and qualitative data in order to assess the development of thinking skills and the changes in values and attitudes.

[110] Assessment of social studies is translated from *Sahoekwa kyoyuk kwajong: Che chimcha kyoyuk kwajong Kyoyukbu gosi che 1997-15 ho* [Social studies curriculum: The 7th National Curriculum (# 1997-15)] (MOE, 1998, p. 104).

8. The assessment for each domain of social studies should be carried out with respect to the elements of assessment from the objectives presented in the curriculum.
9. The elements of assessment should be geared toward synthetic and balanced assessment that considers the domain of skills and value-attitude, not just toward measuring the domain of knowledge.
10. The assessment for each domain of knowledge should focus on the acquisition of factual knowledge and a certain degree of understanding of the basic concepts, principles, and generalization necessary for the explanation of social phenomena and problem-solving for an understanding of results of achievement. Qualitative and quantitative assessments should be balanced.
11. The assessment for each domain of skills should focus on measuring the skills of information acquisition and its uses, inquiry, decision-making, and community participation needed for the acquisition of knowledge and life in a democratic society.
12. The assessment for the domain of value-attitude should measure the degree of internalizing the desirable and rational values and the ability to analyze and evaluate values.
13. The results of assessment should be used not only to judge the academic achievement of students, but also to diagnose and assess their learning abilities and the appropriateness of instruction and learning.

## Educational Objectives of Korean History for 9th Grade[111]

### 5. The Establishment and Development of the Chosōn Dynasty[112]

General objectives

- To compare and explain the differences between the Koryō and the Chosōn societies.
- To comprehend the central and local political and educational systems in the Chosōn society.
- To explain the concerns with traditional cultures and the growth of national identity in the early Chosōn period, using specific cases.
- To explain the increase of Sarim power and its political changes.
- To explain the contents and meanings of Chosōn foreign policy toward the neighboring countries, Ming China, Yōjin, and Japan.
- To explain the causes, processes, and effects of Oaeran and Horan, using specific cases.

Specific objectives

- To compare the differences of the Chosōn society from the Koryō society. To explore both their hard line and soft line of foreign policies toward Japan and Yōjin.
- To infer the political meaning of publishing Kyoung-guk Tae-jeon.
- To explore the conversion of foreign policy to Chin-myōung Pae-kem (favoring Ming China and rejecting Kem) after In-jo's coup d'etat and political situations in East-Asia after the two wars.

### 6. The Changes in the Chosōn Society

General objectives

- To comprehend the features of factional politics, using specific cases, and explain its positive and negative influences.

---

[111] Educational objectives of Korean history for 9th grade are translated from *Chunghakgyo kyoyuk kwajong haesōl: Kukō, yunri, sahoe* [Middle school curriculum commentary: Korean, ethics, and social studies] (MOEHRD, 1999, pp. 310-321).

[112] Students learn Korean history up to unit 4 in their grade eight.

- To understand the background and intentions of Yōng-jo implementing the Tang-pyōng (impartiality) policy.
- To understand the social situations and the scholarly dispositions of Shil-hak by studying the various-reformative theories that Shil-hak scholars proposed in order to solve the political and social problems at the end of the Chosōn society.
- To comprehend situations that peasants deployed various types of resistances in their efforts to end against social disorder and taxation corruptions in the Sedo government by the use of specific cases, and to use these facts to understand the impact of Tōng-hak and Catholicism on peasant society.

Specific objectives

- To compare physiocrats to mercantilists.
- To explore the political and social background of the diffusion of Catholicism and Tōng-hak.

**7. The Enlightenment and Independence Movement**

General objectives

- To comprehend the purposes and meanings of reform policies by Hung-sōn Tae-won'gun by analyzing specific cases.
- To understand the characteristics and meanings of the Kanghwa-do Treaty.
- To explain the differences between and meanings of Kae-wha and Wi-jōng Chōk-sa Movements by comparing and analyzing their purposes and activities.
- To understand the characteristics of Kae-wha faction (Enlightenment power) and their reformative purposes by studying the reformative features of the Kab-shin Chōng-pyōn.
- To understand situations that peasants who suffered from economic intervention by foreign power and exploitation by corrupted public officials gathered on Tōng-hak and advanced the modernization movement against foreign power and feudalism.
- To infer that the Kabō Reform involved both its significances and limitations as a modern reform by studying its implementation process and activities.

Specific objectives

- To explore current people's reactions on the isolationist foreign policy (Shōae-guk Chōung-chaek) of Hung-sōn Tae-won'gun.
- To analyze the content of the Kanghwa-do Treaty.
- To infer the characteristics of the Tōng-hak Peasant Movement and the Kabō Reform.

## 8. The Deployment of Movement for National Sovereignty Safeguard

General objectives

- To understand the purposes of Tōng-nip Hyōp-hoe in connection with the activities of Man-min Kong-dong-hoe members.
- To comprehend the establishment of the Tae-han Che-guk and its significance, and the purposes and results of the Gwang-mu Reform.
- To understand the international situation that the Chosōn faced by studying the background and results of the war between Russia and Japan.
- To follow the attitudes of overcoming national crises by understanding the specific facts of a variety of movements for the national sovereignty safeguard deployed against the infringement of sovereignty by Japan.
- To understand that there are differences in the characteristics between Ui-byong and Yaeguk-gyemong Movement, but also similarities in terms of national movement to protect national sovereignty from infringement by Japan.

Specific objectives

- To infer the background and intention behind Tōng-nip Hyop-hoe promoting the enlightenment movement to the people.
- To discuss the ideological limitations of Confucians who led Ui-byong against Japan at the end of the Chosōn Dynasty.
- To explore the significance of the Shil-ryok Yang-song Movement after the Elsa Treaty.

## 9. The National Independence Movement

General objectives

- To explain the main content of Japanese colonial policies during each period and the background of and reasons for the changes in the policies.
- To explain the background, processes, effects, and significances of the 3.1 Movement synthetically.
- To mark and explain the organizations, places, and times on the map in which armed independence resistances acted.
- To understand the aspects and features of various national movements after the 3.1 Movement, using specific facts.
- To explain the relationship between the changes of Japanese colonial policies and national independence struggles.
- To develop a historically accurate chronological table in connection with the lives and contributions of fighters to national independence. To have an attitude following the examples of the fighters' independent spirits and patriotism.

Specific objectives

- To explore the ultimate purposes of the economic policies of the Japanese colonization.
- To compare independence movements before and after the 3.1 Movement.
- To study the situations that national independence movements faced during the struggles against Japan after the 3.1 Movement.
- To explore the actual circumstances and significance of the protection movements for Korean culture during the colonial period.

## 10. The Development of the Tae-han Min-guk

General objectives

- To understand the process of the establishment of the Tae-han Min-guk in the middle of ideological conflicts and disorder following the 8.15 Liberation.
- To explain the background, processes, effects, and influences of the 6.25 War.

- To explore the graft and corruption under the Rhee Syngman Administration and the process of the 4.19 Revolution through the comprehension of the specific situations by interviewing local people and gathering and analyzing data such as newspapers.
- To infer the specific situations of economic growth after the 5.16 Military Coup d'etat by using of diverse-economic-statistical data and understanding the changes of necessities for life. To explain the motives of the economic growth.
- To infer the specific situations of democratic movements under the Yushin System, the 5.18 Democratic Movement, and the June Democratic Resistance. To explain their significance.
- To list the efforts made in order to establish a peaceful reunification after 7.4 South-North Joint Statement. To have an attitude that contributes to a peaceful reunification of our nation.

Specific objectives

- To explore our nation's confrontation concerning the proposal of Trusteeship and the movement of negotiation between the South and the North.
- To discuss the results of the Rhee Syngman administration's maneuvers to grasp political power for a long term.
- To infer the people's awareness of national problems for the last 30 years.

## Educational Objectives of Korean History for 10th Grade[113]

### A. Overall Objectives

- To understand our history subjectively because it is our past and the origins of our national identity.
- To comprehend our history extensively because it is the root of the present and the clue to the future.
- To comprehend our history synthetically because it is the whole of life of our nation.
- To develop the ability to solve problems by improving the ability to analyze, evaluate, and synthesize historical data.
- To have an attitude of actively participating in the creation of a new culture and the development of a society by understanding our history as the process of life.

### B. Objectives of Each Unit

### 1. An Understanding of Korean History

General objectives

- To comprehend the meaning of history in various perspectives.
- To understand the various perspectives of understanding of history and their characteristics.
- To understand the commonality and particularity of national tradition and culture.
- To understand the purposes of studying history.

Specific Objectives

- To understand the historical awareness in the East and the West and the proper judgment on the value of historical materials.
- To explore the commonalities of Korean history to world history and the particularities of Korean history.

[113] Educational objectives of Korean history for 10th grade are translated from *Kodūnghakgyokyoyuk kwajong haesōl: Sahoe* [High school curriculum commentary: Social studies] (MOEHRD, 2001, pp. 71-84).

## 2. Culture of the Prehistoric Age and the Establishment of A Nation

General objectives

- To understand our national living places and the features of our national race and language in the prehistoric age.
- To infer the living features of the prehistoric age by understanding remains and artifacts from that period.
- To comprehend the relationship of the developmental conditions of the culture and the changes of society to the establishment of the nation.
- To understand the background and process of the establishment of the Ko-Chosōn.

Specific objectives

- To organize the process of human development in the prehistoric age systemically.
- To infer the background of the change from the Paleolithic Age to Neolithic Age.
- To explore the social phenomena of the Neolithic Age.
- To infer the relationship between the growth of patriarchal power and social changes.
- To study the record about the establishment of a nation by Tan-gun and comprehend its significance.
- To explore the relationship between the culture of the Iron Age and social changes.

## 3. Administrative Structure and Political Activities

General objectives

- To understand the dominant power of the kings in ancient society.
- To comprehend the development of Koguryō, Paekche, and Shilla as the establishment of ancient nations.
- To understand political characteristics meaning that the establishment of Koryō turned its society into a medieval nation.
- To Recognize the modern elements present in the Chosōn society in the last 18$^{th}$ century.

Specific objectives

- To infer the characteristics and differences between ancient nations in the East and the West.
- To understand the development of the three kingdoms in relation to the changes in Chinese societies and the activities of northern nations.
- To analyze and infer the background of Shilla's unification of the three kingdoms in various perspectives.
- To explore the influence of the Kolp'um system on the political and social problems in Shilla society.
- To infer the particularities and differences of medieval societies in the East and the West.
- To comprehend that the establishment of Koryō marked the beginning of medieval society.
- To infer the reformative administration of King Kong-min in relation to national and international political situations.
- To compare between the modern era of the Eastern society to the Western society.
- To explore the centralized policies of the Chosōn Dynasty.
- To explore the modern elements presented at the end of the Chosōn society.
- To explore the features and problems in politics at the end of the Chosōn society.

**4. Economic Structure and Life**

General objectives

- To understand that the economic life of our nation was agriculture-based since early times.
- To understand institutionalized management of labor force and productive resources in ancient periods.
- To understand that trade and agriculture during the Koryō period was well developed.
- To understand circumstances that the Chosōn Dynasty reinforced an agriculture-first policy grounded in the Confucian ideology reinforced by the Chosōn Dynasty.

- To understand that in the last 18th century, economy activation was improved through an increased productive capacity and a brisk market.

Specific objectives

- To explore the process of growth where a sea power became a political power at the end of the Shilla Kingdom.
- To explore the background that monastery could participate in manual industries and commerce during the Koryō era.
- To analyze the problems of the Kwa-jeon[114] system.
- To infer the influence of Neo-Confucianism as an administrative ideology in the industrial policies of the Chosōn society.
- To infer that a germination of Capitalism was introduced to each industry.
- To understand the relationship between the division of the peasant class and the conversion of the economy to Capitalism.

## 5. Social Structure and Life

General objectives

- To understand that in an ancient society social strata were formed, and familial social status was regarded with great importance rather than an individual's ability.
- To understand that the Kol-pum system was established in the process of Shilla's growth as a centralized ancient nation.
- To comprehend that Munbol aristocracy (a noble lineage) was regarded with great importance during the Koryō era in accordance with the consolidation of social stratification.
- To explain the social structure of Chosōn in connection with the order to Confucianism.
- To understand that social structures were changed in the late 18th century in accordance with the agitation of social status order, and the movements of rising social status were actively advanced.

[114] Land distribution system during the Chosōn Dynasty and the financial background of the ruling class.

Specific Objectives

- To explore the background of the Kol-pum system that could be maintained in an ancient society.
- To infer that the society of the Koryō era was an open one.
- To explore the efforts to rationalize the administrating order of the ruling class during the Chosōn era.
- To infer the relationship between the Sarim power and the Confucian clan rules.
- To explore the relationship between the social constitutions and social changes.
- To infer the thoughts that influenced social changes.

## 6. The Development of National Culture

General objectives

- To understand the establishment of process and features of national culture.
- To understand that Confucianism and Buddhism influenced the development of a national ancient culture.
- To comprehend that during the Koryō era, a higher level of culture was established through the prevalence of Daoism and the theory of geomancy as well as Confucianism and Buddhism.
- To understand the significance of the creation of Hangul with respect to the development of national culture.
- To comprehend the learning and arts of the Chosōn era in relation to a governing order.
- To understand the elements of modern, national, and popular cultures presented during the quickening period of modern society.

Specific Objectives

- To infer the influences of Buddhism on ancient society and culture.
- To infer political, social, and cultural influence of Buddhism during the Koryō period.
- To explore the positive and negative aspects of the Sarim culture.

- To infer the relationship between the awareness of the ruling class on current situations and the development of culture.
- To explore the elements of popular and Korean culture during the quickening period of modern society.

## C. Assessment

1. To evaluate students using the assessment elements from the curriculum objectives.
2. To evaluate the degree of historical knowledge achieved the learning content, the concept comprehension, and historical thinking and problem-solving.
3. To evaluate the ability to analyze data that is used to interpret and evaluate by using a variety of instructional materials.
4. To evaluate historical heuristic skills and the changes of attitudes that are difficult to measure through paper-pencil tests by using a variety of methods, such as observation and homework.
5. To use the results of assessment as data to improve instruction and learning by evaluating the degree and element of achievement through frequent and various assessments.

## REFERENCES


American Psychological Association, American Educational Research Association & National Council on Measurement in Education (APA, AERA, & NCME) (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

An, G., Lee, S., Cho, Y., Yu, B., Lee, G., Han, G., Lee, Y., Kim, Y., Namgung, Y., & Jong, S. (1995). *Hanguk kūn-hyōndae kyoyuksa* [The history of Korean modern and contemporary education]. Seoul, Korea: Hanguk chōnshin munhwasa yōngu.

An, G., Jong, Ch., Kim, Ch., Choe, J., Shin, H., Jong, I., & Kim, Y. (1998). *Hanguk KyoyukKaehyōk ui chōngchihak* [The politics of Korean education reform]. Seoul, Korea: Hakjisa.

Anderson, L. W., Krathwohl, D. R., Airsian, P. W., Cruikshank, K. A., Mayor, R. E., Pintrich, P. R., Rath, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Now York: Longman.

Andretti, K. (1993). *Teaching history through primary evidence*. London: David Fulton.

Airasian, P. W. (1994). *Classroom assessment* (2nd Ed.). New York: McGraw-Hill.

Airasian, P. W. (2000). *Assessment in the classroom: A concise approach*. New York: McGraw-Hill Higher Education

Ashby, R. and Lee, P. (1987). Children's concepts of empathy and understanding in history. In C. Portal (Ed.), *The History Curriculum for Teachers* (pp. 62-88). London: The Falmer Press.

Bain, R. B. (2000). Into the breach: Using research and theory to shape history instruction. In P. N. Stearns, P. Seixas & S. Wineburg (Eds.*), Knowing, Teaching, and Learning History* (pp. 331-352). New York: New York University Press.

Baker, E. L. (1994). Learning-based assessments of history understanding. *Educational Psychologist, 29*(2), 97-106.

Barker, D., & Hapkiewicz, W. G. (1979). The effects of behavioral objectives on relevant and cidental learning at two levels of Bloom' taxonomy. *The Journal of Educational Research*, *72*(6), 334-339.

Bartul, J. (1993). Teaching the value of inquiry through the essay question. In R. Blackey (Ed.), *History a new: Innovations in the teaching of history today* (pp. 85-101). Long Beach, CA: The University Press.

Beck, I. L. & McKeown, M. G. (1994). Outcomes of history instruction: Paste-up accounts. In M. Carretero & J. Voss (Eds.), *Cognitive and Instructional Processes in History and the Social Sciences* (237-256). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Bhola, D. S., Impara, J. C., & Buckendahl, C. W. (2003). Aligning tests with states' content standards: Methods and issues. *Educational Measurement: Issues and Practice, 22*(3), 21-29.

Black, P. (1998). *Testing: friend or foe? Theory and practice of assessment and testing*. London: Falmer Press.

Blackey, R. (1993). A guide to the skill of essay construction in history. In R. Blackey (Ed.), *History a new: Innovations in the teaching of history today* (pp. 49-58). Long Beach, CA: The University Press.

Blanco, F. & Rosa, A. (1997). Dilthey's dream: Teaching history to understand the future. In Voss, J. M (Ed.), Explanation and Understanding in Learning History. *International Journal of Education Research, 27*(3), 198-200.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1987). *Taxonomy of educational objectives: The classification of educational goals*. Now York: Longman.

Boix-Mansilla, V. (2000). Historical understanding: Beyond the past and into the present. In P. N. Stearns, P. Seixas & S. Wineburgh (Eds.), *Knowing, Teaching, and Learning History* (pp. 390-418). New York: New York University Press.

Bong, M. (2003). *Choices, evaluations, and opportunities for success: Academic motivation of Korean adolescents*. In F. Pajares & T. Urdan, International perspectives on adolescence (pp. 319-341). Greenwich, CT: Information Age Publishing.

Booth, M (1987). Ages and concepts: A critique of the Piagetian approach to history teaching. In C. Portal (Ed.), *The History Curriculum for Teachers* (pp. 62-88), London: The Falmer Press.

Britt, M. A, Charles, A., Perfetti, C. A., Julie A., Van Dyke, J. A., & Gabrys, G. (2000). The sourcer's apprentice: A tool for document-supported history instruction. In P. N. Stearns, P. Seixas & S. Wineburg (Eds.), *Knowing, Teaching, and Learning History* (pp. 437-470). New York: New York University Press.

Carr, E. H. (1961). What is history? London Great Britain: New York. St Martin's Press.

Carretero, M. & Voss, J. F. (1994). *Cognitive instructional processes in history and the social sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Choi, Y. (2000). Kusōngjuūi Haksup ūl wihan yōksa naeyong gusōng bangan [Alternative of the content structure of history textbook for constructivist learning. *Yōksa kyoyuk, 73*, 178-197.

Chung, B. (1994). The taxonomy in the Republic of Korea. In L. W. Anderson, & L. A. Sosniak (Eds.). (1994). *Bloom's taxonomy: A forty-year retrospective: Ninety-third yearbook of the National Society for the Study of Education* (pp. 164-173). Chicago: University of Chicago Press.

Cizek, G. J. (1997). Learning, achievement, and assessment: Constructs at crossroads. In G. D. Phye (Ed.), *Handbook of Classroom Assessment* (pp. 1-32). New York: Academic Press.

Collingwood, R. G. (1972). *The idea of history*. New York: Oxford university Press.

Collins (1995). *English dictionary*. London: Collins.

Costa, A. L. (1991). Mediating the metacognitive. In A. L. Costa (Ed.), *Developing minds: A resource book for teaching thinking*. Alexiandria, VA: Association for Supervision and Curriculum Development.

Crocker, L. (2003). Teaching for the test: Validity, fairness, and moral action. *Educational Measurement: Issues & Practice, 22*(3), 5-11.

Cunningham, G. K. (1998). *Assessment in the classroom: constructing and interpreting tests*. Washington, D.C.: The Falmer Press.

Daniels, R. V. (1972). *Studying history: How and why*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Danto, A. C. (1965). *Analytical philosophy of history*. Cambridge: University Press.

Darling-Hammond, L., & Ancess, J. (1996). Authentic assessment and school development. In J. B. Baron & D. P. Wolf, *Performance-based student assessment: Challenges and possibilities: Ninety-fifth yearbook of the national society for the study of education*. Chicago, IL: The University of Chicago Press.

Darling-Hammond, L., Ancess, J., & Falk, B. (1995). The new Bronx school: Weaving assessment into the fabric of teaching and learning. In *Authentic assessment in action: Studies of schools and students at work.* New York: Teachers College Press.

Downing, S. M., Dawson-Saunders, B., Case, S. M., & Powell, R. D. (1991). *The psychometric effects of negative stems, unfocused questions, and heterogeneous options on NBME Part I and Part II item characteristics*. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. Baron & R. Sternberg (Eds.), *Teaching Thinking Skills: Theory and Practice*. New York: Freeman.

Ennis, R. H. (1991). Goals for a critical thinking curriculum. In A. L. Costa (Ed.), *Developing minds: A resource book for teaching thinking*. Alexiandria, VA: Association for Supervision and Curriculum Development.

Fairbrother, R. W. (1975). The reliability of teachers' judgments of the ability being tested by multiple-choice items. *Educational Researcher, 17*(3), 202-210.

Fines, J. (1987). Making sense out of the content of the history curriculum. In C. Portal (Ed.), *The History Curriculum for Teachers* (pp. 103-115), London: The Falmer Press.

Fosnot, C. T. (1996). *Constructivism: theory, perspectives, and practice*. New York: Teachers College, Columbia University.

Foucault, M. (1972). *The archaeology of knowledge*. Translated from the French by A. M. Sheridan Smith. New York: Pantheon Books.

Fountain, G., & Fusco, E. (1991). A strategy to support metacognitive processing. In A. L. Costa (Ed.), *Developing minds: A resource book for teaching thinking*. Alexiandria, VA: Association for Supervision and Curriculum Development.

Friedman, W. J. (1982). Conventional time concepts and children's structuring of time. In W. J. Friedman (Ed.), *The Developmental Psychology of Time*. London: Academic Press.

Fulbrook, M (1999). *National identity and German history. In German national identity after the holocaust.* Cambridge, UK: Polity Press.

Furst, E. J. (1994). Bloom's taxonomy: Philosophical and educational issues. In L. W. Anderson, & L. A. Sosniak (Eds.). (1994). *Bloom's taxonomy: A forty-year retrospective: Ninety-third yearbook of the National Society for the Study of Education* (pp. 28-40). Chicago: University of Chicago Press.

Glaser, R. & Silver, E. (1994). Assessment, testing, and Instruction: Retrospect and prospect. *Review of Research in Education, 20*, 393-419.

Good, T. L., & Brophy, G. E. (1986). *Educational psychology* (3rd Ed.). New York: Longman.

Grant, S. G. (2003). *History lessons: Teaching, learning, and testing in U.S. high school classrooms*. Mahwah, NJ: Lawrence Erlbaum Associates.

Greene, S. (1994). The problems of learning to think like a historian: Writing history in the culture of the classroom. *Educational Psychologist*, *29*(2), 89-96.

Hahn, C. L. (1994). Controversal issues in history instruction. In M. Carretero & J. Voss (Ed), *Cognitive and Instructional Processes in History and the Social Sciences* (pp. 201-219). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Haladyna, T. M., & Downing, S. M. (1989). A taxonomy of multiple choice item-writing rules. *Applied Measurement in Education. 2*(1), 51-78.

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item writing guidelines for classroom assessment. *Applied Measurement in Education, 15*(3), 309-334.

Ham, J. (2003). *Hanguk Kyoyuk Kwajong Pyōchōnsa Yōngu: Chosōnjo malbutō che 7 cha kyoyuk kwajongki kkaji* [The study of the changes of Korean curriculum]. Seoul, Korea; Kyoyuk Kwahaksa.

Hambleton, R. K., & Murphy, E. (1992). A psychometric perspective on authentic measurement. *Applied Measurement in Education, 5*, 1-16.

Harnett, P. (1998). Identifying progression in children's understanding: The use of visual materials to assess primary school children's learning in history. *Cambridge Journal of Education*, *23*(2), 137-154.

Hill, P. W., & McGaw, B. (1981). Testing the simplex assumption underlying Bloom's Taxonomy. *American Education Research Journal, 18*, 92-101.

Holt, T. (1990). *Thinking historically: Narrative, imagination and understanding*. New York: College Entrance Examination Board.

Hoodless, P. (1996). *Time and timelines*. London: Historical Association.

Husband, C. (1996). *What is history teaching?: Language, ideas and meaning in learning about the past*. Philadelphia, PA: Open University Press.

Hvolbek, R. H. (1993).History and humanities: Teaching as destructive of certainty. In R. Blackey (Ed.), *History a new: Innovations in the teaching of history today* (pp. 3-9).

Hwang, G. (1998). *Current trend and issues in Korean school curriculum development*. National Institute for Educational Research of Japan.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence. in English* (Ed.), Routledge and Kegan Paul.

Jacobs, H. H. (1997). Redefining assessment. *Alternative Assessment: Social Studies Educator's Handbook* (pp. 2-5). Upper Saddle River, NJ: Prentice Hall.

Jong, S. (2001). Haksūpja Jungshim yoksa kyoyuk ui kwaje wa banghyang [The task and direction of learner-centered history education]. *Yoksa kyoyuk nonjip, 26*, 29-55.

Ju, Y. (2001). Che chilcha kyoyuk kwajōng ūl tonghaebon yōksa suōp hwangyōng ui kujo wa sōngkyōk [The structures and characteristics of history teaching environment in terms of the 7th Curriculum]. *Yōksa kyoyuk nonchip, 26*(2), 143-195

Kim, D. (2000a). Hanguksa yōngu wa kuksa kyoyuk ui banghyang [The study of Korean history and the direction of history education]. *Yōksa kyoyuk, 76*, 1-33.

Kim, D., & Lee, B. (1998). The impacts of college entrance examination on the curriculum implementation and classroom teaching-learning activities in Korea's high schools. *Korean Journal of Curriculum and Evaluation, 1*(1), 1-15.

Kim, Y. (2000b). The 21st century's vision of the Korean teaching profession: Issues and policy plans. *Asia-Pacific Journal of Teacher Education and Development, 3*(1), 35-54.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212-218.

Kottke, J. L., & Schuster, D. H. (1990). Developing tests for measuring Bloom's learning outcomes. *Psychological Reports, 66*, 27-32.

Kunen, S., Cohen, R., & Solomon, R. (1981). A levels of processing analysis of Bloom's Taxonomy. *Journal of Educational Psychology, 73*, 202-212.

Kuhn, D., Michael, W., & Flaton, R. (1994). Historical reasoning as theory-evidence coordination. In Carretero, M. & Voss, J. (Eds.), *Cognitive and instructional Processes in History and the Social Sciences* (pp. 377-402). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lane, S., Parke, C. S., Stone, C. A., Hansen, M. A., & Cerrillo, T. L. (2000). *MSPAP impact study: Social Studies*. Maryland Assessment System Project supported by a grant from the U.S. Department of Education, Assessment Development and Evaluation Grants Program.

Lane, S., Parke, C. S., & Stone, C. A. (2002). The impact pf a state performance-based assessment and accountability program on mathematics instruction and student learning: Evidence from survey data and school performance. *Educational Assessment, 8*(4), 279-315.

Lane, S. & Stone, C. A. (in press). Performance assessment.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, England: Cambridge University Press.

Lee, H. (1993). *Kyoyuk kwajong yongu* [Curriculum Inquiry]. Seoul, Korea: Pakyōungsa.

Leinhardt, G., Stainton, C. & Virji, S. M. (1994a). A sense of history. *Educational Psychology*, *29*(2), 79-88.

Leinhardt, G., Weidman, C., Virji, S. M., & Odoroff, E. (1994b). Learning to reason in history: Mindlessness to mindfulness. In Carretero, M. & Voss, J. (Eds.), *Cognitive and Instructional Processes in History and the Social Sciences* (pp. 131-158). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Leinhardt, G. (2000). Lessons on teaching and learning in history from Paul's pen. In P. N. Stearns, P. Seixas & S. Wineburg (Eds.), *Knowing, Teaching, and Learning history* (pp. 223-245). New York: New York University Press.

Lerner, G. (1997). *Why history matters: Life and thought*. New York Oxford: Oxford University press.

Levstik, L. S., & Barton, K. C. (2001). *Doing history: Investigating with children in elementary and middle schools*. London: Lawrence Erlbaum Associates, Publishers.

Lichtman, A. J., & French, V. (1978). *Historians and the living past*. Arlington Heights, Il: Ahm Publishing Corporation.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15-21.

Linn, R. L., & Gronlund, N. E. (1995). *Measurement and assessment in teaching* (7th Ed.). Englewood Cliffs, NJ: Prentice Hall.

Lomas, T. (1993). *Teaching and assessing historical understanding*. London: Historical Association.

Lynn, S. (1993). Children's reading pictures: History visuals at key stages 1 and 2. *Education 3-13, 21(3*), 23-29.

Macintosh, H. (1987). Testing skills in history. In C. Portal (Ed.), *The History Curriculum for Teachers* (pp. 183-219), London: The Falmer Press.

Masterman, E., & Rogers, Y. (2002). A framework for designing interactive multimedia to scaffold young children's understanding of historical chronology. *Instructional Science, 30*, 221-241.

Mathien, T. (1991). History and the moralist. *The Monist, 74*(2), 240-267.

Marzano, R. J. (2001). Designing a new taxonomy of educational objectives. Thousand Oaks, CA: Corwin Press, Inc.

Marzano, R. J., Brandt, R. S., Hughes, C. S., Jones, B. F., Presseisen, B. Z., Rankin, S. C., & Suhor, C. (1988). *Dimensions of thinking: A framework for curriculum and instruction*. Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, R. J., Pickering, D., & McTighe, J. (1993). *Assessing student outcomes: Performance assessment using the dimensions of learning model*. Alexandria, VA: Association for Supervision and Curriculum Development.

McMillan, J. H. (1997). *Classroom assessment: Principles and practice for effective instruction*. Boston, MA: Allyn and Bacon.McMillan, J. H. (1997). *Classroom assessment: Principles and practice for effective instruction*. Boston, MA: Allyn and Bacon.

McMillan, J. H. (2000). *Essential assessment concepts for teachers and administrators*. Thousand Oaks, CA: Corwin Press, Inc.

McTighe, J. & Ferrara, S. (1998). *Assessing learning in the classroom*. National Education Association.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher, 18*(2), 5-11.

Messick, S. (1995). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice, 14*(4), 5-8.

Minstry of Education (1996). *Education in Korea: 1995-1996*. Seoul, Korea: National Institute for Educational Research & Training.

Ministry of Education (MOE, 1998). *Sahoekwa kyoyuk kwajong: Che chimcha kyoyuk kwajong Kyoyukbu gosi che 1997-15 ho* [Social studies curriculum: The 7th National Curriculum (# 1997-15)]. Seoul, Korea: Daehan kyokwasō.

Ministry of Education and Human Resources Development (MOEHRD, 1999). *Chunghakgyo kyoyuk kwajong haesōl: Kukō, yunri, sahoe* [Middle school curriculum commentary: Korean, ethics, and social studies]. Seoul, Korea: Daehan kyokwasō.

Ministry of Education and Human Resources Development (MOEHRD, 2001). *Kodūnghakgyo kyoyuk kwajong haesōl: Sahoe* [High school curriculum commentary: Social studies]. Seoul, Korea: Daehan kyokwasō.

Ministry of Education and Human Resources Development (MOEHRD, 2003a). *Junghakgyo kuksa*. Seoul, Korea: Kuksa p'yoch'an wiwonhoe.

Ministry of Education and Human Resources Development (MOEHRD, 2003a). *Kodūnghakgyo kuksa*. Seoul, Korea: Kuksa p'yoch'an wiwonhoe.

Ministry of Education and Korean Educational Development Institute (MOE & KEDI: 2000). *Kyoyuk tonggye yōnbo* [Statistical yearbook of education]. Seoul, Korea: Hanguk kyoyuk gaebalwon.

Mitchell, R. (1992). *Testing for learning: How new approaches to evaluation can improve America's schools*. New York: The Free Press.

National Center (1994). *National standards for world history: Exploring paths to the present*. Los Angeles, CA: National Center for History in the Schools.

Nitko, A. J. (1996). *Educational assessment of students* (2nd Ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.

Oxford (1993). *Concise Oxford dictionary*. Oxford: Oxford University Press.

Paul, R. (1990). *Critical thinking: What every person needs to survive in a rapidly changing world*. Rohnert Park, CA: Center for Critical Thinking and Moral Critique.

Paul, R. & Elder, L. (2001). *The miniature guide to critical thinking concepts & tools*. Foundation for Critical Thinking.

Paul, R. W. (1985). Bloom's taxonomy and critical thinking instruction. *Educational Leadership*, 36-39.

Paul, R. & Elder, L. (2002). Critical thinking: *Tools for taking charge of your professional and personal life*. Upper Saddle River, NJ: Publishing as Financial Times Prentice Hall.

Perfetti, C. A., Britt, M. A., Rouet, J., Georgi, M. C. & Mason, R. A. (1994). How students use texts to learn and reason about historical uncertainty. In M. Carretero & J. Voss (Eds.), *Cognitive and Instructional Processes in History and the Social Sciences* (pp. 257-284). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Perfetti, C. A., Britt, M. A., & Georgi, M. C. (1995). *Text-based learning and reasoning*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Phye, G. D. (Ed.) (1997). *Handbook of classroom assessment: Learning, adjustment, and achievement*. New York: Academic Press.

Poole, R. L. (1972). Characteristics of the taxonomy of educational objectives, cognitive domain: A replication. *Psychology in the Schools, 9*(1), 83-88.

Portal, C. (1987). Empathy as an objective for history teaching. In C. Portal (Ed.), *The History Curriculum for Teachers* (pp. 89-99). London: The Falmer Press.

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher, 31*(7), 3-14.

Pratt, D. (1974, May). The functions of teaching history. *History Teacher*, 7(3), 410-425.

Presidential Commission for New Education Community (PCNEC) (2000). *Iship-il segi hanguk kyoyuk ui kaehyōk banghyang gwa kwaje* [Reform direction and task of Korean education in the 21st century*]*. Seoul, Korea: Presidential Commission for New Education Community.

Quellmalz, E. S. (1991). Needed: Better methods of testing higher-order thinking skills. In A. L. Costa (Ed.), *Developing minds: A resource book for teaching thinking*. Alexandria, VA: Association for Supervision and Curriculum Development.

Quellmalz, E. S. (1991). Developing criteria for performance assessments: The missing link. *Applied Measurement in Education, 4*(4), 319-331.

Quellmalz, E. S. (1985). Needed: Better methods for testing higher-order thinking skills. *Educational Leadership, 43*(2), 29-35.

Quellmalz, E. S., & Hoskyn, J. (1997). Classroom assessment of reasoning strategies. In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, adjustment, and achievement*. New York: Academic Press.

Reid, D. K., & Stone, C. A. (1991). Why is cognitive instruction effective? Underlying learning mechanisms. *Remedial and Special Education, 12*(3), 8-19.

Research Committee for the Revision of Social Studies Curriculum (1997). *Research and development of a proposal for the revision of the 7th national curriculum*. The Korea National University of Education.

Resnick, L. B., & Klopfer, L. E. (1989). Toward the thinking curriculum: An overview. In L. B. Resnick, & L. E. Klopfer (Eds.), *Toward the thinking curriculum: Current cognitive research* (pp. 1-18). Alexandria, VA: Association for Supervision and Curriculum Development.

Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction* (pp. 37-75). Totowa, NJ: Barnes & Noble.

Robert, D. M. (1993). An empirical comparison of three- and four-choice items and tests: Susceptibility to testwiseness and internal consistency reliability. *Educational and Psychological Measurement, 59*, 234-247.

Rodrigo, M. J. (1994). Promoting narrative literacy and historical literacy. In M. Carretero & J. Voss (Eds.), *Cognitive and Instructional Processes in History and the Social Sciences* (pp. 309-320). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Rogers, P. (1987). History—the past as a frame of reference. In C. Portal (Ed.), *The History Curriculum for Teachers* (pp. 1-22). Philadelphia: The Falmer Press.

Rosa, A. (1994). How do people consume history for?: Learning history as a process of knowledge consumption and construction of meaning. In M. Carretero & J. Voss (Eds.), *Cognitive and Instructional Processes in History and the Social Sciences* (pp. 221-233). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Rothman, R. Slattery, J. B., & Vranek, J. L. (2002). *Benchmarking and alignment of standards and testing*. CSE Technical Report 566. Achieve, Inc.

Ruder, L. M., & Boston, C. (1994). Performance assessment. *The ERIC Review, 3*(1), 2-12.

Sansom, C. (1987). Concepts, skills and content: A development approach to the history syllabus. In C. Portal (Ed.), *The History Curriculum for Teachers* (pp. 116-141), London: The Falmer Press.

Scott, A. M. (1993). Life is a multiple-choice question. In R. Blackey (Ed.), *History a new: Innovations in the teaching of history today* (pp. 3-9). Long Beach, CA: The University Press.

Seixas, P. (1996). Contextualizing the growth of historical understanding. *Handbook of Education and Human Development* (pp. 765-783).

Seixas, P. (1999). Beyond 'content' and 'pedagogy': in search of a way to talk about history education. *Journal of Curriculum Studies, 31*(3), 317-338.

Seixas, P. (2000). Schweigen! Die Kinder! Or does postmodern history have a place in the schools? In P. N. Stearns, P. Seixas & S. Wineburg (Eds.), *Knowing, Teaching, and Learning history* (pp. 19-37). New York: New York University Press.

Seo, O. (2001). P'ostmodern shidae ūi hanguksa inshik kwa kuksa kyoyuk [The awareness of Korean history and Korean history education in postmodern era]. *Yoksa kyoyuk, 80*, 1-30.

Seo, O. (2000). Hanguksa inshik kwa kuksa kyoyuk ūi mokchōk [The awareness of Korean history and the purpose of history education. *Yoksa kyoyuk, 76*, 35-61.

Seth, M. J. (2002). *Educational fever: Society, politics, and the pursuit of schooling in South Korea*. Honolulu, Hawaii: University if Hawaii.

Shemilt, D. (1984). Beauty and the philosopher: empathy in history and classroom. In A. K. Dickinson, P. J. Lee, and P. J. Rogers (Eds.), *Learning History.* Heinemann Educational Books.

Shin, S. & Huh, K. (1991). Curriculum development in South Korea. In C. Marsh & P. Morris (Eds.), *Curriculum Development in East Asia*. New York: The Falmer Press.

Shemilt, D. (2000). The Caliph's Coin: The currency of narrative frameworks in history teaching, In P. N. Stearns, P. Seixas & S. Wineburg (Eds.), *Knowing, Teaching, and Learning history* (pp. 83-101). New York: New York university Press.

Shemilt, D (1987). Adolescent ideas about evidence and methodology in history. In C. Portal (Ed.), *The History Curriculum for Teachers* (pp. 62-88), London: The Falmer Press.

Sinatra, G., Beck, L., & Mckeown, M. (1992). A longitudinal characterization of young students' knowledge of their country's government. *American Educational Research Journal, 29*.

Shepard, L. A. (1991). Interview on assessment issues with Lorrie Shepard. *Educational Researcher, 20*(3), 21-23.

Shepard, L. A. (1991). Psychometricians' beliefs about learning. *Educational Researcher, 20*(6), 2-16.

Shepard, L. A. (2000). *The role of classroom assessment in teaching and learning*. CSE Technical Report 517.

Shim, S. (1998). *Hanguk kyoyuk ūi saeroun mosaek* [New alternative of Korean education]. Seoul, Korea: Naeil ūl yōnūn chaek.

Smith, J. K., Smith, L. F., & De Lisi, R. (2001). *Natural classroom assessment: Designing seamless instruction & assessment.* Thousand Oaks, CA: Corwin Press, Inc.

Song, C. (1999a). Yōksa kyoyuk yōngu ūi sōngkwa wa kwaje [The fruits and problems of the study in history education]. *Yōksa kyoyuk, 71*, 1-33.

Song, S. (1999b). Iron yōksa kyoyuk ūi sōngkwa wa hangye [The fruits and limitations of theoretical history education]. *Yōksa kyoyuk. 70,* 160-167.

Song, I. (1998). Kwangbok ihu hanguk ūi yōksa kyoyuk ūi donghyang punsuk [The analysis of the research tendency in history education in Korea after it liberation]. *Sahoikwa kyoyuk*, *31*, 407-430.

Sorensen, C. (1994). Success and education in south Korea. *Comparative Education Review, 38*(1), 10-35.

Stanley, W. B. (1991). Teacher competence for socials studies. In J. P. Shaver (Ed.), *In Handbook of research on socials studies teaching and learning: A project of the National Council for the Social Studies* (pp. 249-62). New York: Macmillan.

Stearns, P. N. (2000). Getting specific about training in historical analysis. In P. N. Stearns, P. Seixas & S. Wineburg (Eds.), *Knowing, Teaching, and Learning history* (pp. 419-436). New York: New York University Press.

Stearns, P. N., Seixas, P., & Wineburg, S. (Eds.) (2000). *Knowing, Teaching, and Learning History*. New York: New York University Press.

Stern, L. & Ahlgren, A. (2002). Analysis of students' assessments in middle school curriculum materials: Aiming precisely at benchmarks and standards. *Journal of Research in Science Teaching, 39*(9), 889-910.

Stiggins, R. J. (1994). Student-centered classroom assessment. New York: Macmillan College Publishing Company.

Stiggins, R. J., Rubel, E., & Quellmalz, E. (1988). *Measuring thinking skills in the classroom* (Revised edition). Washington, D.C.: National Educational Association.

Stow, W. & Haydn, T. (2000). Issues in the teaching of chronology. In J. Arthur & R. Phillips, *Issues in History Teaching* (pp. 83-97). New York: Routledge.

Thornton, S. & Vukelich, R. (1988). Effects of children's understanding of time concepts on historical understanding. *Theory and Research in Social Education, 16*(1), 69-82.

United Nations Development Programme (2001). Human development indicators. Retrieved May 23, 2002 from http://www.undp.org/hdr2001/back.pdf

UNESCO (2002). Statistics: Enrolment in tertiary education. Retrieved July 30, 2004 from http://www.uis.unesco.org/TEMPLATE/html/Exceltables/education/enrol_tertiary.xls.

UNESCO (2004). Statistics: Country profile (Republic of Korea). Retrieved July 30, 2004 from http://www.uis.unesco.org/countryprofiles/html/EN/countryProfile_en.aspx?code=4070.htm.

VanSledright, B. & Brophy, J. (1992). Storytelling, imagination, and fanciful elaboration in children's historical reconstructions. *American Educational Research Journal, 29*.

Voss, J. F., & Wiley, J. (2000). A case study of development historical understanding via instruction: The importance of integrating text components and constructing arguments. In P. N. Stearns, P. Seixas & S. Wineburg (Eds.), *Knowing, Teaching, and Learning history* (pp. 375-389). New York: New York University Press.

Voss, J. F., & Wiley, J. (1997). Developing understanding while writing essays in history. *International Journal of Education Research*, *27*(3), 255-265.

Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education*. Washington, DC: Council of Chief State School Officers.

Webb, N. L. (2001). *Alignment analysis of four states: Language arts standards and assessments grades 5, 8, and 11*. Madison, WI: Wisconsin Center for Educational Research.

White, H. (1984). The question of narrative in contemporary historical theory. *History & Theory, 8*(1), 1-33.

Wilson, S., & Wineburg, S. (1993). Wrinkles in time and place: Using performance assessments to understand the knowledge of history teachers. *American Educational Research Journal, 30*, 729-769.

Wineburg, S. (1991a). Historical problem-solving: A study of the cognitive process used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83* (1), 73-87.

Wineburg. S. (1991b). On the reading of historical texts: Notes on the breach between school and academy. *American Educational Research Journal*, *28*(3), 495-519.

Wineburg, S. (2000). Making historical sense. In P. N. Stearns, P. Seixas & S. Wineburg (Eds.), *Knowing, Teaching, and Learning history* (pp. 306-325). New York: New York University Press.

Wineburg, S. (2001). *Historical thinking and other unnatural acts*. Philadelphia: Temple University Press.

Wiggins, G. P. (1989). Teaching to the (authentic) test. *Educational Leadership, 46*, 41-47.

Wiggins, G. (1990). *The case for authentic assessment*. (EDD-TM-9010). Washington, DC: ERIC Clearinghouse on Tests, Measurement, and Evaluation, American Institutes for Research.

Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco, CA: Jossey-Bass Publishers.

Wood, S. (1995). Developing and understanding of time—Sequencing issues: *Teaching History, 79*, 11-14.

Yang, J. (2001). Chilcha kyoyuk kwajong kwa yōksa kyoyuk ūi jinro [The 7th curriculum and the direction of history education]. *Kyoyuk madang, 8*, 106-107.

Yeager, E. A. & Foster, S. J. (2001). The role of empathy in the development of historical understanding. In O. L. Davis Jr., E. A. Yeager, and S. J. Foster (Eds.) *Historical Empathy and Perspective Taking in the Social Studies* (pp. 13-20). New York: Rowman & Littlefield Publishers, Inc.

Yeager, E. A, & Wilson, E. K. (1997). Teaching historical thinking in the social studies methods. *Social Studies, 88*(3), 121-126.

Yu, B. (1995). *Chungdūng kyoyuk* [Secondary education]. In G. An, S. Lee, Y. Cho, B. Yu, G. Lee, G. Han, Y. Lee, Y. Kim, Y. Namgung, & S. Jong, *Hanguk kūn-hyōndae kyoyuksa* [The history of Korean modern and contemporary education]. Seoul, Korea: Hanguk jōnshin munhwasa yōngu.

Yun, J., Song, G., Cho, D., Kim, B., & Yu, M. (1991). *Hanguk ui kyoyuk chōngchaek* [Educational policy of Korea]. Seoul, Korea: Kyoyuk kwahaksa.