AN EXAMINATION OF HIGH SCHOOL TEACHERS' REPORTED KNOWLEDGE, USE, PREPARATION, AND IMPORTANCE OF HIGHER-LEVEL THINKING QUESTIONS IN SUMMATIVE ASSESSMENTS

by

Jason J. Stock

B.A., Saint Cloud State University, 1997
M.S., Saint Cloud State University, 2002

A Doctoral Dissertation

Submitted to the Graduate Faculty

of

Saint Cloud State University

in Partial Fulfillment of the Requirements

for the Degree

Doctor of Education

St. Cloud, Minnesota

February, 2013

This dissertation submitted by Jason J. Stock is in partial fulfillment of the requirements for the Degree of Doctor of Education at St. Cloud State University and is hereby approved by the final evaluation committee.

_____
Chairperson

_____

_____

_____

_____
Dean
School of Graduate Studies

AN EXAMINATION OF HIGH SCHOOL TEACHERS' REPORTED KNOWLEDGE, USE, PREPARATION, AND IMPORTANCE OF HIGHER-LEVEL THINKING QUESTIONS IN SUMMATIVE ASSESSMENTS

Jason Stock

In this empirical study, teacher reported knowledge, use, preparation, and importance of higher-level thinking summative assessment questions were analyzed. This study was conducted using a survey distributed to seven Minnesota high schools that represented urban, suburban, and rural settings. Survey responses were cross tabulated with demographic and background variables of the sample group including gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status. Statistically significant findings were qualitatively and quantitatively examined.

The following research questions were investigated.

1. What relationships exist between teachers' demographics and their accuracy in identification of higher-level thinking questions?
   a. Demographic information and respondents who correctly identified at least two of the three higher-level thinking questions.
   b. Demographic information and respondents who correctly identified all three of the higher-level thinking questions.

2. What relationships exist between teachers' demographics and:
   a. Their use and importance of higher-level thinking questions in summative assessments.
   b. The strategies used by their building principals to convey the importance of using higher-level thinking assessment questions.
   c. The preparation they received at their university/college in developing higher-level thinking questions in summative assessments and the quality/frequency of school or district staff development opportunities on developing higher-level thinking questions in their summative assessments.
   d. Their knowledge/expertise in developing higher-level thinking questions in summative assessments.
   e. The identification of a person responsible for preparing teachers for developing higher-level thinking questions in summative assessments.

3. What relationship exists between teacher ratings of importance of higher-level thinking questions in summative assessments and:
    a. Their school making Adequate Yearly Progress.
    b. Their district employing a designated person to monitor development of higher-level thinking questions in teacher constructed assessments.

Findings revealed in this study were that less than one in four teachers could correctly identify all three higher-level thinking questions from a sampling of three higher-level and three lower-level thinking questions. Furthermore, a significant gender difference within these results was also discovered: more male than female teachers correctly identified all three of the higher-level thinking questions. Another important finding was that the majority of teachers surveyed believed that higher-level thinking questions in summative assessments were very important, but less than one in five reported that they have very high knowledge and expertise in this area. Administrators would be interested to know that over three in four teachers rate their staff development programs as average to poor in helping them develop higher-level thinking questions on summative assessments. In another important finding, respondents ranked institutions of higher learning as the most responsible for preparing teachers to incorporate higher-level thinking questions into their assessments while the building principal was ranked as least responsible for this preparation.

The survey population was exclusively high school teachers, so there was a close correlation between what the researcher was seeking to find through this study and the ability of the respondents to provide information.

Approved by Research Committee

_____
Month Year

_____
John Eller Chairperson

ACKNOWLEDGEMENTS

The process of completing my dissertation has been a very rewarding experience, which would not have been possible without the help of several people who played influential roles in my success. I would like to take this opportunity to extend my gratitude to those individuals.

First, I wish to thank my professors and dissertation committee at St. Cloud State University. I sincerely appreciate the hours of time that they provided to guide me through this process. This input and these suggestions significantly enriched my education.

Additionally, I extend my appreciation to Dr. Kay Worner for her constant support and encouragement through my sixth-year certificate and doctorate programs. Her advice and feedback greatly enhanced my program of study. It truly has been an honor to work with Dr. Worner!

Finally, I want to thank my parents, Dr. and Mrs. John and Lucille Stock, who provided me with a strong personal and academic foundation. This solid base played a key role in my life. I want to especially thank my mother. She always believed in my potential and the person who I would become. Thank you!

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Chapter I

INTRODUCTION

Introduction

Attention on how to teach and assess higher-level thinking skills began in earnest when the book, *Taxonomy of Educational Objectives* (Bloom, 1956) was written. This book focused on the classification of levels of intellectual behavior in learning based on three domains: cognitive, psychomotor, and affective. As Bloom specified, the cognitive domain was subdivided into six levels: knowledge, comprehension, application, analysis, synthesis, and evaluation. These six levels, known as Bloom's Taxonomy, are still referenced in discussions about higher-level thinking skills.

Bloom's Taxonomy (Bloom, 1956) has provided a guide for teachers to use in determining if their teaching objectives and test questions are distributed between lower and higher-level thinking skills. Both are needed to assure students' success in the classroom and beyond. The three levels typically identified as lower-level thinking skills are knowledge, comprehension, and application. In the first level, knowledge, students are expected to recall, list, repeat, and memorize information. In the second level, comprehension, students are expected to classify, discuss, identify, describe, and explain information. In the third level, application, students interpret, write about, and

demonstrate what they have learned and whether or not they can solve problems using their learning.

The next three levels in Bloom's Taxonomy (Bloom, 1956) are identified as higher-level thinking skills. The first of these is analysis, which focuses student learning on the ability to compare and contrast, distinguish between facts and inferences, and break learning down into parts. The next level is synthesis in which students create something new by combining ideas to form a unique or original product. The highest level among the higher-level thinking skills is evaluation in which students make value decisions about issues, argue and defend opinions, and evaluate their positions.

The importance of Bloom's Taxonomy and higher-level thinking skills, in general, continue to be part of the educational dialogue in the 21st century, especially following the passage of the Federal initiative, No Child Left Behind (No Child Left Behind [NCLB], 2002). This initiative focused on student achievement, testing, closing the "gap," and accountability. As a result, educators have taken a more comprehensive examination of what is taught, how it is taught, and how learning is assessed in American schools. The conclusion, among others, is that teachers and administrators must implement strategies to increase student achievement, as well as administer tests that best measure achievement.

Closing the achievement gap is a goal of nearly all, if not all schools in the United States. Strategies that particularly target the lower performing and/or disadvantaged students are viewed as critical to addressing the achievement gap. Furthermore, a curriculum emphasizing higher-order thinking skills has been found to

substantially increase math and reading comprehension scores of economically disadvantaged students (Pogrow, 2005). The argument for teaching critical or higher-level thinking skills is further strengthened when there is a link between those skills and increased student achievement. Studies show that teaching higher-level thinking skills results in higher test scores on both national standardized tests and state high-stakes test (Moore & Stanley, 2010).

In a rapidly changing world, students and adults must be able to adjust and adapt to new ideas, different ways of learning, and thinking critically about solutions to complex problems. According to Sousa (2006), students are born with the brain structure that originates thinking. The foundation for higher-level thinking is already there. Students can be assisted in developing more complex reasoning if they are assisted in organizing the content of their thinking. Sousa supports Bloom's Taxonomy as a compatible structure or guide for organizing student thinking.

Unfortunately, according to Byrnes (2001), "We can safely say that emphasis on facts is the "norm" for the United States and the emphasis on thinking represents an occasional deviation from this norm" (p. 3). Byrnes further stated that the emphasis on higher-level thinking skills in the classroom helps students retain information longer, transfer acquired learning to new contexts, and are better able to solve new problems.

Even though teaching and assessing student learning at higher-levels of thinking is important, evidence suggests that not all students have learned to think critically. This is evident in the fact that an important goal of higher education institutions is to develop students' higher-level thinking skills. Ideally, this goal involves an emphasis on

teaching students to approach evidence from multiple perspectives in order to arrive at their own arguments (Bok, 2006). Students entering post-secondary institutions rarely have this skill developed. They generally enter as "dualistic" thinkers who may be prepared to evaluate arguments as true or false, or correct or incorrect, but have not developed the ability to think beyond their own experiences to evaluate the ways in which different perspectives might be equally legitimate (Roberts, 2008). Several studies have found that high school students often complete assigned tasks that require little cognitive engagement (Conley, 2007). As the demand for higher-level thinking continues to grow, test scores appear to be moving in the opposite direction. This fact has given rise to a national higher-level thinking movement (Geersten, 2003).

Especially in the last decade, an increasing number of employers have suggested that newly hired employees today do not seem to be able to think and solve problems (Moore & Stanley, 2010). In a report titled, *Are They Really Ready to Work? Employers' Perspectives on the Basic Knowledge and Applied Skills of New Entrants to the 21st Century U.S. Workforce* (Conference Board, 2006), more than half of the employers (58%) stated that critical thinking and problem solving skills are "very important" for incoming high school graduates" successful job performance. Of these same employers, nearly three-quarters (70%) rated recently hired high school graduates as deficient in critical thinking. There is a growing consensus that the nation's future workforce lack the skills and knowledge needed if the content and standards of our current K-12 system are not revamped (Dickman, Schwabe, Schmidt, & Henken, 2009).

> The number of low-skilled, blue-collar jobs is shrinking as technology replaces workers who were once trained to do such tasks. We're going primarily from a product industry to a service industry, where critical thinking skills are a prerequisite for success. Regardless of educational background or placement upon graduation from high school, no one would argue that critical thinking skills *hinder* the ability of one's success in the work force. In fact, when faced with the day-to-day decision-making required of most jobs, one would argue that proficiency in critical thinking skills would actually *enhance* the ability of one's success in the work force, again reinforcing the idea that raising the rigor in schools makes sense. (Moore & Stanley, 2010, p. 12)

Employers today are looking for workers who can think through problems and solve them. This means that teaching students to think at higher-levels could result in better career opportunities for them (Moore & Stanley, 2010). Unfortunately, most industry leaders agree that the nation's K-12 schools are not doing an adequate job preparing students for the workforce (Dickman et al., 2009).

Teaching higher-order thinking skills is important for students to be successful in developing needed skills to be successful in the K-12 classroom, in post-secondary education and in the work world. Few would argue that teachers should incorporate cognitive skill development (higher-level thinking skill development) and assessment in curriculum design and instructional strategies. In order to do so, it is essential that teachers have a basic understanding of cognitive thinking and learning levels. If they do not, it is likely that their instructional delivery and assessment practices will foster students who are not able to advance beyond being "dualistic" thinkers.

The following critical question was identified for the purposes of this study in order to address the problem of teacher use and understanding of higher-level thinking questions:

What do high school teachers from randomly selected Minnesota high schools report as their knowledge, use, preparation, and importance of higher-level thinking questions in the development of summative assessments?

Purpose of the Study

Research reveals that there is a link between higher-level thinking skills and increased student achievement. Raising the rigor, no matter what the circumstance, almost always increases achievement. Simply put, giving students practice thinking at a higher-level causes them to become better thinkers, not just for the duration of their school experience, but also in their lives beyond the classroom (Moore & Stanley, 2010).

"Brain researchers suggest that teachers (should) use a variety of higher-order questions in a supportive environment to strengthen the brain" (Cardellichio & Field, 1997, p. 33). "Meaningful learning requires teachers change their role from sage to guide, from giver to collaborator, from instructor to instigator" (Ó Murchú, 2003, p. 10). "Since students learn from thinking about what they are doing, the teacher's role becomes one of stimulating and supporting activities that engage learners in critical thinking" (Bhattacharya, 2002, p. 6).

If the development of higher-level thinking skills is important for students to be successful in the classroom and beyond, it is plausible to conclude that teachers should incorporate cognitive (thinking) skill development in curriculum and instruction. To do

this, teachers must include activities to enhance higher-level thinking and construct assessments that provide feedback on how well the students are progressing in developing critical (higher-level) thinking skills. Mentoring students in higher-level thinking means helping them to acquire the cognitive tools they need to move beyond a basic understanding of individual problems and, instead, see the collective dimensions to both their causes and solutions (Mills, 1959).

Given the importance of incorporating higher-level thinking into the learning process, there is a need to examine what teachers report regarding the importance and use of higher-level thinking questions in the development of summative assessments. Since very little research was found on this topic, this study provides valuable insights for school administrators on teacher responses to the study research questions.

The purpose of this study is to investigate a select sample of Minnesota high school teachers' reported knowledge, use, preparation, and importance of higher-level thinking summative assessment questions. Demographic and background variables gathered from the sample group members include gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status.

The research questions for the study are:

1. What relationships exist between teachers' demographics and their accuracy in identification of higher-level thinking questions?
    a. Demographic information and respondents who correctly identified at least two of the three higher-level thinking questions.

   b. Demographic information and respondents who correctly identified all three of the higher-level thinking questions.

2. What relationships exist between teachers' demographics and:
   a. Their use and importance of higher-level thinking questions in summative assessments.
   b. The strategies used by their building principals to convey the importance of using higher-level thinking assessment questions.
   c. The preparation they received at their university/college in developing higher-level thinking questions in summative assessments and the quality/frequency of school or district staff development opportunities on developing higher-level thinking questions in their summative assessments.
   d. Their knowledge/expertise in developing higher-level thinking questions in summative assessments.
   e. The identification of a person responsible for preparing teachers for developing higher-level thinking questions in summative assessments.
3. What relationship exists between teacher ratings of importance of higher-level thinking questions in summative assessments and:
   a. Their school making Adequate Yearly Progress.
   b. Their district employing a designated person to monitor development of higher-level thinking questions in teacher constructed assessments.

This study will be a quantitative, empirical examination in which teacher reported knowledge, use, preparation, and importance of higher-level thinking questions will be analyzed. This examination will be conducted using a survey distributed to seven Minnesota high schools that represent urban, suburban and rural settings. Survey data will be analyzed and conclusions will be formulated to address the study's questions.

"Many educators believe that specific knowledge will not be as important to tomorrow's workers and citizens as the ability to learn and make sense of new information" (Gough, 1991, p. 10). If students are to develop complex skills and critical thinking abilities, it is essential that teachers have the knowledge and background in higher-level thinking instruction and developing questions for higher –level thinking assessments.

Paul and Elder (1999) pointed out thinking is not driven by answers but by questions. Questions stimulate thought so that students can practice thinking through information or an idea, or generate new questions for further learning. As a result, it is important that teachers know how to develop assessments that actually test student abilities in the higher-level thinking levels.

Studies have found that the use of higher-order thinking skills (interchange with higher-level thinking skills) led to better student achievement (Redfield & Rousseau, 1981). Taylor, Pearson, Clark, and Walpole (2000) found that the number of higher-order thinking questions asked in the classroom was related to both teacher and school effectiveness. There is little doubt that a teacher's use of higher-level thinking questions

and, subsequently, assessing student learning from this instructional method, benefit the student in his or her classroom.

Significance of the Study

This study is important because it examined whether teachers can identify higher-level thinking skill questions, their use of higher-level thinking questions in assessments, their preparation in developing higher-level thinking questions, and who they believe should be responsible to assist in developing higher-level thinking questions in assessments. The results of the study are important to teachers and their professional development, to administrators in their focus on achievement and assessment, and most importantly, to the field of education because of the direct link of higher-level thinking questions in the classroom to closing the achievement gap. The results of the study will be instrumental in providing feedback to educators to improve curriculum, instruction and assessment-all of which increase student achievement.

Limitations of Study

According to Roberts (2009), limitations of a study are factors that affect a study but are not under the control of the researcher. Limitations can further be explained as the constraints that influence the generalizability or interpretation of the results of the study.

The following are some possible limitations of the study:

1. This study will only survey teachers in Minnesota. Teachers outside of Minnesota will not be surveyed.

2. This study will only survey a small sample of teachers as compared to the total population of teachers in Minnesota.
3. This study will only survey Minnesota high school teachers. Teachers of lower grades will not be surveyed.
4. This study will only survey public school teachers in urban, suburban, and rural school districts. Private school teachers will not be included in this survey.
5. The study must be delivered and accessed through the filtering systems of school email and internet services. Failure to receive the invitation to take part in the study could result in a limited participation rate.
6. Respondents need to accurately report data. Survey results are based on the honest response of participants.
7. The online survey delivery system must work reliably to provide results consistent with the responses of the participants.
8. Teachers may simply choose not to participate due to concern that their information may not be confidential and could be accessed by their principal.

Definition of Terms

1. Allied classes: These include courses in theater, music and band.
2. Assessment: The process of determining performance through evaluation.
3. Bloom's Taxonomy: The classification of levels of intellectual behavior in learning based on three domains: cognitive, psychomotor, and affective. The

cognitive domain is subdivided into six levels: knowledge, comprehension, application, analysis, synthesis, and evaluation (Bloom, 1956).

4. Cognitive: Faculty for the processing of information, applying knowledge, and changing preferences (Bloom, 1956).
5. Core classes: These include courses in math, English, social, and science.
6. Critical thinking skills: The process of actively conceptualizing, applying, analyzing, synthesizing, and/or evaluating information gathered from observation, experience, reflection, reasoning, or communication. This is used synonymously with higher-level thinking skills
7. Fine Arts classes: These include courses in art, computer graphics, etc.
8. FTE: Full-Time Equivalent. An FTE is a measure of how many full-time teachers a school employs.
9. Higher-level thinking skills: The cognitive levels of analysis, synthesis, and evaluation.
    a. Analysis Level Question: At this level the teacher expects the student to examine elements and the relationships between elements or the operating organizational principles undergirding an idea.
    b. Synthesis Level Question: At this level the teacher expects the student to put conceptual elements or parts together in some new plan of operation or development of abstract relationships.
    c. Evaluation Level Question: At this level the teacher expects the student to understand the complexity of ideas so that he/she can recognize how

concepts and facts are either logically consistent or illogically developed (Bloom, 1956).

10. Lower-level thinking skills: The cognitive levels of knowledge, comprehension, and application.
    a. Knowledge Level Question: At this level the teacher expects the student to recognize and recall information.
    b. Comprehension Level Question: At this level the teacher expects the student to be able to arrange or, in some way, organize information.
    c. Application Level Question: At this level the teacher expects the student to use abstractions to describe particular ideas or situations (Bloom, 1956).
11. No Child Left Behind (NCLB): The reauthorization of the Elementary and Secondary Education Act, which sets standards of student performance, teacher quality and leadership expectations (No Child Left Behind, 2002).
12. Summative Assessment: Cumulative evaluations used to measure student growth at the end of a course in order to determine whether long term learning goals have been met.

Summary

Higher-level thinking skills and student performance have been topics of discussion and debate within the educational community since the work of Benjamin Bloom spurred a plethora of research studies in the 1960s and 1970s. Over the past 50

years, this issue has been the focus of numerous studies and investigations. Some studies have examined student success in college. Others investigated state test scores and college entrance expectations. Still others researched K-12 curriculum and test rigor. The underlying question is, are higher-level thinking skills being developed and assessed in order to prepare high school students for success in the classroom and beyond?

Recently, the increased attention on accountability, testing, and academic or content rigor has once again amplified attention of teaching and assessing higher-level thinking skills. Political and global pressures have created an environment in which educators must examine how they can better develop and assess higher-level thinking skills of their students.

This study will investigate teachers' accuracy in identifying higher-level thinking skill questions, their reported use of, and training in, higher-level thinking skills in the development of summative assessments. It will also examine their rating of the importance of using higher-level thinking skills in summative assessments, and the perceived importance their school administration places on the use of higher-level thinking skills in assessments. Therefore, the Purpose of the Study is summarized in the following question:

> What do high school teachers from randomly selected Minnesota high schools report as their knowledge, use, preparation, and importance of higher-level thinking questions in the development of summative assessments?

The findings of this investigation should indicate the level of reported use and value teachers place on higher-level thinking skills in the teaching/learning process. With this information, administrators, curriculum developers, and teachers will be able to accurately identify and provide necessary professional development needed to assure incorporation of higher-level thinking skills in the classroom. This increased capacity of staff members is necessary so that the achievement of students is increased, the achievement gap is reduced, and teachers are more effective in the classroom.

Chapter II

REVIEW OF LITERATURE

Development of Higher-Level Thinking Discussions

Higher-level thinking skills are not a new topic of discussion. Documented discussions began emerging as early as 400 BC in Athens, Greece during the time of Plato and Socrates. In Plato's work, The Republic, he said " . . . compulsory learning never sticks in the mind" (360 BC) (Reeve & Grube, 1992).

"During the eighth to the nineteenth century, classical education was the prominent form of teaching which focused on answering the questions 'who, what, when, where and how' while neglecting the theoretical 'why' and 'which' questions" (K12 Academics, n.d., para. 6). Classical scholars were known for being able to remember and restate words and ideas yet they demonstrated little understanding of what they meant. Classical education remained the primary form of education until the late 18th and early 19th centuries (K12 Academics, n.d.).

In 1780, Abigail Adams, wife of John Adams stated, "Learning is not attained by chance; it must be sought for with ardor and attended to with diligence" (quotationspage.com, 2013). This same time period brought with it the advent of a Lancaster education. Joseph Lancaster was born of a shopkeeper in Southwark, London. In 1798, he opened his first free elementary school in his hometown. Lancaster used a

unique approach in his schools. Slightly higher-level students were to instruct and teach less advanced students.

Despite their initial success, Lancastrian schools were criticized for their poor standards and harsh discipline. Lancaster died in New York in 1838. At that time, there were between 1200 and 1500 schools using his principles.

The 1940s brought with it an increasing awareness of different levels of thinking. Benjamin S. Bloom led a group that extensively contemplated the nature of thinking. In 1956, after 8 years of work, they completed a handbook commonly referred to as *Bloom's Taxonomy*. The intent of the group was to develop a classification method that was believed to be important to the process of thinking. This "Taxonomy" was divided into three domains:

1. The cognitive—knowledge based domain—six levels.
2. The affective—attitudinal based domain—five levels.
3. The psychomotor—skills based domain—six levels.

The cognitive portion of Bloom's Taxonomy has received the most attention over the years as it "has been transformed into a basic reference for all educators worldwide. Unexpectedly, it has been used by curriculum planners, administrators, researchers, and classroom teachers at all levels of education" (Anderson & Sosniak, 1994, p. 1). Bloom himself considered the Handbook, "one of the most widely cited yet least read books in American education" (Anderson & Sosniak, 1994, p. 1). While other cognitive theories have been developed, Bloom's Taxonomy has become the standard for nearly 50 years.

Bloom's Taxonomy consists of six cognitive levels or steps. The levels are arranged as a hierarchy with the lowest three levels being knowledge, comprehension, and application. The three highest levels also referred to as higher-level thinking skills are analysis, synthesis, and evaluation. Due to the hierarchical structure, learners who are functioning at the application level have also mastered the understanding and remembering levels.

Bloom's Taxonomy has remained popular within the educational community. Over the years, proponents have attempted to condense, expand, and reinterpret it in a variety of ways. "Research findings have led to the discovery of a veritable smorgasbord of interpretations and applications falling on a continuum ranging from tight overviews to expanded explanations" (Forehand, 2005, para. 9).

During the 1990s, Lorin Anderson, a former student of Bloom's, worked with a group of his peers to update the taxonomy. Even "the original group always considered the [Taxonomy] framework a work in progress, neither finished nor final" (Anderson & Krathwohl, 2001b, p. xxvii). Bloom himself wrote in a memorandum from 1971 in which he stated, "Ideally each major field should have its own taxonomy of objectives in its own language—more detailed, closer to the special language and thinking of its experts, reflecting its own appropriate sub-divisions and levels of education, with possible new categories, combinations of categories, and omitting categories as appropriate" (Bloom circa 1971, cited in Anderson & Krathwohl, 2001b, pp. xxvii-xxviii). Cognitive psychologists, curriculum theorists, instructional researchers, and assessment specialists revisited the levels within the taxonomy. After 6 years of

examination, the New Bloom's Taxonomy was released and published in 2001. The revisions included terminology, structure, and emphasis.

The greatest change to the taxonomy was in its terminology. The six major categories were changed from noun to verb forms. Additionally, the lowest level of knowledge was renamed to remembering while comprehension and synthesis were re-titled to understanding and creating. Finally, the top two levels were essentially exchanged making creating the highest level of cognitive function. Figure 1 provides a graphical representation of the changes that occurred.

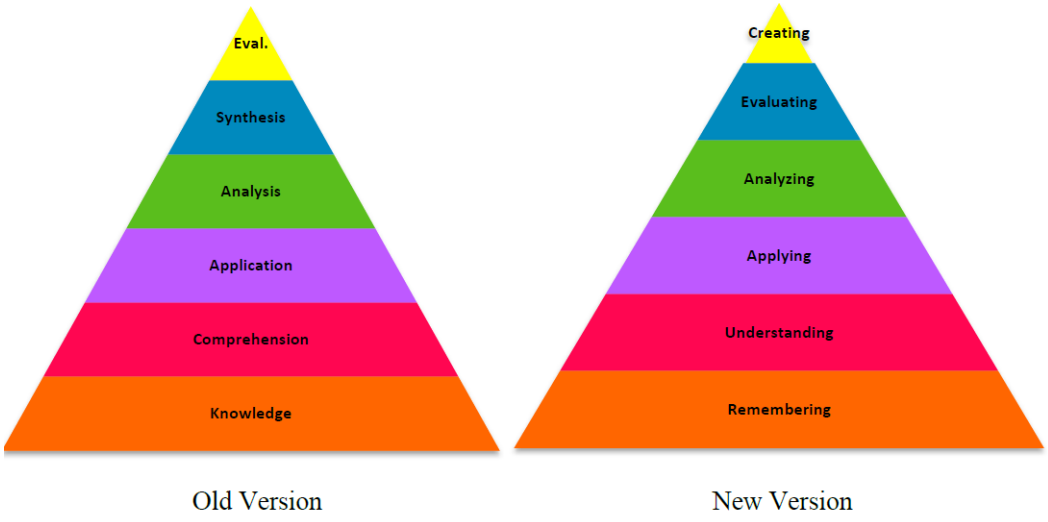


Figure 1

Graphical Representation of Changes to Bloom's Taxonomy
(Anderson & Krathwohl, 2001b, p. 67)

The new terms are defined as:

- **Remembering**: Retrieving, recognizing, and recalling relevant knowledge from long-term memory.

- **Understanding**: Constructing meaning from oral, written, and graphic messages through interpreting, exemplifying, classifying, summarizing, inferring, comparing, and explaining.
- **Applying**: Carrying out or using a procedure through executing, or implementing.
- **Analyzing**: Breaking material into constituent parts, determining how the parts relate to one another and to an overall structure or purpose through differentiating, organizing, and attributing.
- **Evaluating**: Making judgments based on criteria and standards through checking and critiquing.
- **Creating**: Putting elements together to form a coherent or functional whole; reorganizing elements into a new pattern or structure through generating, planning, or producing. (Anderson & Krathwohl, 2001, pp. 67-68)

The revised taxonomy improves on Bloom's ideas and can be used in the classroom as a valued resource for today's teachers. The structure "provides a clear, concise visual representation" (Krathwohl, 2002, p. 212) of higher-level thinking skills. "Today's teachers must make tough decisions about how to spend their classroom time. Clear alignment of educational objectives with local, state, and national standards is a necessity" Foreland (2005, para. 20). The revised taxonomy provides a clear framework for teachers to use to assess their curriculum goals.

Background of Assessment Strategies

Student achievement and assessment has been discussed for decades. Quantifying student learning can be challenging for teachers at every grade level. A variety of methods have been developed in the quest for accurately and meaningfully assessing student knowledge and progress in learning what has been taught. "If what students learned as a result of the instructional practices of teachers were predictable, then all forms of assessment would be unnecessary; student achievement could be

determined simply by inventorying their educational experiences" (Wiliam, 2010, p. 254). The challenge in assessing student learning is to connect what is taught with how it is taught and ultimately, to understand how well students have learned.

According to Muirhead (2002), "Assessment is a flash point and catalyst for controversy" (para. 1). Assessment is an important component in the learning process that encourages learners to consider assessment methods that match the needs of today learners (Muirhead, 2002). At a more fundamental level, "A primary aim of assessment is (to) provide the necessary information to improve future educational experiences" (Muirhead, 2002, para. 5). It requires the time and effort to ask questions that will help assess the effectiveness of the teaching strategies that are being used (Huba & Freed, 2000). Vella, Berardinelli, and Burrow (1997) related that an important purpose of evaluation is "to determine if all of the learners developed important knowledge, skills, and attitudes as a result of the program" (p. 16). These insights into assessment highlight that instruments designed to measure student performance not only assess student, but also evaluate teacher performance.

Assessment Forms/Styles

Assessment can be conducted using several methods including student-centered, alternative, rubric, and journaling. The focus of student-centered assessment puts the learner at the center of the process. This evaluation method reflects the needs, gifts and talents of the learner. Teachers must recognize that students are individual learners who are independent and desire relevance in the evaluation of their work (Caffarella, 1993).

"The student-centered model of learning encourages teachers to view their students as academic partners who work together to produce relevant and meaningful learning experiences. It requires professors who are willing to change their standard teaching methods" (Muirhead, 2002, para. 7). Boud (1995) commented on the role of the teacher with they will need to become researchers of student perceptions, designers of multifaceted assessments strategies, managers of assessment processes and consultants assisting students in the interpretation of rich information about their learning (cited in Muirhead, 2002, para. 8).

The eight features below are considered integral components of learner-centered teaching (Huba & Freed, 2000):

- Learners are actively involved and receive feedback.
- Learners apply knowledge to enduring and emerging issues and problems.
- Learners integrate discipline-based knowledge and general skills.
- Learners understand the characteristics of excellent work.
- Learners become increasingly sophisticated learners and knowers.
- Professors coach and facilitate, intertwining teaching and assessing.
- Professors reveal they are learners, too.
- Learning is interpersonal, and all learners---students and professors---are respected and valued. (p. 33)

The student-centered instructor must constantly provide the connection between academic knowledge and student skills creating meaningful assessments that promote achievement.

Alternative assessments are another method of measuring student growth. Alternative assessment demonstrates by the very nature of its name that it is in direct contrast to more traditional methods of evaluation. Alternative assessment originates from educators who have become frustrated with more traditional evaluation methods

(Sanders, 2001). "There are two major differences between the traditional educator and those who use alternative assessment. The first is that the traditional educator is more dependent upon fewer assignments to evaluate student performance" (Muirhead, 2002, para. 10). In contrast, instructors who use alternative assessments might use portfolios, presentations, book reviews and/or interviews (Travis, 1996).

> Alternative assessment methods are promoted as a way to encourage authentic learning. Students are given a diversity of learning opportunities to display their critical thinking skills and greater depth of knowledge. They connect their daily lives to the learning, develop deeper dialogue over the course material and foster both individual and group orientated learning activities. (Muirhead, 2002, para. 11)

Teachers who choose to use this method are constantly challenged to produce and communicate consistent evaluation techniques that truly assess student learning. If used correctly, this method should encourage students to develop their thinking on a variety of subjects without concern about whether their responses are right or wrong; assessment questions that only require recall from the student cannot accomplish the same result.

Grading rubrics are another form of assessment used by teachers. Rubrics are public statements that strive to produce trust between teacher and student. They reject the perceived secret nature of evaluation criteria and allow learners to actually see the instructor's grading process (Muirhead, 2002). Rubrics are valuable due to their ability to reveal vital information to students that encourages them to improve their knowledge and skill levels (Huba & Freed, 2000). Rubrics are effective ways to limit subjective

evaluation and encourage objective evaluation. Huba and Freed (2000, cited in Muirhead, 2002) outlined five elements of effective rubrics:

1. Levels of Mastery. Achievement is described according to terms such as excellent, good, needs improvement and unacceptable.
2. Dimensions of Quality. Assessment can address a variety of intellectual or knowledge competencies that target a specific academic discipline or involve multiple disciplines.
3. Organizational Groupings. Students are assessed for multidimensional skills such as teamwork that involves problem solving techniques and various aspects of group dynamics.
4. Commentaries. This element of the rubric provides a detailed description of the defining features that should be found in the work. The instructor creates the categories for what is considered as being excellent, sophisticated or exemplary.
5. Descriptions of Consequences. This is a unique rubric feature that offers students insight into various lessons of their work in a real life setting (i.e. professionalism). (para. 15)

The five rubric components offer educators great guidelines to use as they develop their evaluation procedures that fit their students' population and academic discipline (Muirhead, 2002).

Journal writing is another form of assessment used by today's educators. Journaling can yield powerful insights into a student's thoughts and understanding of materials. Muirhead (2001, cited in Muirhead, 2002) shared seven major advantages to journal writing:

1. Provides an aid to our memory—researchers and writers have learned the value of recording their ideas for future use.
2. Provides a basis for creating new perspectives—it creates a framework to explore relationships and arguments between ideas.
3. Enhances critical thinking skills—learning to analyze the underlying assumptions of our actions and those of others is a very liberating process.
4. Provides psychological/emotional advantages—it enables individuals to work through difficult work or personal situations that can promote healing and growth.

5. Offers opportunities to increase empathy for others—individuals can address social issues and enhance their understanding of our society and world.
6. Provides a practical way to understand books/articles—writing creates a format to regularly examine reading materials and improve our ability to comprehend and recall knowledge.
7. Provides support for self-directed learning activities—journal writing requires personal discipline and establishing individual learning goals to complete journaling assignments. (para. 17)

Journaling can be an effective way to measure student performance. Evaluators must remember that it is imperative to provide constructive feedback in a timely manner in order to allow adequate student time to make the necessary changes (Muirhead, 2002). If used correctly, journaling can be an enriching and valuable technique used to enhance a variety of classroom experiences.

Formative and Summative Assessment

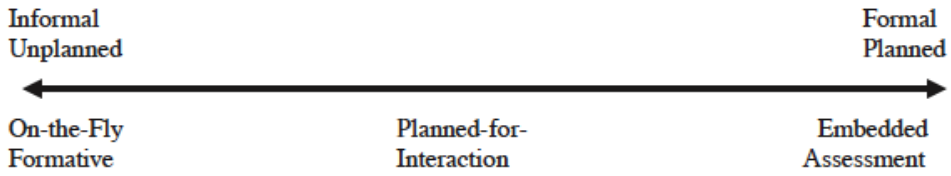
Formative assessment is the tool teachers use to measure if curriculum is being comprehended by students. Formative assessment's purpose is to inform teachers and students as to the gap between what students know, and what they are expected to know (Shavelson, 2006). When incorporated into the classroom, it provides the information teachers need to adjust their curriculum as it is happening (Garrison & Ehringhaus, 2007).

> Since the goal of formative assessment is to gain an understanding of what students know (and don't know) in order to make responsive changes in teaching and learning, techniques such as teacher observation and classroom discussion have an important place alongside analysis of tests and homework. (Boston, 2002, para. 7)

Black and William (1998) made the following recommendations:

- Frequent short tests are better than infrequent long ones.
- New learning should be tested within about a week of first exposure.
- Be mindful of the quality of the test items and work with other teachers and outside sources to collect good ones. (p. 139)

According to Shavelson et al. (2008, p. 295), formative assessment can range anywhere on the continuum as depicted below:



Impromptu formative assessment occurs when teachable moments occur in a classroom. The teacher uses these moments to encourage growth in understanding. Planned-for-interaction formative assessment occurs when a teacher plans in certain central questions within the lesson. Embedded-in-the-curriculum formative assessments are evaluation tools that are built ahead of time at key junctures within the class. Embedded assessments give the teacher a snapshot of any gaps that students may have with the adopted curriculum (Shavelson et al., 2008).

Summative assessments are given periodically to see what level of understanding a student has at a particular point in time (Garrison & Ehringhaus, 2007). "Formative assessment is assessment *for* learning, and summative assessment is assessment *of* learning" "Teachers and trainers use summative assessment to discover what a learner has achieved during the program of study" (Garrison & Ehringhaus, 2007, p. 1).

Many individuals associate summative assessments only with state testing and similar type high-stakes tests. However, summative tests encompass a large variety of assessment tools, which include the following:

- State assessments
- District benchmark or interim assessments
- End-of-unit or chapter tests
- End-of-term or semester exams
- Scores that are used for accountability of schools (AYP) and students (report card grades). (Garrison & Ehringhaus, 2007, p. 1)

The true value of summative assessment is that teachers can assess, at a particular point in time, student learning. "Summative assessments happen too far down the learning path to provide information at the classroom level and to make instructional adjustments and interventions during the learning process. It takes formative assessment to accomplish this" (Garrison & Ehringhaus, 2007, p. 1).

Standardized and High Stakes Tests

Validity of a test is defined as the degree to which a test is actually measuring what it is designed to measure (Allpsychonline, n.d.). As Nuttall (1987) suggested, "The fidelity of the inference drawn from the responses to the assessment is what is usually called the *validity* of the assessment" (p. 110). Because increasing attention today is being placed on high stakes tests, the validity of assessment tools is constantly being examined and scrutinized. To further complicate this issue, test validity can be different for various groups of people. Hence, as Nuttall pointed out, "In practice, an assessment does not have a single validity; it can have many according to its different uses and the different kinds of inference made, in other words, according to the universe of

generalizations" (p. 110). A test can be an effective measure for one purpose, and not for another (Wiliam, 2010). According to Cureton (1951,) "The essential question of test validity is how well a test does the job it is employed to do. The same test may be used for several different purposes, and its validity may be high for one, moderate for another, and low for a third" (p. 621).

College admission departments are a primary consumer of standardized tests. They use these test scores as predictive validities of higher educational success. Admission departments are also increasingly using standardized testing for entering freshmen. In 1901, fewer than 1000 examinees participated in the first "College Boards." Contrast that to today where nearly 3 million high school seniors take SAT or ACT tests annually (Atkinson & Geiser, 2009).

Since the first "College Boards" in 1901, standardized testing for college admission has rapidly grown. In 1926, the new "Scholastic Aptitude Test" claimed to measure a student's general analytic ability. In 1959, the creation of the ACT as a competitor to the SAT was intended as a measure of achievement rather than ability. Current efforts focus attention on adapting K-12 standards-based tests as a measure of college success in the future (Atkinson & Geiser, 2009).

> . . . it is evident that we have come full circle to a renewed appreciation for the value of achievement tests." "A century of admission testing has taught us that this initial premise may have been sounder than anyone realized at the time. (Atkinson & Geiser, 2009, p. 665)

Although standardized tests are valuable tools, it is important to recognize that high school grades are still the best indicator of student success in college (Atkinson &

Geiser, 2009). "Irrespective of the quality or type of school attended, cumulative grade-point average in academic subjects in high school has proven consistently the best overall predictor of student performance in college" (Atkinson & Geiser, 2009, p. 666).

Related Studies

Higher-level thinking skills have been a topic of conversation for nearly 2500 years. Particular interest in this topic was intensified in the 1950s with the publishing of Bloom's Taxonomy. During the 1960s and 1970s, interest continued resulting in several studies that measured the test writing skills of teachers and organizations that produced standardized tests. Although interest seemed to wane after 1980, resurgence has recently developed. In the following section of this paper, comparable and analogist studies will be examined and reviewed.

In a study conducted by Davis and Pfeiffer (1965), which took place from 1963-1964, semester examinations for all ninth grade courses were secured from a junior high school in Northeastern Ohio. Classes included were Civics, College Preparatory World History, College Preparatory Algebra, General Mathematics, College Preparatory Biology, General Science, Beginning and Advanced French, English, Home Economics, and Business Training. The test questions were categorized according to Blooms Taxonomy.

Tests were collected and individually analyzed for the level of cognitive ability required to complete the problems. A committee of teachers was used to collect and interpret the tests. To establish reliability on the classification process, a sampling of

items from each test was reanalyzed after a 3-week period. No change in classification resulted. In addition, a scorer experienced in the use of the Taxonomy analyzed another sample of test items. The inter-scorer agreement was 0.87. The test questions were not only classified into the six areas according to Blooms, but also further classified into sub-categories. Students were also divided into the three groups of prevocational, business and college preparatory. Within these sub groups, the six knowledge levels were broken down into percentages (Davis & Pfeiffer, 1965).

Findings uncovered that over half of all questions in each of the three areas were at a knowledge level. The next largest area was application, which peaked at 34%. The remaining areas were very consistent between the three divisions. The higher-level thinking skills of synthesis and evaluation were virtually identical in the prevocational and college preparatory areas. The study concluded that the lack of higher-level thinking processes, while not unusual was surely depressing. In a way, the students at this Ohio school were intellectually deprived (Davis & Pfeiffer, 1965).

In a study conducted by Pancella (1971), several commercially developed biology tests were examined. The tests were evaluated according to Bloom's taxonomy. Two research questions were posed:

1. What percent of test questions found in standardized and commercial tenth grade biology examinations is represented by each of the six levels of the Taxonomy?
2. Do tests of the Biological Sciences Curriculum Study (BSCS) contain more items which measure higher cognitive levels than do other standardized or commercially prepared tests? (Pancella, 1971, para. 3)

A total of 2689 test items were classified with 71.88% of questions identified at the knowledge level; 15.17% at the comprehension level; 11.49% at the application level; 1.37% at the analysis level; 0.04% at the synthesis level; and 0.04% at the evaluation level. Only 1.45% of the questions were in the top three levels of Bloom's Taxonomy. Very few tests other than BSCS tests reflected more than 10% questions above the second level of Bloom's. Five tests were totally at the knowledge level. Eleven of the tests had knowledge level questions in excess of 90% of their total. Only BSCS tests reflected levels above application (Pancella, 1971).

In a study conducted by Funk (1972), pre-service teachers' test building abilities were examined in relation to Bloom's Taxonomy. This study was designed to investigate the effectiveness of a method intended to develop improved test construction. The purpose of the study was to:

1. Examine the variables affecting test-construction practices of pre-service science teachers.
2. Examine the effect of instruction on the attitudes of pre-service science teachers.
3. Examine the relationship between test-construction behavior and attitudes
4. Determine whether the testing practices of pre-service science teachers were different from a population of science teachers. (Funk, 1972, para. 2)

Pre-service teachers were asked to create a series of tests at different points of their teacher development programs. These tests were classified according to Bloom's Taxonomy. Several veteran teachers' tests were also examined in the same manner. An examination of the data showed that the tests prepared by the study population, both in the methods course and during the student teaching experience, contained a lower

proportion of knowledge level items than the tests prepared by a population of experienced science teachers. It was also discovered that the tests prepared by the study population during the student teaching assignment contained significantly higher proportions of knowledge level items than the tests they prepared during the methods course. An inter-scorer reliability coefficient of 0.89 and an intra-scorer reliability coefficient of 0.94 were obtained (Funk 1972).

In a study conducted by Billeh (1974), the tests of several teachers in Beirut were analyzed according to Bloom's Taxonomy. The classes that were examined were Chemistry, Physics and Biology. The answers to the following questions were sought:

1. What relationship, if any, exists between the level of the test item as classified by Bloom's Taxonomy and each of the following teachers' characteristics: (a) years of science teaching experience, (b) professional in-service training, (c) academic specialization, and (d) status as part-timer versus full-timer.
2. Is there a relationship between the science subject taught and the level of the test item?
3. Do teachers of different grade levels emphasize different cognitive levels of test items?
4. What percentage (proportion) of the test items asked by secondary school science teachers falls into each of the categories of Bloom's classification system. (Billeh, 1974, p. 313)

Many interesting conclusions were discovered. One example was that items requiring lower-levels of cognition seemed to be asked by teachers with more teaching experience. Billeh (1974) additionally found that there was not a significant correlation between topics covered and questioning cognitive levels. By far, the heaviest emphasis in science examinations was on the lowest level of the classification system-the knowledge level. Nearly 72% of the examination time was devoted to recall of facts of one sort or

another. Furthermore, 60% of that time requires only the lowest cognitive levels. Only 7% of the examination's time was devoted to questions requiring the application of science principles, theories, or other abstractions to new situations. Test items requiring comprehension constituted 21% of the examinations. Test items requiring the highest cognitive levels, namely analysis, synthesis, and evaluation were absent.

In a study conducted by Fast (1974), the standardized tests in chemistry were analyzed according to Bloom's Taxonomy. The following questions were asked:

1. What are the classifications of the test items in the twelve ACS-NSTA High School Chemistry Tests according to Bloom's six cognitive levels?
2. What trend, if any, do the more recent tests show in the six cognitive levels classification?
3. How does the difficulty of a test time compare with the cognitive level classification of a test item?
4. How does the discrimination factor of a test item compare with the cognitive level classification?
5. Do the advanced series contain more of the upper cognitive level questions? (Fast, 1974, p. 17)

The 12 tests in this study contained a total of 955 items, which were classified according to the taxonomy. After the 955 items were classified in the major levels of the taxonomy, a stratified random sample of 40 items was selected for validation by a panel of judges who were experienced in using the Taxonomy. The 93% agreement in the sample set compares favorably with the percentage of agreement reported by the other investigators (Fast, 1974).

It was discovered that 40% of the items in the 12 tests examined were classified as knowledge level items. The comprehension and application level items each made up approximately 25%. The remaining 10% were analysis items. The levels of synthesis

and evaluation were absent from their findings. The Fast (1974) study supports earlier evidence that the cognitive levels of synthesis and evaluation are difficult to produce in a forced choice format such as multiple choice. Perhaps these higher-level-thinking questions would be easier to represent using questions with free response or an essay type format.

In a study conducted by Funk (1977), several middle school standardized tests were examined. Although this study was conducted with middle school students in mind, it provides a snapshot of what was being required at this level. The researchers in this case examined a number of factors. Among them were such things as reliability, validity, and construction.

A test's reliability refers to its ability to consistently produce the same results. For example, if a student is taking an IQ test, the IQ that the test predicts should be consistent every time that pupil takes the test. A test can be reliable, but not valid if the instrument produces consistent results, but the results are not correct. In terms of science, this would be a precise answer that was not accurate (Funk, 1977).

A test is valid only if the test reflects the actual correct results. In the example above, a test that repeatedly predicts the same IQ score is only reliable. In order to be valid, the IQ score has to be correct. Finally, the construction of the test can be determined in many ways. In this examination, the levels of cognitive questions were analyzed according to Bloom's Taxonomy.

The results of the construction portion of this analysis uncovered that of the 12 standardized tests that were selected, only two of the instruments had questions above

the application level. None of the tests had questions in the synthesis or evaluation levels. Of the two tests that did have analysis, both had less than 2%. Although all of the examinations had questions in the application level, the highest percentage was just 12 percent with the lowest being one percent.

In a study conducted by Kracht (1978), various geography-standardized tests were examined to determine if certain themes were covered and if the questions required higher-level thinking skills. The following questions were asked:

1. Are test items contained in geographic achievement tests or the social studies components of national achievement tests directed primarily toward the testing of knowledge at the recall/memory level?
2. Are test items contained in geographic achievement tests or the social studies components of national achievement tests comprehensive in their treatment of the discipline? (Kracht, 1978, para. 2)

In reference to the first research question, the majority of the tests provided a balance of questions between the lower and higher cognitive questions. Because it was found that the majority of tests emphasized interpretation of maps, graphs, and charts, and because by definition those items are higher-order questions, it was decided to further analyze the items excluding these types of questions (Kracht, 1978). In regard to the second research question, it was found that only three tests included questions from 60% or more of the content categories included in the analysis (Kracht, 1978).

In a study conducted by Black (1980), which was conducted in Nigeria, the tests of chemistry, physics, and biology teachers were examined and screened using Bloom's Taxonomy. The influence of the following variables was also investigated:

1. Teachers' educational backgrounds. Is there a difference in cognitive emphasis on exams of these three categories of teachers: Untrained

university graduates, Nigerian Certificate in Education holders from the advanced teacher training colleges, and untrained graduates?
2. Subjects taught. Do different science subjects tend to place varying emphasis on problem solving, concept development, and memorization?
3. Grade level taught. Do teachers' perceptions of capabilities in the five forms of secondary school influence the cognitive emphasis on their tests?
4. School Certificate examinations. Taken at the end of secondary school, performance on these determines whether a student receives a Certificate or not. Two questions were of interest:
    a. How much do teachers use past West African School Certificate Examinations as a guide in writing their own class tests?
    b. Is there a difference in cognitive emphasis between School Certificate and teachers' examinations? (Black, 1980, pp. 301-302)

Seventy-five randomly selected secondary schools, of a possible 207 in the East Central State, were contacted by mail in 1971. The principal was asked to complete a brief questionnaire and each science teacher was asked to send an examination recently given, plus complete a brief questionnaire. Forty-eight schools (64%) responded. Questions on all examinations were classified according to Bloom's Taxonomy. An inter-scorer reliability rating of 0.89 determined for a panel of three on a set of 295 questions (Black, 1980).

Findings from the Black (1980) study included that no biology questions were found to be above the comprehension level and for chemistry, physics and general science, none above application. Teachers' qualifications did not appear to have a significant effect on the level of questions being asked. The subject taught had the greatest influence. The fact that physics teachers asked significantly fewer knowledge-level questions than chemistry teachers, and that chemistry teachers asked fewer knowledge level questions than biology teachers seemed to reflect a strong discipline

influence. The form (grade level) taught proved to be a significant influence only for the physical science and physics groups. Finally, the West African School Certificate Examinations appeared to be a weaker influence than originally expected.

In this study by Sultana (2001), the lesson plans of 67 teacher interns in Kentucky were examined to determine the extent to which their lesson objectives were designed to develop higher-order thinking skills in their students. The lesson plans represented 43 elementary, 15 middle, and 9 high school teacher interns. The study took place during a 3-year period from 1995-1998.

Two researchers individually and independently categorized the data of the new teachers into cognitive levels using Bloom's Taxonomy. A Pearson's "r" correlation was performed to determine inter-rater reliability between the two raters. The test resulted in an inter-rater reliability of $r = 0.98$. The raters discussed the items in which they disagreed and came to agreement through consensus (Sultana, 2001).

The Sultana (2001) study found that 41.3% of material covered was at the knowledge level. An additional 19% was at the comprehension level, 16.7% was at the level of application, and 10.3% was analysis. Synthesis made up 9.5% of the material covered, and evaluation only 3.2%. Of the material that the beginning teachers taught, 23% was directed toward the highest three levels of Bloom's Taxonomy, and 77% of the objectives were aimed at the three lowest levels of Bloom's Taxonomy.

These findings suggest that if teachers teach as they have been taught in teacher development programs, the programs designed to teach aspiring teachers to be effective in the classroom should more adequately prepare new teachers to challenge pupils of the

future through the use of higher-level thinking materials in the materials developed for the student learning.

In a study conducted by Oliver and Dobele (2007), the assignments and tests of six entry-level IT (Information Technology) classes at a university in Australia were examined. The six courses examined were from within the School of Computing Sciences at Central Queensland University, which constitute three-quarters of the first year program of study. The following results were found:

Bloom Rating: Assignments

Course	Knowl-edge	Compre-hension	Applica-tion	Analy-sis	Synth-esis	Evalua-tion	Weight-ing	Bloom Rating
SF	30.5	5.3	0.2	0.0	0.0	0.0	36.0	1.2
CFC	7.3	32.6	0.0	0.1	0.0	0.0	40.0	1.8
SAD	10.0	8.5	15.3	2.0	4.2	0.0	40.0	2.6
PROG1	8.8	9.9	11.3	10.0	0.0	0.0	40.0	2.6
PROG2	3.8	4.7	15.0	11.5	0.0	0.0	35.0	3.0
WI	1.5	3.0	0.0	19.5	6.0	10.0	40.0	4.4

Bloom Rating: Examinations

Course	Knowl-edge	Compre-hension	Applica-tion	Analy-sis	Synth-esis	Evalua-tion	Weight-ing	Bloom Rating
SF	42.0	22.0	0.0	0.0	0.0	0.0	64.0	1.3
SAD	42.0	15.0	3.0	0.0	0.0	0.0	60.0	1.4
CFC	12.0	48.0	0.0	0.0	0.0	0.0	60.0	1.8
PROG1	13.2	22.8	12.0	12.0	0.0	0.0	60.0	2.4
PROG 2	4.0	27.0	15.5	18.5	0.0	0.0	65.0	2.8
WI	0.0	6.0	0.0	18.0	0.0	36.0	60.0	5.0

Bloom Rating: Overall

Course	Knowl-edge	Compre-hension	Applica-tion	Analy-sis	Synth-esis	Evalua-tion	Weight-ing	Bloom Rating
SF	72.5	27.3	0.2	0.0	0.0	0.0	100.0	1.3
SAD	52.0	23.5	18.3	2.0	4.2	0.0	100.0	1.8
CFC	19.4	80.4	0.0	0.2	0.0	0.0	100.0	1.8
PROG1	22.0	32.7	23.3	22.0	0.0	0.0	100.0	2.5
PROG2	7.8	31.7	30.5	30.0	0.0	0.0	100.0	2.8
WI	1.5	9.0	0.0	37.5	6.0	46.0	100.0	4.8

(Oliver & Dobele, 2007, p. 348)

In the Oliver and Dobele (2007) study, each class was given a Bloom Rating for both assignments and examinations. An overall rating was also given. The Bloom Rating was a number between one and six depending on the difficulty of the curriculum.

The results showed that there was significant difference between the six classes examined. Although many of the classes did focus on lower-level thinking skills, this could have been predicted due to the entry-level nature of these classes. Significant amounts of synthesis and evaluation were found in the WI (Workplace Issues) classes, which earned a 4.8 Bloom Rating. This study seems to be the exception to the general absence of higher-level thinking skills in examinations.

In a study conducted by Khorsand (2009), tests given by EFL (English as a Foreign Language) teachers in Iran were evaluated. Twenty teachers participated in the study and from their tests, generated 215 questions that were classified according to Bloom's Taxonomy. The teachers selected were experienced, and it was believed that they would generate effective questions. This study sought to answer the following questions:

1. What cognitive levels of text processing are indicated by questions generated by EFL teachers in advanced reading texts?
2. Which levels of questions do teachers emphasize more: high or low cognitive level of questions? (Khorsand, 2009, para. 2)

The Khorsand (2009) study was found that 57.21% of the questions were at the knowledge level. Fewer questions, 36.74%, were written at the comprehension level while only 1.86% were application level questions. Higher-level thinking questions were even lower with analysis at 0.47%, synthesis at 2.33%, and evaluation at 1.39%. As this data shows, 93.95% of the questions were written at the lowest two levels of the taxonomy. Regarding high and low cognitive levels, teachers emphasized low-level

questions, since 95.81% of the questions targeted low-level cognitive abilities while only 4.19% were at a high level.

In general, the Khorsand (2009) findings suggested that teachers demanded little deep text processing from students based on the testing used in the classroom. The data from this study indicated that Iranian EFL (English as a Foreign Language) teachers in advanced reading comprehension classes were designing their testing primarily at the lowest cognitive levels of Bloom's taxonomy. Possible explanations for teachers' preference for low level questions are listed below:

1. Low-level questions take little time for teachers to generate.
2. Teachers are mostly unaware of different cognitive levels of questions and learning.
3. Maybe teachers, themselves, do not pay attention to higher-level processes and do not expect their students to pay attention to these things.
4. Maybe our teachers consider knowledge and comprehension more important and focus on those things.

In a study conducted by Kocakaya and Gonen (2010), examinations from 19 physics teachers in Diyarbakir, Turkey were examined. Four types of high schools were used in this study: (a) three "Ordinary," (b) two "Vocational and Commercial," (c) one "Anatolian," and (d) one "Science." Test were evaluated to determine the level at which questions were being asked according to Bloom's Taxonomy. These results were compared to college entrance exams. The purpose of this study was to:

1. Compare the cognitive levels of questions asked at high school physics exams and university entrance exams according to Bloom's Taxonomy.
2. Determine the differences at the level of asked questions at high school and university entrance exams. (Kocakaya & Gonen, 2010, p. 5)

Figure 2 breaks the findings of the 876 questions according to the school type and cognitive level.

School Types	**OHS**		**VCHS**		**AHS**		**SHS**		**Total**	
Question Levels	**f**	**%**	**f**	**%**	**f**	**%**	**f**	**%**	**f**	**%**
Knowledge	33	9.6	20	8.9	-	0.0	2	1.2	55	6.3
Comprehension	77	22.4	29	12.8	2	1.5	14	8.1	122	13.9
Application	171	49.7	147	65.0	56	42.1	84	48.4	458	52.3
Analysis	26	7.6	8	3.5	63	47.4	50	28.9	147	16.8
Synthesis	33	9.6	20	8.9	8	6.0	19	11.0	80	9.1
Evaluation	4	1.1	2	0.9	4	3.0	4	2.4	14	1.6
Total	344		226		133		173		876	

Figure 2

Distribution of Exam Questions According to School Type and Cognitive Level (Kocakaya & Gonen, 2010, p. 6)

Kocakaya and Gonen (2010) found that 72.5% of the questions were of the lower-order cognitive skills type. Only about 27.5% of the questions asked were at the higher-levels of cognitive domain such as analysis, synthesis and evaluation. On the other hand, only 6.3% were at the knowledge and 13.9% at the comprehension levels. The other 52.3% were at the application level.

Figure 3 shows the university entrance exams according to their levels of cognitive progress and frequencies of cognitive level questions.

Years	2000*		2001*		2002*		2003*		2004		2005		Total	
Questins Level	f	%	f	%	f	%	f	%	f	%	f	%	f	%
Knowledge	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Comprehension	0	0.0	0	0.0	1	5.3	0	0.0	3	15.7	0	0.0	4	3.5
Application	9	47.4	10	52.6	12	63.1	12	63.1	4	21.2	5	26.3	52	45.6
Analysis	10	52.6	7	36.8	5	26.3	5	26.3	9	47.4	13	68.4	49	43.0
Synthesis	0	0.0	2	10.6	1	5.3	1	5.3	0	0.0	0	0.0	4	3.5
Evaluation	0	0.0	0	0.0	0	0.0	1	5.3	3	15.7	1	5.3	5	4.4
Total	19		19		19		19		19		19		114	

*Azar(2005)

Figure 3

Distribution of the OSS Physics Questions According to the Cognitive Level as to Years (Kocakaya & Gonen, 2010, p. 7)

One noted difference is the increased emphasis on analysis type questions in the university entrance exam.

Summary

Higher-level thinking skills have been an area of interest for educational scholars since formal educational systems began. Although Benjamin Bloom's work did inspire additional interest in the 1960s, his work was only a result of ongoing educational discussions. Before the 1960s, higher-level thinking skills research studies were limited, however, the 1960s and 1970s produced many such studies. These studies ranged from investigations that examined tests in relation to teacher experience, content area being

taught, standardized tests and passing rates. The 1980s and 1990s provided very few new studies in this area. The new millennium has shown an increased interest in this area of research with several studies in Kentucky, Australia, Iran, and Turkey. Although the greatest amount of discussion within the United States occurred in the 1960s and 1970s, the past decade has rendered an increased attention on these areas primarily due to changes within the United States political system. Pressure from politicians and the public for accountability measures has produced an atmosphere where testing and assessment is an area of concern. This increased scrutiny will inevitably create many similar studies as educators continually strive to better understand higher-level thinking skills and their impact on learning.

Chapter III

METHODOLOGY

Purpose of the Study

Research shows that there is a link between higher-level thinking skills and increased student achievement. Raising the rigor, no matter what the circumstance, almost always increases achievement. Giving students the opportunity to practice thinking at higher-levels causes them to become better thinkers, not just for the duration of their school experience, but also in their lives beyond the classroom (Moore & Stanley, 2010).

"Brain researchers suggest that teachers (should) use a variety of higher-order questions in a supportive environment to strengthen the brain" (Cardellichio & Field, 1997, p. 33). "Meaningful learning requires teachers change their role from sage to guide, from giver to collaborator, from instructor to instigator" (Ó Murchú, 2003). "Since students learn from thinking about what they are doing, the teacher's role becomes one of stimulating and supporting activities that engage learners in critical thinking" (Bhattacharya, 2002).

If the development of higher-level thinking skills is important for students to be successful in the classroom and beyond, then teachers should incorporate cognitive (thinking) skill development in curriculum and instruction. This could be done by

teachers including activities to enhance higher-level thinking and constructing assessments that provide feedback on how well the students are progressing in developing critical thinking skills. Mentoring students in higher-level thinking means helping them to acquire the cognitive tools they need to move beyond a basic understanding of individual problems and, instead, see the collective dimensions to both their causes and solutions (Mills, 1959).

Given the importance of incorporating higher-level thinking into the learning process, there is a need to examine what teachers report regarding the importance and use of higher-level thinking questions in the development of summative assessments. Little research to date has been conducted on this topic.

The purpose of this study is to investigate a select sample of Minnesota high school teachers' reported knowledge, use, preparation, and importance of higher-level thinking summative assessment questions. Demographic and background variables of the sample group include gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status.

This chapter reports the methodology of the study regarding the following questions:

1. What relationships exist between teachers' demographics and their accuracy in identification of higher-level thinking questions?
   a. Demographic information and respondents who correctly identified at least two of the three higher-level thinking questions.

b. Demographic information and respondents who correctly identified all three of the higher-level thinking questions.

2. What relationships exist between teachers' demographics and:
    a. Their use and importance of higher-level thinking questions in summative assessments.
    b. The strategies used by their building principals to convey the importance of using higher-level thinking assessment questions.
    c. The preparation they received at their university/college in developing higher-level thinking questions in summative assessments and the quality/frequency of school or district staff development opportunities on developing higher-level thinking questions in their summative assessments.
    d. Their knowledge/expertise in developing higher-level thinking questions in summative assessments.
    e. The identification of a person responsible for preparing teachers for developing higher-level thinking questions in summative assessments.
3. What relationship exists between teacher ratings of importance of higher-level thinking questions in summative assessments and:
    a. Their school making Adequate Yearly Progress.
    b. Their district employing a designated person to monitor development of higher-level thinking questions in teacher constructed assessments.

Population

This study surveyed teachers from a select sample of urban, suburban, and rural high schools in Minnesota. The sample included two urban, two suburban, and three rural high schools.

The survey was distributed electronically to all core, allied, and fine arts subject area classroom teachers in the sample high schools. The principal of the high schools received an email stating the purpose of the study and a request for their participation. Subsequently, the principal authorized the survey and informed staff members about the survey and encouraged their responses.

According to the Minnesota Department of Education, in the 2011-2012 school year, there were 432.13 FTE (full-time equivalent) teachers employed at the schools in which surveys were distributed. An FTE is a measure of a single full-time teacher or its equivalent. As an example, one FTE could represent two half-time teachers, four 0.25 teachers, or one full-time teacher. The number of survey respondents was 216, a return of 50.0%.

The survey method allowed the researcher to make statistical inferences about the population using the sample. The success of survey research was dependent upon the representativeness of the population being studied. The survey population was exclusively high school teachers, so there was a close correlation between what the researcher was seeking to find through this study and the ability of the respondents to provide information.

Instrument

The instrument employed in the study was a 21 question survey that was designed to gather data to address the research questions. Once the survey instrument was finalized and formal IRB approval obtained (Appendix C), it was electronically uploaded to a secure survey analysis database.

The survey instrument followed the guidelines for writing quality survey questions and survey methods as developed through Survey Monkey (Survey Monkey Help Center-Tips for Writing Effective and Relevant Survey Questions). The survey has face and content validity in that it sought information from a qualified population that reflects the research questions of the study.

An expert panel was created to determine question validity in the survey section in which teachers were asked to identify those questions which involved higher-level thinking and those which involved lower-level thinking. Panel member agreement was 0.90. The panel was comprised of four doctoral level Professors of Education at St. Cloud State University. All four are experienced teachers and administrators knowledgeable in curriculum development, assessment, instruction, and Bloom's Taxonomy.

Advantages of the survey method, especially the online versions, include:

1. Less costly and less time to administer than mail, focus groups, case studies, and observations.
2. Quick response time is possible.
3. Large groups can be accessed at once.

4. Statistical applications can be applied with greater confidence with larger numbers of respondents.
5. Comparisons between respondent groups can be easily made.
6. The sample population can be more representative of the greater population if the survey method is properly employed since larger numbers of respondents are possible in a broader arena.
7. They can be used to study many kinds of information such as attitudes, beliefs, values, and past behaviors.
8. They can be focused to obtain specific data (Adapted from *Wikipedia*) (Survey methodolog, n.d.).

Disadvantages of using the survey method include:

1. Non-response or non-response on one or a few items.
2. Depends totally on respondents' motivation, honesty, memory.
3. Respondents may lean toward presenting themselves in a favorable light.
4. Bias: due to the survey designer in developing questions.
5. Bias: due to the respondents. Those who respond, as a group, may be more motivated to do so because of a whole range of things. The same is true for those who do not respond, as a group.
6. Strength of choice on some types of survey questions is lessened. For example, "Moderately agree" is not as definitive an answer as a yes or no response. But, even yes or no responses can also lead to the same issue since

some may pick “no” even if s/he really believes it is just once in a while. It is harder to get strong data connections.

Data Collection and Procedures

The electronic survey was completed using Survey Monkey (Appendix B). Respondents were able to complete the survey electronically in about 10-15 minutes and submit the document anonymously. The survey included demographic information, as stated in the research questions. Survey items were created to secure responses to the following:

1. Identify higher-level thinking questions from a list of six questions which included both higher and lower-level thinking questions.
2. Rate the importance of higher-level thinking questions in summative assessments.
3. Rate their own expertise, use, and preparation for incorporating higher-level thinking questions into summative assessments.
4. Rate teacher use of higher-level thinking questions in summative assessments.
5. Rate to what extent administration evaluates teacher use of higher-level thinking questions.

Nine questions focused on demographic information, including gender, years teaching, years in current position, subject area taught, highest level of education

achieved, location of school, size of school, grade level configuration of school, and AYP status.

Study participants received a brief email explanation of the purpose and procedures of the study (Appendix A), the link to the on-line survey (Appendix B), and a statement that the survey process was endorsed by the building principal. The survey permitted responses only by the participant and once completed and submitted, the survey was no longer accessible to the participant from the same computer. Participants were allotted 3 weeks to respond to the survey; two email reminders were sent at 7 and 18 days. The survey was removed from on-line access at the end of 28 days.

Data Analysis

Descriptive statistics were calculated to determine the frequency distributions and percentages of participant responses for each item. Correlation of differences between demographic information from the participant and questionnaire items was analyzed.

Pearson's Chi Square and Pearson's correlations were used to determine the strength of relationship between participant accuracy on identification of higher-level thinking questions and the variables represented by the demographic, and internal and external factors listed in research questions. A p value less than 0.05 was identified as significant.

Statistical analysis was conducted using the software program available through Survey Monkey and Statistical Package for the Social Sciences.

Limitations

According to Roberts (2008), limitations of a study are factors that affect a study but are not under the control of the researcher. Limitations can further be explained as the constraints that influence the generalizability or interpretation of the results of the study.

The following are some possible limitations of the study:

1. This study will only survey teachers in Minnesota. Teachers outside of Minnesota will not be surveyed.
2. This study will only survey a small sample of teachers as compared to the total population of teachers in Minnesota.
3. This study will only survey Minnesota high school teachers. Teachers of lower grades will not be surveyed.
4. This study will only survey public school teachers in urban, suburban, and rural school districts. Private school teachers will not be included in this survey.
5. The study must be delivered and accessed through the filtering systems of school email and internet services. Failure to receive the invitation to take part in the study could result in a limited participation rate.
6. Respondents need to accurately report data. Survey results are based on the honest response of participants.
7. The online survey delivery system must work reliably to provide results consistent with the responses of the participants.

8. Teachers may simply choose not to participate due to concern that their information may not be confidential and could be accessed by their principal.

Summary

The purpose of this study is to examine teachers reported knowledge, use, preparation, and importance of higher-level thinking questions in summative assessments. The results will be used to assist teachers and administrators in understanding the perceived status of teachers' training and use of higher-level thinking questions in helping students achieve greater success in the classroom and beyond.

Chapter IV

FINDINGS

Purpose of the Study

The purpose of this study is to investigate a select sample of Minnesota high school teachers' reported knowledge, use, preparation, and importance of higher-level thinking summative assessment questions. Demographic and background variables of the study sample include gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status.

This chapter reports the findings of the study. The data are analyzed and findings organized according to each research question:

1. What relationships exist between teachers' demographics and their accuracy in identification of higher-level thinking questions?
    a. Demographic information and respondents who correctly identified at least two of the three higher-level thinking questions.
    b. Demographic information and respondents who correctly identified all three of the higher-level thinking questions.

2. What relationships exist between teachers' demographics and:
    a. Their use and importance of higher-level thinking questions in summative assessments.
    b. The strategies used by their building principals to convey the importance of using higher-level thinking assessment questions.
    c. The preparation they received at their university/college in developing higher-level thinking questions in summative assessments and the quality/frequency of school or district staff development opportunities on developing higher-level thinking questions in their summative assessments.
    d. Their knowledge/expertise in developing higher-level thinking questions in summative assessments.
    e. The identification of a person responsible for preparing teachers for developing higher-level thinking questions in summative assessments.
3. What relationship exists between teacher ratings of importance of higher-level thinking questions in summative assessments and:
    a. Their school making Adequate Yearly Progress.
    b. Their district employing a designated person to monitor development of higher-level thinking questions in teacher constructed assessments.

The data were collected using Survey Monkey as a data collection tool. Statistical Package for the Social Sciences (SPSS) Version 19, Release 19.0.0.2 and Minitab 16.1.1 were used for the data analysis. Pearson Chi Square and p-values were

used to determine if relationships were statistically significant at a p-value < 0.05. Tables indicate percentages of responses in each research question by demographics. Tables also illustrate respondent ratings, prioritization, or multiple-choice selections.

Description of the Sample

In this study, high school teachers in two urban, two suburban, and three rural schools in Minnesota were surveyed. The schools were selected by enrollment and location to represent a cross section of high schools in Minnesota.

The principals of the high schools were sent an email from the researcher stating the purpose of the study and a request to participate (Appendix A). The principals authorized the survey and informed their staff members about the survey and encouraged response. The surveys were then distributed electronically to all core, allied, and fine arts subject area classroom teachers in the selected high schools.

According to the Minnesota Department of Education, in the 2011-2012 school year, there were 432.13 FTE (full-time equivalent) teachers employed at the schools in which surveys were distributed. An FTE is a measure of a single full-time teacher or its equivalent. As an example, one FTE could represent two half-time teachers, four 0.25 teachers, or one full-time teacher. The number of survey respondents was 216, a return of 50%.

The study results will provide school leaders with information regarding the sample group teachers' perceptions about the importance, knowledge, preparation, and use of higher-level thinking questions in their summative assessments.

Research Question 1

The survey asked respondents to identify three higher-level thinking questions from a list of six example questions. Data were analyzed to determine if any statistically significant relationships were found between those who correctly identified two or all three of the higher-level thinking questions and demographic information which include gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status.

What relationships exist between teachers' demographics and their accuracy in identification of higher-level thinking questions?

a. *Demographic information and respondents who correctly identified at least two of the three higher-level thinking questions.*
b. *Demographic information and respondents who correctly identified all three of the higher-level thinking questions.*

Research Question 1 Part A has one finding. It is described below.

The analysis of demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), and whether the respondent correctly identified at least two higher-level thinking questions yielded p-values > 0.05, and, therefore, were dismissed as lacking statistically significant data. The frequency distribution for this item is reported below in Table 1.

Table 1

Frequency Distribution of Correct Identification of Two or More Higher-Level Thinking Questions by Gender (n=183)

Selections	Male			Female		
	N	ACT	%	N	ACT	%
Correct identification of two or more higher-level thinking questions.	61	49	80.3	122	102	83.6

ACT = Actual number who correctly identified two or more correctly.

Research Question 1 Part B has one finding. It is described below.

Analysis of demographics and those who correctly identified all three higher-level thinking questions showed a relationship with gender. Using a Fisher's proportions test, a $p = 0.024$ was calculated. This means that response choice(s) and gender were likely not independent.

The significant p-value resulted from the fact that 23.0% (n=61) of males, and 9.8% (n=122) of females correctly identified all three of the higher-level thinking questions. The frequency distribution for this item is reported below in Table 2.

Table 2

Frequency Distribution of Correct Identification of Three Higher-Level Thinking Questions by Gender (n=183)

Selections	Male			Female		
	N	ACT	%	N	ACT	%
Correct identification of three higher-level thinking questions.	61	14	23.0	122	12	9.8

ACT = Actual number who correctly identified all three.

Research Question 2

The survey asked respondents to answer questions related to the use, importance, and their preparation in using higher-level thinking questions within their summative assessments. The data were analyzed to determine if any statistically significant relationships were found between their responses and demographic information which included gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status.

What relationships exist between teachers' demographics and:

a. *Their use and importance of higher-level thinking questions in summative assessments.*
b. *The strategies used by their building principals to convey the importance of using higher-level thinking assessment questions.*
c. *The preparation they received at their university/college in developing higher-level thinking questions in summative assessments and the quality/frequency of school or district staff development opportunities on developing higher-level thinking questions in their summative assessments.*
d. *Their knowledge/expertise in developing higher-level thinking questions in summative assessments.*
e. *The identification of a person responsible for preparing teachers for developing higher-level thinking questions in summative assessments.*

Research Question 2 Part A has four findings. They are described below.

Respondents were asked to identify how often they incorporated higher-level thinking questions into their summative assessments. Of a total of 189 study respondents who answered this question, 44.4 % reported between 50% and 79% of the time and, 24.9% reported between 20% and 49% of the time. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of

school, size of school, grade level configuration of school, and AYP status), no statistically significant data were found when applying the Pearson Chi Square test. The frequency distribution for this item is reported below in Table 3.

Table 3

Frequency Distribution of Respondents' Reported Use of Higher-Level Thinking Questions in Summative Assessments (n=189)

Selections	N	%
More than 80%	38	20.1
About 50% to 79%	84	44.4
About 20% to 49%	47	24.9
Less than 20%	20	10.6
Did not answer	32	

Gender demonstrated a relationship with response selection when participants were asked to select the range most representative of how often they incorporate higher-level thinking questions into summative assessments. A Pearson Chi Square value of 9.216, a $p = 0.027$, n=189, and three degrees of freedom were found. This means that response choice(s) and gender were likely not independent.

The significant Chi Square likely resulted from a larger percentage of males (32.8%, n=21) who selected the range of 20-49% of the time that they incorporated higher-level thinking questions into summative assessments than did females (20.8%, n=26). A larger percentage of males (26.6%, n=17) also selected the range of more than 80% of the time when compared to females

(16.8%, n=21). However, a larger percentage of females (52.0%, n=65) selected the 50-79% range when compared to males (29.7%, n=19). Males tended to select both a lower and a higher range of time that they incorporated higher-level thinking questions into summative assessments than females. However, females were more consistent in the selection of a specific range (50-79%) than were males. The frequency distribution for this item is reported below in Table 4

Table 4

Frequency Distribution of Respondents' Reported Incorporation of Higher-Level Thinking Questions in Summative Assessments by Gender. (n=189)

	Male			Female		
Selections	N	EXP	%	N	EXP	%
More than 80%	17	12.9	26.6	21	25.1	16.8
About 50% to 79%	19	28.4	29.7	65	55.6	52.0
About 20% to 49%	21	15.9	32.8	26	31.1	20.8
Less than 20%	7	6.8	10.9	13	13.2	10.4

Exp = SPSS Expected Value

Location of school (rural, suburban, or urban) showed a relationship when participants were asked to select the range most representative of how often they incorporated higher-level thinking skills into summative assessments. A Pearson Chi Square value of 18.166, a $p = 0.006$, n=189, and six degrees of freedom was found. This means that response choice(s) and location of school were likely not independent.

The significant Chi Square likely resulted from a lower percentage of urban teachers (5.9%, n=4) who selected the range of more than 80% compared

to either suburban (25.0%, n=22) or rural teachers (36.4% n=12). A higher percentage of urban teachers (35.3%, n=24) selected the range of 20% to 49% of the time compared to either rural (18.2%, n=6) or suburban teachers (19.3%, n=17). A lower percentage of rural teachers (33.3%, n=11) reported incorporating higher-level thinking questions into summative assessments in the range of 50% to 79% of the time than suburban, 45.5% (n=40) or urban teachers, 48.5% (n=33). Interestingly, nearly 70% of suburban and rural teachers reported incorporating higher-level thinking questions into summative assessments 50% or more of the time compared to 54% of urban teachers. The frequency distribution for this item is provided below in Table 5.

Table 5

Frequency Distribution of Respondents' Reported Incorporation of Higher-Level Thinking Questions into Summative Assessments by Location of School (n=189)

	Rural			Suburban			Urban		
Selections	N	EXP	%	N	EXP	%	N	EXP	%
More than 80%	12	6.6	36.4	22	17.7	25.0	4	13.7	5.9
About 50% to 79%	11	14.7	33.3	40	39.1	45.5	33	30.2	48.5
About 20% to 49%	6	8.2	18.2	17	21.9	19.3	24	16.9	35.3
Less than 20%	4	3.5	12.1	9	9.3	10.2	7	7.2	10.3

EXP = SPSS Expected Value

Respondents were asked to rate how important it is to incorporate higher-level thinking questions into summative assessments. Of the 191 respondents, 63.9% (n=122) rated it very important while 33.5% (n=64) rated it as somewhat

important. Only 2.6% (n=5) of respondents rated the importance of incorporating higher-level thinking questions in summative assessments as not really that important or not at all important. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant data were found when applying the Pearson Chi Square test. The frequency distribution for this item is reported below in Table 6.

Table 6

Frequency Distribution of Respondents' Rating of Importance of Incorporating Higher-Level Thinking Questions into Summative Assessments (n=191)

Selections	N	%
Very important	122	63.9
Somewhat important	64	33.5
Not really that important	4	2.1
Not at all important	1	0.5
Did not answer	30	

Research Question 2 Part B has two findings. They are described below.

Respondents were asked to identify strategies that principals use to encourage higher-level thinking questions in summative assessments. Of the 187 who answered this question, 47.6% (n=89) indicated that their principal verbally encouraged teachers to use higher-level thinking questions in their assessments. Nearly 31% (n=57) reported that the principal does not discuss or provide

direction about the development of higher-level thinking questions on summative assessments. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant data were found when applying the Pearson Chi Square test. Although not statistically significant, it is interesting that 30.5% (n=57) of the respondents reported that the principal does not discuss or provide direction to teachers about incorporating higher-level thinking questions in summative assessments. The frequency distribution for this item is reported below in Table 7.

Table 7

Frequency Distribution of Respondents' Reported Principal's Strategies to Convey the Importance of Higher-Level Thinking Questions (n=187)

Selections	N	%
My principal provides professional development opportunities on this topic.	58	31.0
My principal discusses the importance of using higher-level thinking questions with teachers in their performance review.	46	24.6
My principal verbally encourages teachers to use higher-level thinking questions in their assessments.	89	47.6
My principal does not discuss or provide direction to teachers about the development and use of higher-level thinking questions in assessments.	57	30.5
Did not answer.	34	

Respondents were also asked to identify how they believed principals should encourage the use of higher-level thinking questions. Of the 185 who answered this question, 58.9% (n=109) indicated they thought principals should offer staff development and 56.8% (n=105) reported that principals should encourage departmental discussions. Other responses to this question found that 33.0% (n=61) of those who answered this question reported that principals should have discussions with individual teachers during their performance review about the importance of using higher-level thinking questions on summative assessments and 22.7% (n=42) reported that principals should let teachers decide how to assess learning. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant data were found when applying the Pearson Chi Square test. The frequency distribution for this item is reported below in Table 8.

Table 8

Frequency Distribution of Respondents' Reported Strategies Principals Should Use to Encourage Teachers to Use Higher-Level Thinking Questions in Summative Assessments (n=185)

Selections	N	%
Principals need to actively promote and encourage the use of higher-level thinking questions on assessments through providing staff development on the topic.	109	58.9
Principals should have discussions with individual teachers during their performance reviews on the importance of using higher-level thinking questions on their summative assessments.	61	33.0
Principals should encourage departmental discussions on the use of higher-level thinking questions in summative assessments.	105	56.8
Principals should let teachers decide how to assess learning.	42	22.7
Did not answer.	36	

Research Question 2 Part C has five findings. They are described below.

Respondents were asked to identify the preparation they received in their teaching certification program regarding the incorporation of higher-level thinking questions in summative assessments. Of the 183 respondents who answered this question, 65.0% (n=119) reported that their teacher certification program included some discussion about this topic while 18.6% (n=34) reported that their teacher certification program did not include discussions regarding higher-level thinking questions. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant data were

found when applying the Pearson Chi Square test. The frequency distribution for this item is reported below in Table 9.

Table 9

Frequency Distribution Respondents' Reported Preparation Received in Teaching Certification Programs Regarding Higher-Level Thinking Questions in Summative Assessments (n=183)

Selections	N	%
My teacher certification program fully prepared me to incorporate higher-level thinking questions into teacher-developed assessments.	30	16.4
My teacher certification program included some discussions that addressed higher-level thinking questions.	119	65.0
My teacher certification program did not include discussions regarding higher-level thinking questions.	34	18.6
Did not answer.	38	

Respondents were also asked to identify how often higher-level thinking assessment has been a topic of staff development in the schools in which they taught. Of the 184 respondents who answered this question, 47.3% (n=87) reported that the topic had been discussed between one and three times in the past 2 years. A combined total of 37% (n=68) reported that the topic of higher-level thinking questions had been a topic of their staff development program from four to seven or more times in the past 2 years, and 15.8% (n=29) reported that the topic was not discussed in the past 2 years. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant

data were found when applying the Pearson Chi Square test. The frequency distribution for this item is reported below in Table 10.

Table 10

Frequency Distribution of Respondents' Reported Frequency of Higher-Level Thinking Questions as a Topic of Staff Development in the Schools in Which They Have Taught (n=184)

Selections	N	%
It has been discussed more than 7 times in the past 2 years.	16	8.7
It has been discussed 4-7 times in the past 2 years.	52	28.3
It has been discussed 1-3 times in the past 2 years.	87	47.3
It has not been discussed in the past 2 years.	29	15.8
Did not answer.	37	

Respondents were also asked to identify how effective they believed the staff development programs they had attended were, in helping teachers develop higher-level thinking questions in summative assessments. Of the 186 respondents who answered this question, over half of the respondents 58.1% (n=108) rated staff development programs as average. Nearly 24% (n=44) rated staff development programs on the topic of higher-level thinking questions in assessments as below average or poor; 14.0% (n=26) reported not attending any staff development programs on the topic and, only 4.3% (n=8) of respondents rated staff development on the topic of higher-level thinking questions in assessments as excellent. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration

of school, and AYP status), no statistically significant data were found when applying the Pearson Chi Square test. The frequency distribution for this item is reported below in Table 11.

Table 11

Frequency Distribution of Respondents' Reported Effectiveness of Staff Development Programs in Incorporating Higher-Level Thinking Questions in Summative Assessments (n=186)

Selections	N	%
Excellent	8	4.3
Average	108	58.1
Below Average	29	15.6
Poor	15	8.1
I have not attended any staff development programs on this topic.	26	14.0
Did not answer.	35	

Years in their current position showed a statistically significant relationship when data were analyzed on how effective staff development programs were in helping develop higher-level thinking questions in summative assessments. A Pearson Chi Square value of 37.862, a $p = 0.002$, n=186 and 16 degrees of freedom was found. This means that response choice(s) and years in current position were likely not independent.

The significant Chi Square likely resulted from the fact that a larger percentage of teachers with 3 or less years' experience (82.1%, n=32) than teachers with 21 or more years of experience (33.3%, n=5) selected the category of average as their response when asked to rate the effectiveness of their staff

development programs in helping teachers develop higher-level thinking questions into their summative assessments. No teachers with 3 or less years of experience selected categories of below average or poor (0%, n=0) when compared with teachers with 21 or more years of experience (60%, n=9) in rating the effectiveness of staff development programs in helping develop higher-level thinking questions in summative assessments. The frequency distribution for this item is reported below in Table 12.

Table 12

Frequency Distribution Respondents' Reported Effectiveness of Staff Development Programs in Helping Develop Higher-Level Thinking Questions in Summative Assessments (n=186)

	3 or Less			4-11			11-15			16-20			21 or More		
Selections	N	EXP	%	N	EXP	%	N	EXP	%	N	EXP	%	N	EXP	%
Excellent	1	1.7	2.6	2	2.6	6.6	2	1.6	5.4	1	1.5	2.9	0	0.6	0.00
Average	32	22.6	82.1	35	35.4	57.4	20	21.5	54.1	6	19.7	47.1	5	8.7	33.3
Below Average	0	6.1	0.0	7	9.5	11.5	7	5.8	18.9	8	5.3	23.5	7	2.3	46.7
Poor	0	3.1	0.0	3	4.9	4.9	6	3.0	16.2	4	2.7	11.8	2	1.2	13.3
I have not attended any	6	5.5	15.4	12	8.5	19.7	2	5.2	5.4	5	4.8	14.7	1	2.1	6.7

The subject area taught showed a statistically significant relationship when data were analyzed on the perceived effectiveness of staff development programs on the development of higher-level thinking questions within summative assessments. A Pearson Chi Square value of 31.482, a $p = 0.049$, n=186, and 20 degrees of freedom was found. This means that response choice(s) and subject area taught were likely not independent.

The statistically significant Chi Square value likely resulted from a lower percentage of teachers who selected science (4.3%, n=1) or other (9.1%, n=4) as their primary teaching responsibility rated their staff development programs as below average when compared to teachers who selected language arts (21.2% n=7), math (19.4%, n=6), social studies (19.0%, n=4), or allied-fine arts (20.6%, n=7) as their primary teaching responsibility. Lower percentages of teachers who selected math (3.2%, n=1), allied-fine arts (0%, n=0), and other (2.3%, n=1), as their primary teaching responsibility rated the staff development programs as poor than did those who selected language arts (18.2%, n=6), science (13.0%, n=3), or social studies (19.0%, n=4) as their primary teaching responsibility. Lower percentages of teachers who selected language arts (6.1%, n=2), social studies (4.8%, n=1), or allied-fine arts (5.9%, n=2) as their primary teaching responsibility reported that they had not attended any staff development on incorporating higher-level thinking skills into summative assessments when compared to those who selected math (16.1%, n=5), science (26.1%, n=6), or other (22.7%, n=10) as their primary teaching responsibility. The frequency distribution for this item is reported below in Table 13.

Table 13

Frequency Distribution of Respondents' Reported Effectiveness of Staff Development Programs in Helping Develop Higher-Level Thinking Questions in Summative Assessments by Subject Area Taught (n=186)

	Language Arts			Math			Science			Social Studies			Allied-Fine Arts			Other		
Selections	N	EXP	%	N	EXP	%	N	EXP	%	N	EXP	%	N	EXP	%	N	EXP	%
Excellent	1	1.4	3.0	0	1.3	0.0	1	1.0	4.3	1	0.9	4.8	3	1.5	8.8	2	1.9	4.5
Average	17	19.2	51.5	19	18.0	61.3	12	13.4	52.2	11	12.2	52.4	22	19.7	64.7	27	25.5	61.4
Below Average	7	5.1	21.2	6	4.8	19.4	1	3.6	4.3	4	3.3	19.0	7	5.3	20.6	4	6.9	9.1
Poor	6	2.7	18.2	1	2.5	5.3	3	1.9	13.0	4	1.7	19.0	0	2.7	0.0	1	3.5	2.3
I have not attended any	2	4.6	6.1	5	4.3	16.1	6	3.2	26.1	1	2.9	4.8	2	4.8	5.9	10	6.2	22.7

EXP=SPSS Expected Value

Research Question 2 Part D has two findings. They are described below.

Respondents were asked to rate their own knowledge/expertise related to developing higher-level thinking questions in their summative assessments. Of the190 respondents who answered this question, 66.8% (n=127) rated their knowledge level as about average, 18.9% (n=36) rated their knowledge/expertise as very high, and 14.2% (n=27) reported not being confident in knowledge/expertise regarding developing higher-level thinking questions for summative assessments. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant data were found when applying the Pearson Chi Square test. The frequency distribution for this item is reported below in Table 14.

Table 14

Frequency Distribution of Respondents Rating of Knowledge/Expertise in Developing Higher-Level Thinking Questions for Summative Assessments (n=190)

Selections	N	%
Very high knowledge and expertise.	36	18.9
About average knowledge and expertise.	127	66.8
I am not really confident in my level of knowledge and expertise.	27	14.2
Did not answer	31	

Subject area taught showed a relationship with how respondents rated their knowledge/expertise in developing higher-level thinking questions for their summative assessments. A Pearson Chi Square value of 34.721, a $p < 0.001$, n=190 and 10 degrees of freedom were found. This means that response choice(s) about knowledge/expertise in incorporating higher-level thinking questions into summative assessments and subject area taught were likely not independent.

The significant Chi Square likely resulted from a larger percentage of language arts teachers who reported very high knowledge and expertise (45.7%, n=16) in developing higher-level thinking questions when compared with math teachers (6.5%, n=2), science teachers (13.0%, n=3) or those who selected other as a teaching category (6.5%, n=3). A lower percentage of language arts teachers reported about average knowledge and expertise (48.6%, n=17) in developing higher-level thinking questions when compared to math teachers (80.6%, n=25) or social studies teachers (76.2%, n=16). A larger percentage of teachers who selected other as their teaching category (28.3%, n=13) reported that they were not confident in their knowledge and expertise when compared to teachers who selected teaching categories of language arts (5.7%, n=2) or social studies (0%, n=0) as the areas of primary teaching responsibility. The frequency distribution for this item is reported below in Table 15.

Table 15

Frequency Distribution of Respondents' Rating of Knowledge/Expertise in Developing Higher-Level Thinking Questions for Summative Assessments (n=190)

	Language Arts			Math			Science			Social Studies			Allied-Fine Arts			Other		
Selections	N	EXP	%	N	EXP	%	N	EXP	%	N	EXP	%	N	EXP	%	N	EXP	%
Very high knowledge and expertise.	16	6.6	45.7	2	5.9	6.5	3	4.4	13.0	5	4.0	23.8	7	6.4	20.6	3	8.7	6.5
About average knowledge and expertise.	17	23.4	48.6	25	20.7	80.6	16	15.4	69.6	16	14.0	76.2	23	22.7	67.6	30	30.7	65.2
I am not really confident in my level of knowledge and expertise.	2	5.0	5.7	4	4.4	12.9	4	3.3	17.4	0	3.0	0.0	4	4.8	11.8	13	6.5	28.3

EXP = SPSS Expected Value

Research Question Part E has one finding. It is described below.

Respondents were asked to rank, in priority order, levels of responsibility for preparing teachers to develop higher-level thinking questions on summative assessments (1=most responsible to 5=least responsible). The response total (RT) is the weighted sum of the rankings received for each of the possible choices. For example, if the principal was given a "1" ranking 25 times, and a "2" ranking 20 times by respondents, his/her response total would be calculated as follows: 1x25+2x20 = 65. The response count (RC) is the total of the rankings given by respondents for each of the five possible choices, thus the RC for the principal in the above example is 45 (25+20). The response average (RA) would be the average ranking that any one person received (RT divided by RC). Of the 180 teachers who responded to this question, the institutions that prepare teachers were ranked as most responsible (RA=2.26) and the teachers, themselves, were ranked second as most responsible (RA=2.39). The principal (RA=3.87) was ranked as least responsible for helping teachers develop higher-level thinking questions in summative assessments. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant data were found according to the Pearson Chi Square test. The ranking summary for this item is reported below in Table 16.

Table 16

Ranking Summary for Preparing Teachers to Incorporate Higher-Level Thinking Questions in Summative Assessments (n=180)

Selections	RA	RT	RC
Institutions that prepare teachers.	2.26	407	180
Staff development programs offered for teachers at their schools.	2.80	504	180
The principal of the school.	3.87	697	180
Teacher coaches or mentors.	3.58	645	180
Teachers themselves.	2.39	428	179
Did not answer.	41		

RA = Response Average, RT = Response Total, RC = Response Count

Research Question 3

The survey asked respondents to answer questions related to how important they feel it is to use higher-level thinking questions on summative assessments. The data were analyzed to determine if any statistically significant relationships were found between their ratings and their school's Adequate Yearly Progress (AYP) status and, whether or not their school employs someone responsible for oversight of higher-level thinking (e.g., a higher-level thinking coach). This question was developed to determine if there would be any conclusions regarding teacher attitude toward higher-level thinking questions and a standard Minnesota achievement measure such as their school making AYP. This question also was framed to assess whether or not having access to a person who coaches teachers on higher-level thinking questions influences teacher perception of the importance of higher-level thinking questions in assessments.

What relationships exist between teacher ratings of importance of higher-level thinking questions in summative assessments and:

a. Their school making Adequate Yearly Progress.

b. Their district employing a designated person to monitor development of higher-level thinking assessment questions.

Research Question Part A has one finding. It is described below.

Respondents were asked if their school was making Adequate Yearly Progress. Of the 205 respondents, 21.0% (n=43) responded yes, 52.2% (n=107) responded no, and 26.8% (n=55) did not know. When analyzed according to the importance teachers assigned to higher-level thinking skills, no statistically significant results were found when applying the Pearson Chi Square test. The frequency distribution for this item is reported below in Table 17.

Table 17

Frequency Distribution for Response to Whether Their School is Making (AYP) Adequate Yearly Progress (n=205)

Selections	N	%
Yes	43	21.0
No	107	52.2
I don't know	55	26.8
Did not answer	16	

Research Question 3 Part B has one finding. It is described below.

Respondents were asked if their district employed an individual whose responsibilities included overseeing the use/development of higher-level thinking questions in teacher-constructed assessments. Of the 204 respondents, 42.6% (n=87) responded that they did not know, 31.9% (n=65) responded no one is

employed with that responsibility, and 25.5% (n=52) responded yes that a person is employed with that responsibility. When data were analyzed comparing teacher responses to this question and the importance the teachers assigned to higher-level thinking skills, no statistically significant results were found when applying the Pearson Chi Square test. The frequency distribution for this item is reported below in Table 18.

Table 18

Frequency Distribution for Response to Whether Their District Employs a Higher-Level Thinking Coach (n=204)

Selections	N	%
Yes	52	25.5
No	65	31.9
I don't know	87	42.6
Did not answer	17	

Chapter V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Purpose of the Study

The purpose of this study is to investigate a select sample of Minnesota high school teachers' reported knowledge, use, preparation, and importance of higher-level thinking summative assessment questions. Demographic and background variables of the study group including gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status are also considered in the results of this study.

This chapter reports the conclusions and recommendations of the study which are organized according to each research question:

1. What relationships exist between teachers' demographics and their accuracy in identification of higher-level thinking questions?
    a. Demographic information and respondents who correctly identified at least two of the three higher-level thinking questions.
    b. Demographic information and respondents who correctly identified all three of the higher-level thinking questions.
2. What relationships exist between teachers' demographics and:

a. Their use and importance of higher-level thinking questions in summative assessments.
b. The strategies used by their building principals to convey the importance of using higher-level thinking assessment questions.
c. The preparation they received at their university/college in developing higher-level thinking questions in summative assessments and the quality/frequency of school or district staff development opportunities on developing higher-level thinking questions in their summative assessments.
d. Their knowledge/expertise in developing higher-level thinking questions in summative assessments.
e. The identification of a person responsible for preparing teachers for developing higher-level thinking questions in summative assessments.

3. What relationship exists between teacher ratings of importance of higher-level thinking questions in summative assessments and:
   a. Their school making Adequate Yearly Progress.
   b. Their district employing a designated person to monitor development of higher-level thinking questions in teacher constructed assessments.

The data were collected using Survey Monkey as a data collection tool. Statistical Package for the Social Sciences (SPSS) Version 19, Release 19.0.0.2 and Minitab 16.1.1 were used for the data analysis. Pearson Chi Square and p-values were used to determine if relationships were statistically significant at a p-value < 0.05.

Tables indicate percentages of responses in each research question by demographics. Tables also illustrate respondent ratings, prioritization, or multiple-choice selections.

Research Question 1

The survey asked respondents to identify three higher-level thinking questions from a list of six example questions. Data were analyzed to determine if any statistically significant relationships were found between those who correctly identified two or all three of the higher-level thinking questions and demographic information which included gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status.

What relationships exist between teachers' demographics and their accuracy in identification of higher-level thinking questions?

a. *Demographic information and respondents who correctly identified at least two of the three higher-level thinking questions.*
b. *Demographic information and respondents who correctly identified all three of the higher-level thinking questions.*

The following conclusions were found for Research Question 1 Part A:

1. The analysis of demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), and whether the respondent correctly identified at least two higher-level thinking questions yielded p-values > 0.05, and, therefore, were dismissed as lacking statistically significant data.

The following conclusions were found for Research Question 1 *Part B:*

1. Analysis of demographics and those who correctly identified all three higher-level thinking questions showed a relationship with gender. Using a Fisher's proportions test, a $p = 0.024$ was calculated. This means that response choice(s) and gender were likely not independent. The significant p-value resulted from the fact that 23.0% (n=61) of males, and 9.8% (n=122) of females correctly identified all three of the higher-level thinking questions. This finding was not anticipated and as noted later in this chapter, should be further researched to determine the extent to which gender influences or affects the correct identification of higher-level thinking questions. It is important to note as well, that the findings were only significant between gender and correct identification of all three higher-level thinking questions, but not for the correct identification of two or more.

Research Question 2

The survey asked respondents to answer questions related to the use, importance, and their preparation in using higher-level thinking questions within their summative assessments. The data were analyzed to determine if any statistically significant relationships were found between their responses and demographic information which included gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status.

What relationships exist between teachers' demographics and:

a. Their use and importance of higher-level thinking questions in summative assessments.
b. The strategies used by their building principals to convey the importance of using higher-level thinking assessment questions.
c. The preparation they received at their university/college in developing higher-level thinking questions in summative assessments and the quality/frequency of school or district staff development opportunities on developing higher-level thinking questions in their summative assessments.
d. Their knowledge/expertise in developing higher-level thinking questions in summative assessments.
e. The identification of a person responsible for preparing teachers for developing higher-level thinking questions in summative assessments.

The following conclusions were found for Research Question 2 Part A:

1. Respondents were asked to identify how often they incorporated higher-level thinking questions into their summative assessments. Of a total of 189 study respondents who answered this question, 44.4 % reported between 50% and 79% of the time and, 24.9% reported between 20% and 49% of the time. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant results were found when applying the Pearson Chi Square test.
2. Gender demonstrated a relationship with response selection when participants were asked to select the range most representative of how often they incorporate higher-level thinking questions into summative assessments. A Pearson Chi Square value of 9.216, a $p = 0.027$, n=189, and three degrees of

freedom were found. This means that response choice(s) and gender were likely not independent. The significant Chi Square likely resulted from a larger percentage of males (32.8%, n=21) who selected the range of 20-49% of the time that they incorporated higher-level thinking questions into summative assessments than did females (20.8%, n=26). A larger percentage of males (26.6%, n=17) also selected the range of more than 80% of the time when compared to females (16.8%, n=21). However, a larger percentage of females (52.0%, n=65) selected the 50-79% range when compared to males (29.7%, n=19). Males tended to select both a lower and a higher range of time that they incorporated higher-level thinking questions into summative assessments than females. However, females were more consistent in the selection of a specific range (50-79%) than were males.

3. Location of school (rural, suburban, or urban) showed a relationship when participants were asked to select the range most representative of how often they incorporated higher-level thinking skills into summative assessments. A Pearson Chi Square value of 18.166, a $p = 0.006$, n=189, and six degrees of freedom was found. This means that response choice(s) and location of school were likely not independent. The significant Chi Square likely resulted from a lower percentage of urban teachers (5.9%, n=4) who selected the range of more than 80% compared to either suburban (25.0%, n=22) or rural teachers (36.4% n=12). A higher percentage of urban teachers (35.3%, n=24) selected the range of 20% to 49% of the time compared to either rural

(18.2%, n=6) or suburban teachers (19.3%, n=17). A lower percentage of rural teachers (33.3%, n=11) reported incorporating higher-level thinking questions into summative assessments in the range of 50% to 79% of the time than suburban, 45.5% (n=40) or urban teachers, 48.5% (n=33). Interestingly, nearly 70% of suburban and rural teachers reported incorporating higher-level thinking questions into summative assessments 50% or more of the time compared to 54% of urban teachers.

4. Respondents were asked to rate how important it is to incorporate higher-level thinking questions into summative assessments. Of the 191 respondents, 63.9% (n=122) rated it very important while 33.5% (n=64) rated it as somewhat important. Only 2.6% (n=5) of respondents rated the importance of incorporating higher-level thinking questions in summative assessments as not really that important or not at all important. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant results were found when applying the Pearson Chi Square test.

The following conclusions were found for Research Question 2 Part B:

1. Respondents were asked to identify strategies that principals use to encourage higher-level thinking questions in summative assessments. Of the 187 who answered this question, 47.6% (n=89) indicated that their principal verbally

encouraged teachers to use higher-level thinking questions in their assessments. Nearly 31% (n=57) reported that the principal does not discuss or provide direction about the development of higher-level thinking questions on summative assessments. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant results were found when applying the Pearson Chi Square test.

2. Respondents were also asked to identify how they believed principals should encourage the use of higher-level thinking questions. Of the 185 who answered this question, 58.9% (n=109) indicated they thought principals should offer staff development and 56.8% (n=105) reported that principals should encourage departmental discussions. Other responses to this question found that 33.0% (n=61) of those who answered this question reported that principals should have discussions with individual teachers during their performance review about the importance of using higher-level thinking questions on summative assessments and 22.7% (n=42) reported that principals should let teachers decide how to assess learning. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no

statistically significant results were found when applying the Pearson Chi Square test.

The following conclusions were found for Research Question 2 Part C:

1. Respondents were asked to identify the preparation they received in their teaching certification program regarding the incorporation of higher-level thinking questions in summative assessments. Of the 183 respondents who answered this question, 65.0% (n=119) reported that their teacher certification program included some discussion about this topic while 18.6% (n=34) reported that their teacher certification program did not include discussions regarding higher-level thinking questions. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant results were found when applying the Pearson Chi Square test.
2. Respondents were also asked to identify how often higher-level thinking assessment has been a topic of staff development in the schools in which they taught. Of the 184 respondents who answered this question, 47.3% (n=87) reported that the topic had been discussed between one and three times in the past 2 years. A combined total of 37% (n=68) reported that the topic of higher-level thinking questions had been a topic of their staff development program from four to seven or more times in the past 2 years, and 15.8%

(n=29) reported that the topic was not discussed in the past 2 years. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant results were found when applying the Pearson Chi Square test.

3. Respondents were also asked to identify how effective they believed the staff development programs they had attended were, in helping teachers develop higher-level thinking questions in summative assessments. Of the 186 respondents who answered this question, over half of the respondents 58.1% (n=108) rated staff development programs as average. Nearly 24% (n=44) rated staff development programs on the topic of higher-level thinking questions in assessments as below average or poor; 14.0% (n=26) reported not attending any staff development programs on the topic and, only 4.3% (n=8) of respondents rated staff development on the topic of higher-level thinking questions in assessments as excellent. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant data were found when applying the Pearson Chi Square test.
4. Years in their current position showed a statistically significant relationship when data were analyzed on how effective staff development programs were

in helping develop higher-level thinking questions in summative assessments. A Pearson Chi Square value of 37.862, a $p = 0.002$, n=186 and 16 degrees of freedom was found. This means that response choice(s) and years in current position were likely not independent. The significant Chi Square likely resulted from the fact that a larger percentage of teachers with three or less years' experience (82.1%, n=32) than teachers with 21 or more years of experience (33.3%, n=5) selected the category of average as their response when asked to rate the effectiveness of their staff development programs in helping teachers develop higher-level thinking questions into their summative assessments. No teachers with 3 or less years of experience selected categories of below average or poor (0%, n=0) when compared with teachers with 21 or more years of experience (60%, n=9) in rating the effectiveness of staff development programs in helping develop higher-level thinking questions in summative assessments. This finding indicates that teachers considered as highly experienced more consistently rated their staff development programs as less effective than did teachers with 5 or less years in the teaching field. Experience, however, was not found to have a significant relationship with teacher ability to correctly identify two or more higher-level thinking questions.

5. The subject area taught showed a statistically significant relationship when data were analyzed on the perceived effectiveness of staff development programs on the development of higher-level thinking questions within

summative assessments. A Pearson Chi Square value of 31.482, a $p = 0.049$, n=186, and 20 degrees of freedom was found. This means that response choice(s) and subject area taught were likely not independent. The statistically significant Chi Square value likely resulted from a lower percentage of teachers who selected science (4.3%, n=1) or other (9.1%, n=4) as their primary teaching responsibility rated their staff development programs as below average when compared to teachers who selected language arts (21.2% n=7), math (19.4%, n=6), social studies (19.0%, n=4), or allied-fine arts (20.6%, n=7) as their primary teaching responsibility. Lower percentages of teachers who selected math (3.2%, n=1), allied-fine arts (0%, n=0), and other (2.3%, n=1), as their primary teaching responsibility rated the staff development programs as poor than did those who selected language arts (18.2%, n=6), science (13.0%, n=3), or social studies (19.0%, n=4) as their primary teaching responsibility. Lower percentages of teachers who selected language arts (6.1%, n=2), social studies (4.8%, n=1), or allied-fine arts (5.9%, n=2) as their primary teaching responsibility reported that they had not attended any staff development on incorporating higher-level thinking skills into summative assessments when compared to those who selected math (16.1%, n=5), science (26.1%, n=6), or other (22.7%, n=10) as their primary teaching responsibility.

The following conclusions were found for Research Question Part D:

1. Respondents were asked to rate their own knowledge/expertise related to developing higher-level thinking questions in their summative assessments. Of the190 respondents who answered this question, 66.8% (n=127) rated their knowledge level as about average, 18.9% (n=36) rated their knowledge/expertise as very high, and 14.2% (n=27) reported not being confident in knowledge/expertise regarding developing higher-level thinking questions for summative assessments. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant results were found when applying the Pearson Chi Square test.
2. Subject area taught showed a relationship with how respondents rated their knowledge/expertise in developing higher-level thinking questions for their summative assessments. A Pearson Chi Square value of 34.721, a $p < 0.001$, n=190 and 10 degrees of freedom were found. This means that response choice(s) about knowledge/expertise in incorporating higher-level thinking questions into summative assessments and subject area taught were likely not independent. The significant Chi Square likely resulted from a larger percentage of language arts teachers who reported very high knowledge and expertise (45.7%, n=16) in developing higher-level thinking questions when compared with math teachers (6.5%, n=2), science teachers (13.0%, n=3) or

those who selected other as a teaching category (6.5%, n=3). A lower percentage of language arts teachers reported about average knowledge and expertise (48.6%, n=17) in developing higher-level thinking questions when compared to math teachers (80.6%, n=25) or social studies teachers (76.2%, n=16). A larger percentage of teachers who selected other as their teaching category (28.3%, n=13) reported that they were not confident in their knowledge and expertise when compared to teachers who selected teaching categories of language arts (5.7%, n=2) or social studies (0%, n=0) as the areas of primary teaching responsibility.

The following conclusions were found for Research Question 2 Part E:

1. Respondents were asked to rank, in priority order, levels of responsibility for preparing teachers to develop higher-level thinking questions on summative assessments (1=most responsible to 5=least responsible). The response total (RT) is the weighted sum of the rankings received for each of the possible choices. For example, if the principal was given a "1" ranking 25 times, and a "2" ranking 20 times by respondents, his/her response total would be calculated as follows: 1x25+2x20 = 65. The response count (RC) is the total of the rankings given by respondents for each of the five possible choices, thus the RC for the principal in the above example is 45 (25+20). The response average (RA) would be the average ranking that any one person received (RT divided by RC). Of the 180 teachers who responded to this question, the institutions that prepare teachers were ranked as most

responsible (RA=2.26) and the teachers, themselves, were ranked second as most responsible (RA=2.39). The principal (RA=3.87) was ranked as least responsible for helping teachers develop higher-level thinking questions in summative assessments. When analyzed according to demographics (gender, years teaching, years in current position, subject area taught, highest level of education achieved, location of school, size of school, grade level configuration of school, and AYP status), no statistically significant results were found according to the Pearson Chi Square test.

Research Question 3

The survey asked respondents to rate how important it is to use higher-level thinking questions in summative assessments. The data were analyzed to determine if any statistically significant relationships were found between their answers and their school's Adequate Yearly Progress (AYP) status and, whether or not their school employs someone responsible for oversight of higher-level thinking (e.g., a Higher-level Thinking Coach). This question was developed to determine if there would be any conclusions regarding teacher attitude toward higher-level thinking questions and a standard Minnesota achievement measure such as their school making AYP. This question also was framed to assess whether or not having access to a person who coaches teachers on higher-level thinking questions influences teacher perception of the importance of higher-level thinking questions in assessments.

What relationships exist between teacher ratings of importance of higher-level thinking questions in summative assessments and:

a. Their school making Adequate Yearly Progress.

b. Their district employing a designated person to monitor development of higher-level thinking assessment questions.

The following conclusions were found for Research Question 3 Part A:

1. Respondents were asked if their school was making Adequate Yearly Progress. Of the 205 respondents, 21.0% (n=43) responded yes, 52.2% (n=107) responded no, and 26.8% (n=55) did not know. When analyzed according to the importance teachers assigned to higher-level thinking skills, no statistically significant results were found when applying the Pearson Chi Square test.

The following conclusions were found for Research Question 3 Part B:

1. Respondents were asked if their district employed an individual whose responsibilities included overseeing the use/development of higher-level thinking questions in teacher-constructed assessments. Of the 204 respondents, 42.6% (n=87) responded that they did not know, 31.9% (n=65) responded no one is employed with that responsibility, and 25.5% (n=52) responded yes that a person is employed with that responsibility. When data were analyzed comparing teacher responses to this question and the importance the teachers assigned to higher-level thinking skills, no statistically significant results were found when applying the Pearson Chi Square test.

Recommendations for Professional Practice

Listed below are recommendations for professional practice that educational leaders can use to guide their decision making in the future. These suggestions are based on the research findings discovered in this study.

The following are recommendations for professional practice:

1. Administrators should encourage teachers to incorporate higher-level thinking questions into their summative assessments through the use of staff development opportunities and teacher performance discussions.
2. Administrators should consider evaluating the effectiveness of their staff development programs, particularly on the topic of the development of higher-level thinking questions in summative assessments.
3. Higher education institutions should use this data, or select areas, to improve their teacher preparation programs to assure that prospective teachers are knowledgeable about the value and use of higher-level thinking assessments.

Recommendations for Further Research

The findings of this study generated suggestions for further research. These include the following:

1. A study should be conducted to further investigate the relationship between gender and teacher accuracy in the identification of higher-level thinking questions.
2. A study should be conducted to further investigate if teachers who are highly experienced (21 or more years) consistently rate the effectiveness of their

staff development programs as below average when compared to teachers with 5 or fewer years of experience.

3. A study should be conducted to further investigate why language arts teachers rate their knowledge of developing higher-level thinking skills significantly higher than other subject area teachers.
4. A study should be conducted to further investigate why a school's Adequate Yearly Progress status is not correlated with teachers' ratings of the importance of higher-level thinking questions on summative assessments
5. A study should be conducted to further investigate the relationship between the employment of a Higher-level Thinking Coach and teachers' ratings of the importance of higher-level thinking questions on summative assessments.

Summary

The purpose of this study was to examine teachers' reported knowledge, preparation, use, and importance of higher-level thinking questions in summative assessments. The findings of this study revealed some significant results and, even though not statistically significant, also revealed other information that merits further study. Some of the findings have implications for school leaders, university teacher preparation programs, and K-12 professional development.

One of the most critical items revealed in the study was that less than one in four teachers could correctly identify all three higher-level thinking questions from a sampling of three higher-level and three lower-level thinking questions. Furthermore, a

significant gender difference within these results was also discovered: more male teachers correctly identified all three of the higher-level thinking questions than did female teachers. Another important finding was that the majority of teachers surveyed believed that higher-level thinking questions in summative assessments are very important, but less than one in five reported that they have very high knowledge and expertise in this area. Administrators would be interested to know that in this study, over three in four teachers rate their staff development programs as average to poor in helping them develop higher-level thinking questions on summative assessments. In another important finding, institutions of higher learning are ranked by the teachers as the most responsible for preparing teachers to incorporate higher-level thinking questions into their assessments. This has implications for teacher certification programs.

Research confirms that teacher use of higher-level thinking questions decreases the achievement gap and increases overall student achievement. Therefore, many of the findings of this study are important for administrators, curriculum developers, teachers, and institutions of higher learning to know, in order to provide quality teacher development. An increased capacity of staff members to know about and use higher-level thinking questions in assessments will increase the opportunity for all students to achieve greater success in the classroom and beyond.

REFERENCES

REFERENCES

Allpsychonline. (n.d.). *Research methods*. Retrieved from http://allpsych.com/researchmethods/validityreliability.html

Anderson, L. W., & Krathwohl, D. R. (2001a). *A taxonomy for learning, teaching, and assessing*. New York: Longman.

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001b). *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives* (Complete edition). New York: Longman.

Anderson , L. W., & Sosniak, L. A. (Eds.). (1994). *Bloom's taxonomy: A 40-year retrospective* (93rd Yearbook of the National Society for the Study of Education, Pt. 2). Chicago, IL: University of Chicago Press.

Atkinson, R. C., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational Researcher, 38*(9), 665-676.

Bhattacharya, M. (2002). *Creating a meaningful learning environment using ICT* (CDTL Brief 5). Singapore: National University of Singapore.

Billeh, V. Y. (1974). An analysis of teacher-made science test items in light of the taxonomic objectives of education. *Science Education, 58,* 313-319.

Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Black, T. R. (1980). An analysis of levels of thinking in Nigerian science teachers' examinations. *Journal of Research in Science Teaching, 17*(4), 301-306.

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals.* New York: D. McKay Co., Inc.

Bok, D. C. (2006). *Our underachieving colleges: A candid look at how much students learn and why they should be learning more.* Princeton, NJ: Princeton University Press.

Boston, C. (2002). *The concept of formative assessment.* Retrieved from ERIC database. (ED 0TM0206)

Byrnes, J. P. (2001). *Cognitive development and learning in instructional contexts* (2nd ed.). Boston: Allyn and Bacon.

Cardellichio, T.M., & Field, W. (1997). Seven strategies to enhance neural branching. *Educational Leadership, 54*(6), 33-36.

Caffarella, R. S. (1993). Self-directed learning. *New Directions for Adult and Continuing Education, 57*, 25-35.

Conference Board. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century U.S. workforce*. (Collaboration of The Conference Board, Corporate Voices for Working Families, Partnership for 21st Century Skills, and the Society for Human Resource Management). New York: Conference Board.

Conley, D. T. (2007). *Toward a more comprehensive conception of college readiness.* Eugene, OR: Educational Policy Improvement Center.

Cureton, E. F. (1951). Validity. In E. F. Lindquist (Ed.), *Educational measurement* (1st ed.) (pp. 621-694). Washington, DC: American Council on Education.

Davis, O. L., Jr., & Pfeiffer, I. (1965). *Teacher-made examinations: What kind of thinking do they demand?* Retrieved from ERIC database. (ED015170).

Dickman, A., Schwabe, A., Schmidt, J., & Henken, R. (2009). *Preparing the future workforce: Science, technology, engineering and math (STEM) policy in K-12 education.* Retrieved from EBSCOhost.

Fast, K. V. (1974). An analysis and classification of the ACS-NSTA high school chemistry achievement tests using bloom's taxonomy-cognitive domain. *Science Education, 58*(1), 17-21.

Forehand, M. (2005). Bloom's taxonomy: Original and revised. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology*. Retrieved from http://projects.coe.uga.edu/epltt/

Funk, H. J. (1972). *Test construction behavior of pre-service secondary science teachers.* Retrieved from ERIC database. (ED068333)

Funk, H. J. (1977). Standardized tests for middle school science: What do they measure? *School Science and Mathematics, 77*, 551-557.

Garrison, C., & Ehringhaus, M. (2007). *Formative and summative assessments in the classroom.* Retrieved from http://www.amle.org/Publications/WebExclusive/Assessment/tabid/1120/Default.aspx

Geertsen, H. R. 2003. Rethinking thinking about higher-level thinking. *Teaching Sociology, 31*(1), 1-19.

Gough, D. (1991). *Thinking about thinking.* Alexandria, VA: National Association of Elementary School Principals.

Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning.* Boston, MA: Allyn and Bacon.

K12 Academics. (n.d.). *Education reform.* Retrieved from http://www.k12academics.com/education-reform/history

Khorsand, N. (2009). *Cognitive levels of questions used by Iranian EFL teachers in advanced reading comprehension tests.* Retrieved from ERIC database. (ED507869)

Kocakaya, S., & Gonen, S. (2010). Analysis of Turkish high-school physics-examination questions according to Bloom's taxonomy. *Asia-Pacific Forum on Science Learning and Teaching, 11*(1), 5-18.

Kracht, J. B. (1978). An initial analysis of content and cognitive level of questions appearing on intermediate level tests of geographic knowledge and skills. Retrieved from ERIC database. (ED160511)

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212-218.

Limitation. (n.d.). In Merriam-Webster in *Merriam-Webster*.com. Retrieved http://www.merriam-webster.com/dictionary/limitation

McNeil, L. (2010). Beyond the products of higher-order questioning: How do teacher and English-language learner perceptions influence practice? *TESOL Journal, 2*, 74-90.

Mills, C. W. (1959). The sociological imagination. New York: Oxford University Press.

Moore, B., & Stanley, T. (2010). *Critical thinking and formative assessments: Increasing the rigor in your classroom.* Larchmont, NH: Eye on Education.

Muirhead, B. (2002). *Relevant assessment strategies for online colleges and universities*. Retrieved from http://www.usdla.org/html/journal/FEB02_Issue/article04.html

No Child Left Behind (NCLB) Act of 2001. (2002). Pub. L. No. 107-110, § 115, Stat. 1425.

Nuttall, E. V. (1987). Survey of current practices in the psychological assessment of limited-English-proficiency handicapped children. *Journal of School Psychology, 25*(1), 53-61.

Oliver, D., & Dobele, T. (2007). First year courses in IT: A Bloom rating. *Journal of Information Technology Education, 6*, 347-360.

Ó Murchú, D. (2003). *Mentoring, technology and the 21st century's new perspectives, challenges and possibilities for educators.* Second Global Conference, Virtual Learning & Higher Education, Oxford, UK.

Pancella, J. R. (1971). *Cognitive levels of test items in commercial biology examinations*. Retrieved from ERIC database. (ED052045)

Paul, R., & Elder, L. (1999 revised 2000). *Critical thinking: Basic theory and instructional structures handbook.* Retrieved from http://www.criticalthinking.org/pages.the-role-of-questions-in-teaching-thinking-and-learning

Pogrow, S. (2005). HOTS revisited: A thinking development approach to reducing the learning gap after Grade 3. *Phi Delta Kappan, 87*(1), 64-75.

quotationspage.com. (2013). *Quotation 3073 from Laura Moncur's motivational quotations*. Retrieved from http://www.quotationspage.com/quote/3072.html

Redfield, D. L., & Rousseau, E.W. (1981). A meta-analysis of experimental research on teacher questioning behavior. *Review of Educational Research, 51*, 237-245.

Reeve, C. D., (Ed.), & Grube, G. M. (Trans.). (1992). *Plato Republic* (2nd ed.). Indianapolis, IN: Hackett Publishing Company.

Roberts, K. A. (2008). Beyond content: Deep-structure objectives for sociology and social science curricula. *Scaffolding for Student Success,* 2-23.

Sanders, M. (2001). *Understanding dyslexia and the reading process: A guide for educators and parents.* Boston: Allyn and Bacon.

Shavelson, R. J. (2006). *On the integration of formative assessment in teaching and learning with implications for teacher education.* Paper prepared for the Stanford Education Assessment Laboratory and the University of Hawaii Curriculum Research and Development Group.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., Tomita, M., & Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education, 21*(4), 295-314.

Sousa, D. (2006). *How the brain learns.* Thousand Oakes, CA: Corwin Press.

Sultana, Q. (2001). *Scholarly teaching: Application of Bloom's taxonomy in Kentucky's classrooms.* Retrieved from ERIC database. (ED471982)

Survey methodology. (n.d.). In *Wikipedia*. Retrieved from http://en.wikipedia.org/wiki/Statistical_survey

Taylor, B. M., Pearson, P. D., Clark, K., & Walpole, S. (2000). Effective schools and accomplished teachers: Lessons about primary-grade reading instruction in low-income schools. *The Elementary School Journal, 101*, 121-166. Retrieved from http://www.tesol-journal.com

Travis, R. H. (1996). Developing an assessment program for a university health sciences major. *Health Educator: Journal of Eta Sigma Gamma, 28*(1), 33-38.

Vella, J., Berardinelli, P., & Burrow, J. (1997). *How do they know they know? Evaluating adult learning (Jossey-Bass Higher and Adult Education Series*). New York: Jossey-Bass Inc.

Wiliam, D. (2010). What counts as evidence of educational achievement? The role of constructs in the pursuit of equity in assessment. *Review of Research in Education, 34*(1), 254-284.

APPENDICES

APPENDIX A

Survey Introduction Letter to Administrators

Greetings,

Over the past 2 weeks, I have communicated individually with each of you and would like to take this opportunity to thank you for agreeing to distribute my dissertation survey. I really do appreciate your willingness to help out. It is hoped that results will assist teachers, principals and district administrators decide how, when, and to what extent discussions about higher level thinking assessments should be incorporated into staff development and strategic planning activities.

Below is a link to the survey that you can forward to your staff. The survey is only 20 questions and takes about 10 minutes to complete. The survey is only intended for certified staff. As part of the survey process, I will send you two reminders, one on April 23 and the other on April 30. The survey window is 3 weeks and will be closed on May 2, 2012.

https://www.surveymonkey.com/s/J6WMC9H

If you can, please quickly reply to this email and let me know that you have received the link and that my email is not caught in a spam filter. Once again, thank you for your willingness to help out. Final survey results will beavailable through the Educational Administration and Leadership Doctoral Program website at St. Cloud State University at:

http://bulletin.stcloudstate.edu/gb/programs/EDADDoctoralProgram.asp

The researcher can also be contacted at jjs1212@gmail.com.

Jason Stock

PS: If you choose, you can use the intro below to send out to your staff:

Dear Fellow Educator,

Hello, my name is Jason Stock and I am an educator and a doctoral candidate at Saint Cloud State University in the Educational Administration and Leadership Doctoral Program. Your insights and expertise are very important to my research topic: high school teacher thoughts/beliefs about higher level thinking assessments.

Your school was selected as one of several across Minnesota to take part in this survey based on enrollment size and district location. If you are willing, I would appreciate your help by following the link below to a survey which takes about ten minutes to complete.

https://www.surveymonkey.com/s/J6WMC9H

Your responses are completely anonymous and confidential. Your input will help future teachers and administrators better serve their students. Thank you again for your participation and for all that you do for your school and community.

Jason Stock
SCSU Doctoral Candidate

APPENDIX B

Survey

Minnesota Teacher Survey
Topic: Higher Level Thinking Assessments
Conducted by Doctoral Candidate Jason Stock
Educational Administration and Leadership Doctoral Program
St. Cloud State University
Spring 2012

Thank you for taking the time to participate in a research-based study that examines how Minnesota high school teachers view the importance and use of higher level thinking questions on summative assessments. Your school was selected as one of several across Minnesota to take part in this survey due to enrollment size and district location. Because your time is very valuable, this survey is only 20 multiple-choice questions and will take about 10 minutes to complete. Your responses are completely anonymous and confidential. Your input will help future teachers and administrators better serve their students.

Participation is voluntary. Your decision whether or not to participate will not affect your current or future relations with St. Cloud State University or the researcher. If you decide to fill out the survey and there are any questions you are not comfortable answering, you do not need to answer them. We ask you to please remember this information is confidential and is designed to help educators. If you decide to participate, you are free to withdraw at any time without penalty.

The purpose of this study is to examine teachers' reported:

- use of higher level thinking skills assessment questions on summative assessments
- amount and quality of training/preparation on how to incorporate higher level thinking questions in summative assessments
- rating of the importance of using higher level thinking questions in summative assessments, and
- opinion on the inclusion of higher level thinking assessment questions in teacher evaluation processes.

It is hoped that the results will assist teachers, principals, and district administrators decide how, when, and to what extent discussions about higher level thinking assessments should be incorporated into staff development and strategic planning activities.

Please take a minute to click on the following link to find out detailed information about the survey you are about to complete.

Link to IRB Consent Form

Your completion of the survey indicates that you are at least 18 years of age and you consent to participation in the study. Thank you again for all that you do for your school and community! Please click on the link below to begin your survey.

Begin Survey

Demographic Information:

1. What is your gender?
   - Male
   - Female

2. How many years have you been teaching?
   - 3 or less
   - 4-10
   - 11-15
   - 16-20
   - 21 or more

3. How many years have you been in your current position?
   - 3 or less
   - 4-10
   - 11-15
   - 16-20
   - 21 or more

4. In which subject area do you primarily teach (spend half a day or more)?
   - Language Arts
   - Math
   - Science
   - Social Studies
   - Allied-Fine Arts: Physical Education/Health/FACS/Industrial Tech/Business/Computer Science/Media/Foreign Language
   - Other (Comment Box)

5. What is the highest level of education you have achieved?
   - Bachelors' Degree
   - Bachelors' Degree plus Additional Credits
   - Masters' Degree
   - Masters' Degree plus Additional Credits
   - Doctorate

6. What is the location of your school?
   - Rural
   - Suburban
   - Urban

7. What is the size of you school?
   - 250 or Less
   - 251-500
   - 501-750
   - 751-1000
   - 1001-1250
   - 1251-1500
   - Over 1500

8. What is the grade level configuration of your school?
   - 9-12
   - 10-12
   - 7-12
   - K-12
   - Other

9. Is your school currently making AYP in all areas?
   - Yes
   - No
   - I don't know

10. Does your district employ an individual whose responsibilities include overseeing the use/development of higher level thinking questions in teacher constructed assessments?
   - Yes
   - No
   - I don't know

Survey Questions:

11. How would you rate the importance of incorporating higher level thinking questions into summative assessments?
   - Very important
   - Somewhat important
   - Not really that important
   - Not at all important

12. On average, which percentage best describes how often you incorporate higher level thinking questions into your summative assessments?
   - More than 80%
   - About 50% to 79%
   - About 20% to 49%
   - Less than 20%

13. How would you rate your knowledge/expertise related to developing higher-level thinking questions for your summative assessments using the rating scale below?
   - Very high knowledge and expertise
   - About average knowledge and expertise
   - I am not really confident in my level of knowledge and expertise
   - Comment section

14. Which of the following strategies does your principal/administrator use to convey to teachers the importance of incorporating higher-level thinking questions into assessments? Check all that apply
    - o My principal provides professional development opportunities on this topic,
    - o My principal discusses the importance of using higher level thinking questions with teachers in their performance review,
    - o My principal verbally encourages teachers to use higher level thinking questions in their assessments.
    - o My principal does not discuss or provide direction to teachers about the development and use of higher lever thinking questions in assessments.
    - o Comment section

15. Which statement(s) below describes how you believe principals should encourage teachers to use higher level thinking questions on summative assessments? Check all that apply
    - o Principals need to actively promote and encourage the use of higher level thinking questions on assessments through providing staff development on the topic.
    - o Principals should have discussions with individual teachers during their performance reviews on the importance of using higher level thinking questions on their summative assessments
    - o Principals should encourage departmental discussions on the use of higher level thinking questions in summative assessments
    - o Principals should let teachers decide how to assess learning.

16. Which statement below best represents the preparation you received in your teaching certification program regarding the incorporation of higher level thinking questions on summative assessments?
    - o My teacher certification program fully prepared me to incorporate higher level thinking questions into teacher-developed assessments
    - o My teacher certification program included some discussions that addressed higher level thinking questions.
    - o My teacher certification program did not include discussions regarding higher level thinking questions.

17. How often has higher level thinking assessment been a topic of staff development in the schools in which you have taught?
    - o It has been discussed more than 7 times in the past 2 years.
    - o It has been discussed 4-6 times in the past 2 years.
    - o It has been discussed 1-3 times in the past 2 years.
    - o It has not been discussed in the past 2 years.

18. How would you rate the effectiveness of the staff development programs you have attended on helping teachers develop higher level thinking questions within summative assessments?
    - o Excellent
    - o Average
    - o Below average
    - o Poor
    - o I have not attended any staff development programs on this topic

19. Who should be responsible for providing preparation for teacher's on the development of higher level thinking questions in summative assessments? Please prioritize the following list: 1=most responsible to 5= least responsible. (Use each number only once)
____Institutions that prepare teachers
____Staff development programs offered for teachers at their schools
____The principal of the school
____Teacher coaches or mentors
____Teachers themselves

20. Select three of the following six sample questions that you feel best represent a higher level thinking question. (Answers Shuffled)
- What are the main functions of the human circulatory system?
- In your own words, restate the definition of zero based budgeting.
- How does the story "Huckleberry Finn" relate to your own life?
- Outline the major positive and negative aspects of speed limits.
- Develop a plan of action to increase profits for an organization.
- Justify whether it was right or wrong to use the atomic bomb in WWII.

Once again, thank you for taking the time to complete this survey. Final survey results will be available through the Educational Administration and Leadership Doctoral Program website at St. Cloud State University at:

http://bulletin.stcloudstate.edu/gb/programs/EDADDoctoralProgram.asp

The researcher can also be contacted at jjstock@stcloudstate.edu

---

APPENDIX C

St. Cloud State University Institutional Review Board Letter

St. Cloud State University Institutional Review Board (IRB)

Office of Sponsored Programs Administrative Services 210
Website: stcloudstate.edu/osp Email: osp@stcloudstate.edu Phone: 320-308-4932

Name: Jason Stock
Address: 1639 West Oakes Drive
St. Cloud, MN 56303
Email: jjs1212@gmail.com

Co-Investigator:

IRB APPLICATION DETERMINATION:
EXEMPT

Project Title: An examination of high school teachers' reported knowledge, use, and importance of higher level thinking skills in summative assessments

Advisor: John Eller

The Institutional Review Board has reviewed your application to conduct research involving human subjects. Your project has been: **EXEMPT**

We are pleased to advise you that your project has been deemed as exempt in accordance with federal regulations. The IRB has found that your research project meets the criteria for exempt status and the criteria for protection of human subjects in exempt research. Please note the following items concerning our exempt policy:

-- Principal Investigator assumes the responsibilities for the protection of human subjects in this project

--Exempt protocols DO NOT need to be renewed.

--Exempt protocols DO NOT require revisions. However, if changes are made to a protocol that may no longer meet the exempt criteria, a new initial application will be required.

--Adverse events (research related injuries or other harmful outcomes) must be reported to the IRB as soon as possible.

--The IRB reserves the right to review the research while it is in progress or when it is completed.

Good luck on your research. If we can be of further assistance, please contact the Office of Sponsored Programs at 320-308-4932 or email jlkuznia@stcloudstate.edu. Please use the SCSU IRB number listed on any of the forms submitted which relate to this project, or on any correspondence with the IRB.

For the Institutional Review Board:

Jodi Kuznia
IRB Administrator
Office of Sponsored Programs

For St. Cloud State University:

Dan Gregory
Interim Dean, Graduate Studies

OFFICE USE ONLY

SCSUIRB# 943 - 1172
Type of Review:
Today's Date: 2/16/2012
EXEMPT: 2/16/2012
Expiration Date: