The Effects of Quantity and Quality of Teachers' Probing and Guiding Questions on Student Performance[*]

Alpaslan ŞAHİN [**]

Abstract

This study investigated the types, quantity, and quality of teacher questions and their impact on student understanding. In contrast to previous studies, in order to obtain optimum effects of question types, quantity, and quality, this study controlled for variables such as teachers' experience, textbooks used, and teachers' mathematics preparation knowledge, all of which may affect student achievement. The data were collected from 33 7th- and 8th-grade teachers in 2 different states, Texas and Delaware, who participated in a longitudinal project. A total of 103 videotapes were obtained. For the 1st research question, Hierarchical Linear Modeling (HLM) was run with 2 levels; student and teacher. For the 2nd question, inter-correlations were computed between the variables. We found that the quality teachers' probing questions significantly predicted student performance when other variables were controlled. We also found that the quality and quantity of guiding questions and probing questions significantly correlated.

Keywords: Teachers' Questioning, Student Understanding, Probing, Guiding, and Student Performance.

Öğretmenlerin Sorduğu Sorgulayıcı ve Yönlendirici Soruların Sayı ve Niteliğinin Öğrencilerin Başarısı Üzerindeki Etkisi

Öz

Bu çalışmanın amacı öğrenme-öğertme sürecinde öğretmenlerin sordukları soruların kalitesinin, çeşidinin ve sayısının öğrencilerin anlamaları üzerindeki etkisini incelemektedir. Diğer çalışmalardan farklı olarak, öğretmenin soru çeşidinin, kalitesinin ve sayısının etkisini tam olarak ölçmek için; öğretmenin tecrübesi, kullanılan ders kitapları ve öğretmenin aldığı ders sayılarının öğrencinin öğrenmesi üzerindeki etkisi kontrol edilmiştir. Veriler Texas ve Delaware eyaletlerinde 7. ve 8. sınıflarda matemat-

---

ik öğretmenliği yapan toplam 33 katılımcıdan toplanmıştır. Verilerin toplanmasında izin alınarak video kayıtları kullanılmıştır. Bu kapsamda toplam 103 adet video kaydı yapılmıştır. Verilerin analizinde hiyerarşik doğrusal modelleme (HLM) ve korelasyon katsayıları kullanılmıştır. Çalışma sonucunda, ortam değişkenleri kontrol edildiğinde, öğretmenlerin sordukları sorgulayıcı açık uçlu soruların öğrencilerin başarısını etkilediği bulunmuştur. Ayrıca, öğretmenin kaliteli sorduğu açık uçlu sorularla kaliteli sorduğu yön verici sorular birbiri ile anlamlı şekilde ilişkili bulunmuştur.

Anahtar Kelimeler: Ögretmen, Soru sorma tekniği, Sorgulayıcı soru, Yönlendirici soru, Öğrenci başarısı.

1. INTRODUCTION

Asking and answering questions are very common teaching activities among teachers and students. However, researchers have found little evidence on how the types, quantity, and quality of teacher questioning affect learners (Carlsen, 1991; Chin, 2007; Dillon, 1982; Kawanaka & Stigler, 1999; Samson, Strykowski, Weinstein, & Walberg, 1987). The research does not indicate that one type of question is necessarily superior to the other in terms of student performance (Carlsen, 1991; Chin, 2007; Gall, 1984; Winne, 1979). Mills, Rice, Berliner, and Rousseau (1980) grouped the types of empirical studies done during the 1970s into five groups. From the findings of contemporary research, this grouping seems valid today as well. For instance, one of the groups of studies focused on the effects of types and quality of questions on student achievement, which is the issue addressed in this study. According to Mills et al. (1980), the studies in this group provided important yet sometimes contradictory findings on the issue of questioning in the classroom. The following section discusses the relationships between teachers' types, quality, and quantity of questioning and student achievement.

2. RELATIONSHIPS BETWEEN QUESTIONING AND STUDENT ACHIEVEMENT

Winne (1979) reviewed 18 experimental and quasi-experimental studies on this issue in order to determine which question types helped students learn better. In that review,

Winne defined higher order or divergent questions as questions requiring "the student to mentally manipulate bits of information previously learned to create an answer or to support an answer with logically reasoned evidence" (p. 14). This definition of higher-order questions matched the application, analysis, synthesis, and evaluation levels of Bloom's taxonomy (Bloom, Englaehart, Furst, Hill, & Krathwohl, 1956). Lower-order or convergent questions were defined as those asking for exact recall or recognition of facts previously discussed or read by a teacher. This definition aligned with the levels of knowledge and comprehension in Bloom's taxonomy. Winne used the tallying method and categorized studies into three groups: (a) studies yielding significant results, (b) positive or negative results, or (c) non-significant results. When the studies in each group were compared, there was no difference in student outcomes whether the teacher asked more higher-order or lower-order questions.

In another review, Redfield and Rousseau (1981) used a meta-analysis technique to review almost the same group of experimental and quasi-experimental studies as did Winne (1979) (18 out of 20 were the same as in Winne's study). The studies were categorized into two groups, as skills or training experiments, according to Campbell and Stanley's (1966) criteria for internal validity. Teacher training was defined as the independent variable in training experiments, whereas frequency and teaching skills (i.e., higher- or lower-level cognitive questions) were defined as independent varia-

bles in the skills experiments. In both, student achievement was the dependent variable. The results showed that in some experiments the use of higher-level cognitive questions helped student achievement increase, but in other experiments the use of factual questions was more helpful. However, the majority of experiments favored higher-order questions. In contrast to Winne's review, this meta-analysis demonstrated that teachers' predominant use of higher-level cognitive questions has a positive effect on student scores (.729 moderate effect size). The overall findings supported previous conclusions (e.g., Gall, 1970). In summary, in an environment where teachers are trained in questioning skills and in which the validity of program implementation is carefully monitored, student achievement can be improved when the teachers ask more higher- than lower-level questions (Redfield & Rousseau, 1981).

Samson et. al. (1987) did a similar review, synthesizing 44 empirical studies, including most of the Redfield and Rousseau (1981) and Winne (1979) studies, again to determine whether teacher use of predominantly higher-level cognitive questions has a greater effect on student achievement than the use of lower-level cognitive questions. In this review, the selection criteria were similar to that of the previous two studies. The dependent variable was student achievement and the independent variable was types of questioning (higher- or lower-level cognitive) as defined by Winne (1979). Finally, there was sufficient data to calculate an effect size. Samson et al. (1987) found that the effects of higher cognitive level questioning had a small positive effect on student achievement but concluded that these findings did not have educational importance.

Although the three reviews of previous studies found somewhat different results, it seems clear that the effects of question types, either higher-order or lower-order, on student achievement remain unclear. Carlsen (1991), in a review of questioning, proposed three possible reasons why these studies have inconclusive results on the relationship between teacher types of questions and student achievement. According to Carlsen, the role of questions on student achievement was so weak that finding the effect of teacher questioning was very method dependent. In other words, the teachers' questioning technique was not powerful enough by itself to affect student achievement. Other variables, such as content, context, the types of students, types of textbook, teacher experience, teacher math knowledge, and so on, should be controlled or randomized in studies of questioning.

Several more recent studies have investigated teacher questioning of students and student learning and understanding of mathematics (e.g., Harrop & Swinson, 2003; Ilaria, 2002; Martino & Maher, 1999; Yackel & Cobb, 1996). Some studies continued to investigate the relationship between types of teacher questions and student learning (Chin, 2007; Gall, 1984; Ginsburg, 2009; Kawanaka & Stigler, 1999). A synthesis of experimental and quasi-experimental studies, however, has shown that the relationship between the types and quality of teacher questioning and student achievement has yet to be settled by empirical studies (Dillon, 1982; Rosenshine, 1971; Samson et al., 1987). A few of these recent studies have found some possible support for the relationship between the types and quality of teacher questioning on student outcomes. In summary, attempts to show a relationship between student achievement and certain types or cognitive levels of teacher questions have been inconclusive (Chin, 2007). Although there are some correlational studies showing positive relationships between certain types of teacher

questions and student achievement (Gall, 1970; Rosenshine, 1971), experimental studies have been unable to establish causal linkages (Riley, 1986; Carlsen, 1991;Chin, 2007).

2.1. Textbooks

Data from the Third International Mathematics and Science Study (Schmidt, McKnight, & Raizen, 1997) demonstrated that curriculum materials make a difference in achievement. A body of research also found similar findings, that the type of curriculum materials used in mathematics education affects student learning (e.g., Agodini, Harris, Atkins-Burnett, Heaviside, Novak, & Murphy, 2009; Riordan & Noyce, 2001). However, most curriculum materials or textbooks suffer from a lack of coherence and focus, even though the textbook is the primary guide for implementing the curriculum for most teachers. According to Project 2061, textbook quality should be judged by the following characteristics:

> First, good textbooks can play a central role in improving mathematics education for all students; second, the quality of mathematics textbooks should be judged mainly on their effectiveness in helping students to achieve important mathematics learning goals for which there is a broad national consensus; and, third, an in-depth analysis of much more than a textbook's content coverage would be required to evaluate whether there is potential for students' actually learning the desired subject matter (American Association for the Advancement of Science [AAAS], 2000, p. 1).

Studies have found that standard-based, high-quality textbooks enhance student achievement (Kulm & Capraro, 2004; Reys, Reys, Lappan, Holliday, & Wasman, 2003). Kulm and Capra-

ro reported that despite the variation of enacted curriculum delivered by teachers, student achievement was related to the rankings of the textbooks they used, as rated by the AAAS. The higher-rated textbooks produced higher student achievement. Trafton, Reys, and Wasman (2001) proposed explanations of what this high quality meant by saying that a standard-based curriculum should be comprehensive, coherent, and have depth in developing ideas. Curricula should promote sense-making and engagement of students, and provide motivation for learning. In other words, textbooks influence student learning both directly and indirectly through teachers' providing mathematics content knowledge and applying teaching strategies (Kulm & Capraro, 2004; Reys et al., 2003). In a summary of the goals of a professional development research project, DeBoer et al. (2004) proposed a linear relationship between the following aspects: professional development together with curriculum materials, teacher knowledge, skills and attitude, teaching behavior, and student learning.

A study conducted by Project 2061 examined the quality of 13 textbooks, based on a total of 24 criteria classified into seven categories: identifying a sense of purpose, building on student ideas about mathematics, engaging students in mathematics, developing mathematical ideas, promoting student thinking about mathematics, assessing student progress in mathematics, and enhancing the mathematics learning goal (AAAS, 2000). Four textbooks that were used in this study ranked high, high, medium, and low, respectively: *Connected Mathematics* (Lappan, Fey, Fitzgerald, Friel, & Phillips, 1998), *Math in Context (*Romberg et al., 1998*). Middle Gradesl MathThematics* (Billstein et al., 1999) and *Mathematics: Applications and Connections* (Collins et al., 1999). *Connected Mathematics* and *Math in Context* had a median rating of more than 2.5 on a scale of 0–3 points for all of the 24

instructional criteria for all six benchmarks. *Middle Grades MathThematics* was ranked as partially satisfactory, with scores ranging from 1.3 to 3.0 on the corresponding criteria, while *Mathematics: Applications and Connections* was graded as unsatisfactory, with scores ranging from 0.3 to 2.6 on the corresponding criteria.

2.2. Teaching Experience and Quality

Another factor that may affect student achievement is the length of teaching experience. Even though some research indicates that school inputs make little difference in student learning, a growing body of research shows that schools can make a difference, and an extensive portion of that difference is due to teachers (Darling-Hammond, 2000). Recent studies of teacher effects at the classroom level using the Tennessee Value-Added Assessment System and a similar database in Dallas, Texas, have found that teacher effectiveness is a strong determinant of differences in student learning (Jordan, Mendro, & Weerasinghe, 1997; Sanders & Rivers, 1996; Wright, Horn, & Sanders, 1997). Students who are assigned to several ineffective teachers in a row have significantly lower achievement gains than those who are assigned to several highly effective teachers in a row (Sanders & Rivers, 1996).

Teacher qualities that have been examined for their relationship to student learning include measures of academic ability, years of education, years of teaching experience, measures of subject matter and teaching knowledge, certification status, and teaching behaviors in the classroom (Darling-Hammond, 2000). The results of these studies have been mixed. However, some trends have emerged in recent years. Greenwald, Hedges, and Laine (1996) conducted meta-analyses on school resources that were generally associated with student achievement, including administrator qualifications, class size, teacher education, teaching experience, and teacher salaries. They used the ERIC database to search for the period 1966–1993. The conclusion was that school resources were systematically related to student achievement. In addition, they found that quality of teachers (teacher ability, teacher education, and teacher experience) had very strong, positive relationships to student achievement.

Some studies have found that there is a relationship between the effects of teachers' experience and student learning (Klitgaard & Hall, 1974; Murnane & Phillips, 1981), but not always a significant one or a perfectly linear one. Other studies have established that inexperienced teachers (those with less than three years of experience) were typically less effective than more senior teachers (Rosenholtz, 1986). A possible explanation of this non-linear relationship in experience effects is that veteran teachers have not always continued to learn and may have felt tired in their jobs (Darling-Hammond, 2000). Veteran teachers in settings that emphasized continual learning and collaboration continue to improve their performance (Rosenholtz, 1984). Similarly, very well-prepared beginning teachers can be highly effective. For example, some recent studies of five-year teacher education programs—programs that include a bachelor's degree in the discipline and a master's in education as well as a year-long student teaching placement—have found graduates to be more confident than graduates of four-year programs and as effective as more senior teachers (Andrew & Schwab, 1995; Denton & Peters, 1988, p. 9).

2.3. Teachers' Mathematics Preparation

Spencer (1910) asked "What knowledge is of most worth?" Answers have varied over the years but the professional community may be close to agreement regarding what teachers need to know and be able to do (Strudler,

McKinney, Jones, & Quinn, 1999). Some researchers continue to seek the correct response to Spencer's question, believing that finding the link between teacher education and student achievement is worthwhile (Strudler et al., 1999). Indeed, it has been shown that the teacher's influence is one of the most important variables in student learning (Rivkin, Hanushek, & Kain, 1998; Sanders & Rivers, 1996; Strudler et al., 1999). As Dewey (1939) said, there is only one way to increase student learning or achievement: increase the quality of real teaching.

In a cross-cultural study between the U.S. and China, Ma (1999) found powerful evidence that the mathematical content knowledge of teachers played a vital role in mathematics teaching. According to Ma, teachers' mathematical understanding enabled them to teach mathematical ideas clearly to students. Howe (1998) summarized the problem that existed in mathematics education in the U.S.:

> Education involves two fundamental ingredients: subject matter and students. Teaching is the art of getting the student to learn the subject matter. Doing this successfully requires excellent understanding of both. As simple and obvious as this proposition may seem, it is often forgotten in discussions of mathematics education in the U.S., and one of the two core ingredients is emphasized over the other. In K–12 education, the tendency is to emphasize knowing students over knowing subject matter, while at the university level the emphasis is frequently the opposite. (p. 585)

So, there appears to be a link between teacher's content knowledge and student achievement. Such a link has been asserted in the No Child Left Behind Act [NCLB] (2002) as the primary purpose of teacher education. This act requires that teachers pass a subject matter test. Prior to the enactment of NCLB, the American Mathematical Society proposed that all teachers should have a concrete understanding of the mathematics they plan to teach (Howe, 1998). Researchers (Ball, 1991; Ma, 1999; Monk, 1999; Monk & King, 1994) support the claim that knowledge of subject matter is an important variable in accounting for variance between more effective and less effective teachers. Without content knowledge, it is hardly possible to teach content knowledge to students.

This study investigated the types, quantity, and quality of teacher questions and their impact on student understanding. In contrast to the aforementioned studies, in order to obtain optimum effects of question types, quantity, and quality, this study controlled for variables such as teachers' experience, textbook, and teachers' mathematics preparation knowledge, which may affect student achievement. In particular, we are interested in investigating the effect of the quality and quantity of teachers' probing and guiding questions on students' posttest achievement. Second, we aim to look at the relationships between the quantity and the quality of teachers' probing and guiding questions. Therefore, this study will offer an opportunity to understand the role of questioning on student learning and achievement. Two main research questions were investigated:

1. What is the effect of the quality of questions, quantity of questions, and types of questions on student achievement?

2. What is the relationship between the quality of questions, quantity of questions, and types of questions?

3. METHOD

This study was designed to investigate how the quality, quantity, and types of teachers' questioning influenced student achievement.

The data were collected from 33 seventh- and eighth-grade teachers in two different states, Texas and Delaware, who participated in the project during either the 2002–2003, 2003–2004, or 2004–2005 school years. A total of 103 videotapes were obtained, consisting of one to five lessons for each teacher. The teachers used one of four textbooks: *MathThematics* (Billstein et al., 1999), *Connected Mathematics* (Lappan et al., 1998), *Mathematics: Applications and Connections* (Collins et al., 1999), or *Mathematics in Context* (Romberg et al., 1998).

3.1. Participants

The data analyzed in this study were collected through systematic analysis and coding of videotaped lessons taught by a convenience sample consisting of 33 (7 from Texas and 26 from Delaware) seventh- and eighth-grade teachers. There were 15 public schools (7 from Texas and 8 from Delaware) as part of a five-year longitudinal study. Specifically, since the study focused on the content of algebra, the 33 seventh- and eighth-grade teachers who taught lessons on this content were selected. Table 1 presents a summary of the demographic data of the teachers.

Table 1. Teachers' Demographic Data

Teachers' ID	Textbook	# of Courses Taken	# of Years of Teaching	# of Algebra Lesson	Year
1	CMP	4	0-5	5	2002-03
2	MiC	3	6-10	2	2002-03
3	Glencoe	4	6-10	3	2002-03
4	CMP	3	6-10	4	2002-03
5	MiC	Missing	Missing	3	2002-03
6	MiC	3	0-5	3	2002-03
7	CMP	Missing	Missing	4	2002-03
8	MiC	6	6-10	3	2002-03
9	MTh	12	11-15	2	2002-03
10	MiC	1	6-10	3	2002-03
11	CMP	Missing	Missing	4	2002-03
12	CMP	Missing	Missing	4	2002-03
13	Glencoe	6	0-5	4	2002-03
14	CMP	5	0-5	4	2002-03
15	MiC	6	11-15	3	2002-03
16	CMP	5	11-15	5	2002-03
17	CMP	6	11-15	4	2002-03
18	CMP	Missing	Mssing	3	2002-03
19	Glenceo	16	6-10	3	2003-04
20	CMP	4	6-10	3	2003-04
21	MiC	Missing	Missing	3	2003-04
22	MiC	1	0-5	3	2003-04
23	MiC	6	6-10	2	2003-04
24	MTh	4	6-10	2	2003-04
25	MiC	Missing	Missing	2	2003-04
26	MiC	Missing	Missing	2	2003-04
27	CMP	9	15-20	3	2003-04
28	CMP	Missing	Missing	3	2003-04
29	Glencoe	6	6-10	1	2003-04
30	CMP	Missing	Missing	3	2003-04
31	MiC	8	6-10	3	2003-04

| 32 | MiC | Missing | Missing | 3 | 2003-04 |
| 33 | CMP | 6 | 11-15 | 5 | 2003-04 |

Six teachers were new teachers. Sixteen teachers' information was available on their number of years of experience, varying from 6 years to 20. Eleven teachers' information was missing on their number of years of teaching. Four teachers were male; 29 teachers were female. Some of the teachers had taken several mathematics courses during either their undergraduate or master's programs; others had completed only a few math courses. Ten teachers' information was missing concerning the number of math courses taken.

3.2. Variables

The independent variables were (a) types of teacher questioning, (b) quantity of teacher questioning, and (c) quality of teacher questioning, (d) teacher experience, (e) teachers' mathematics preparation, and (f) textbook used. Student achievement was used as the dependent variable. The types and quality of questions were determined by the indicators used in the IERI project (Nelson, Kulm, & Manon, 2000). Teachers' types of questions were categorized as probing or guiding as defined by criteria developed by the project (AAAS, 2002). Criterion V-A focused on teacher questions that encouraged students to explain their ideas. This criterion reflected teachers' use of probing and follow-up questions to encourage each student to express, clarify, justify, interpret, and represent his or her knowledge and understanding of the learning goals (e.g., with tasks, real-world examples, representations, and/or readings related to the learning goals) and get feedback. In this context, the indicators that defined the quality of questions for Criterion V-A were the following:

Indicator 1: The teaching encourages students to express their knowledge/ understanding relevant to the learning goals.

Indicator 2: The teaching encourages students not only to express but also to clarify, justify, interpret, and/or represent their knowledge/understanding.

Indicator 3: The teaching provides opportunities for each student (rather than just some students) to clarify, justify, interpret, and/or represent their knowledge/ understanding.

Criterion V-B reflected teachers' use of questions that guide interpretation and reasoning of students. Indicators for Criterion V-B were the following:

Indicator 1: The teaching includes specific questions and/or tasks to address a mathematical dilemma that confronts the student(s) and to support student progress toward a more complete conceptual understanding of the learning goals, without leading.

Indicator 2: The guiding questions/tasks are responsive to evidence of student thinking rather than generic in nature and directly target the students' mathematical dilemma regarding the learning goals.

Indicator 3: The teacher is persistent in supporting student progress toward a deeper understanding of the learning goals.

3.3. Lesson Selection

The lessons used in this study were videotaped in 2002–2003, 2003–2004, and 2004–2005. Videotapes of 103 lessons (29 eighth-grade and 74 seventh-grade), one to five lessons for each teacher, were used. Only the first years' lessons for each teacher were selected in order to control for teacher participation in professional

development workshops. The teachers used different textbooks, but the lessons addressed the same algebra content dealing with variables, equality and equations, and change (see Table 2). Three of the textbooks are intended to support teachers in reform-oriented approaches, including the use of student-centered learning and inquiry strategies. These three textbooks, *Connected Mathematics* (Lappan et al., 1998), *Math in Context* (Romberg et al., 1998), and *Middle Grades MathThematics* (Billstein et al., 1999), were rated as high, medium, and low satisfactory respectively in the AAAS (2000) textbook evaluation study. The other textbook, *Mathematics Connections and Applications* (Collins et al., 1999) was a widely used commercial textbook that reflected more traditional instruction. It was rated as unsatisfactory by the AAAS study. As shown in Table 1, two teachers used *MathThematics* (MTh), 14 teachers used *Connected Mathematics* (CMP), 14 teachers used *Mathematics in Context* (MiC), and three teachers used *Mathematics: Connections and Applications* (Glencoe).

Table 2. Learning Goals of the Lessons

Content	Description of Content
Variables	Students were asked to recognize both variables and non-variables in problem situations, to recognize variable expressions as representations of problem situations, and to recognize that variables can be used to represent a generalized rule or principle.
Equality and Equations	Students were asked to demonstrate understanding of the idea that the equals sign indicates equivalence between two expressions. They were asked to find a set of ordered pairs to solve a simple equation, to recognize the representation of a problem situation with a one-variable equation, and to solve simple one-variable equations
Change	The questions used to assess this group of ideas were mostly conceptual in nature. Students were asked to demonstrate understanding about change in a variable over time, as well as how the change in one variable relates to change in another. They were asked to recognize when the relationship between two variables is linear and when the relationship between two variables is represented in the form of an equation.

3.4. Instrumentation

The algebra pretest and posttests were developed by the IERI project researchers. The algebra test was designed to measure the knowledge of algebra concepts and skills regarding variables, change, equality, and equations in middle school students in seventh and eighth grades. The test was specifically designed to evaluate the following benchmark from the *Benchmarks for Science Literacy* (AAAS, 1993): "Symbolic equations can be used to summarize how the quantity of something changes over time or in response to other changes." This benchmark was aligned to the *Principles and Standards for School Mathematics* (NCTM, 2000) middle school objective: "Use symbolic algebra to represent situations and to solve relationships." There were two forms of the test, each consisting of 18 items ranging from multiple-choice questions (7 items), short answer questions (8 items), and a three-part extended-response question (3 items). The coefficient alpha reliability of the test was .81. The summary information for each of the test items is displayed in Table 3.

Table 3. Algebra Field Test Categorization

Item Number	Item Description	Item Type	Content
1	43 = [] – 28	Mc	Equality & Equations
2	Represent trading cards X + 3X = 36	Mc	Equality & Equations
3	Represent Girl Scouts N X 6 = 48	Mc	Equality & Equations
4	Jacob's rule	Mc	Variables
5	Rule in a table	Mc	Change
6	Y=2t	Mc	Change
7	What's true about Y = 2X + 5	Mc	Change
8	Tachi and Bill	Scr	Equality & Equations
9	a =3 and b=5	Scr	Change
10	Small boy raises a flag	Scr	Change
11	Missing number in table	Scr	Change
12	Age of cars	Scr	Change
13	Phone company	Scr	Change
14	Donuts	Scr	Variables
15	19 = 3 + 4X	Scr	Equality & Equations
16A	Garden Patterns A	Scr	---
16B	B	Scr	Change
16C	C	Scr	Equality & Equations Change
16D	D	Ecr	Chang Benchmark

Note: Mc = Multiple choice, Scr = Short answer question, Ecr =Extended response type questions

3.5. Data Coding

Measures of the types, quality, and quantity of teacher questioning were obtained by coding videotapes of the lessons. A computer program was developed by the project to analyze videotapes of teachers' lessons. The lesson was first analyzed to identify the parts of the lesson that addressed one of the intended algebra learning goals. Next, trained analysts identified and time-coded segments of the lesson (sightings) according to their match with one or more of five criteria (types of questioning). For the identified criterion sighting, the analyst rated each indicator as met, or partially met, or not met. A measure of the quality of questioning was obtained by first rating each sighting using the indicators. For each indicator, a rating of 1 (Met), 0.5 (Partially Met), or 0 (Not Met) was assigned. The rating of the criterion

is obtained by adding the indicator ratings; thus a criterion rating can range from 0 to 3. The mean of the ratings for a criterion across all sightings for all lessons was calculated to obtain the measure of the teacher's quality of questioning for each of the two types of questions. The quantity of questions was obtained by finding the percentage of minutes of the class used for criteria V-A and V-B.

To ensure reliability, graduate students and mathematics specialists were trained to do classroom observations. Using videotapes of teaching, protocols were followed for video-analysis to ensure that analysts applied the coding procedure in standard ways (e.g., the training should include at least three people watching a tape together and then sharing their observations) (Gallagher & Parker, 1995; Schoenfeld, 1992). The data were collected at

the different sites where the research was being conducted. Protocols for analysis were developed by the researchers to ensure that the same standards were used at the different sites for collection and analysis. All statistical analyses were correlational in nature. Obtained results were attenuated by the reliability of the data; therefore, reliability scores must be reported (Capraro, Capraro, & Henson, 2001; Thompson, 2003; Vacha-Haase, 1998).

To measure reliability of the video analyses, multiple coders were randomly assigned to code a random sample of videos. Reliability is defined as agreement at the macro level for criterion applied to a similar time frame, using Cohen's kappa to evaluate nonrandom agreement rate between pairs of reviewers for a video. The kappas were averaged to produce a mean reliability estimate across videos and raters.

High rates of inter-rater reliability have been achieved through extensive training of scorers. Each scorer was trained on the open-response items for one of the content area tests and was required to reach and maintain a high error tolerance level. The online nature of the scoring utility allowed close monitoring of the reliability of the scorers. At the completion of each test administration we conducted inter-rater reliability studies on a randomly selected 10 percent of the papers for each content area, and the results were consistently high (90–99%).

3.6. Data Analysis

Quantitative analytic methods were used to analyze the videotape data and student scores. For the first research question, Hierarchical Linear Modeling (HLM) was run with two levels: student and teacher. HLM was applied by creating student- and teacher-level data sets for the model. In level 1, students' pretest scores were used as the predictor and posttest scores were used as the dependent variable. Level 2 investigated the effects of teachers' experience, teacher mathematics preparation, textbook used, quality of probing questioning, quantity of probing questioning, quality of guiding questioning, and quantity of guiding questioning on student performance (posttest scores). Before running HLM, missing data were imputed using NORM (Shafer, 1997) in two steps. First, the mean vector and the covariance matrix were obtained using the EM (expectation-maximization) algorithm. Second, with the obtained estimates, data augmentation was carried out in order to obtain multiply-imputed values to generate three multiply-imputed data sets with 99 iterations. Once multiply-imputed data sets were obtained, HLM was conducted to analyze each separately. Finally, the means of three parameter estimates for each predictor variable were calculated to address both research questions 1 and 2. For the analysis of the model results, the alpha level was set to 0.05 with 1.96 critical values for normal t-statistics, meaning that a t-test value greater than 1.96 is significant.

In order to run HLM for these data sets, all variables should have been interval variables. Since textbooks were a categorical variable, three contrasts were set for four textbooks, comparing higher versus lower rankings. CMP and MiC are the higher ranked and *MathThematics* and Glencoe are the lower ranked according to AAAS (2000) research. The first contrast compared the average of CMP and MiC (higher ranked) with the average of *MathThematics* and Glencoe (lower ranked); the second contrast compared CMP with MiC, and the third contrast compared *MathThematics* with Glencoe.

4. RESULTS

4.1. Research Question 1: What is the effect of the quality of questions, quantity of questions, and types of questions on student achievement?

As illustrated in Table 4, pretest scores predicted posttest scores with a significant average t value of 13.689 (p < .05). The effect of teacher mathematics preparation was significant (t = 1.994, p < .05), which means that the number of math courses teachers took predicted student achievement. The quality of teachers' probing questions significantly predicted student performance (t = 2.276, p < .05), indicating that higher quality probing questions resulted in higher student test performance when other variables were controlled. Contrast C1 had a significant effect (t = 4.295, p < .05), which means that higher-rated textbooks predicted student performance positively as found by Kulm and Capraro (2004).

Table 4. Model Results

	(Est. /S.E)-1	(Est. /S.E)-2	(Est. /S.E)-3	Average
Within Level				
Posttot on Pretot	13.921*	12.939*	14.208*	13.689*
Between Level				
Posttot on				
Tprep	2.112*	2.124*	1.746	1.994*
Texp	-0.824	-1.091	-1.775	-1.230
Faql	2.277*	1.899	2.654*	2.276*
Faqn	-0.816	-0.834	-0.556	-.735
Fbql	0.184	0.862	0.697	.581
C1	4.084*	5.058*	3.743*	4.295*
C2	1.279	1.585	2.134*	1.666
C3	0.478	0.410	0.349	.412

Note. * p < 0.05

Tprep= teachers' mathematics preparation, Texp= number of years of teaching experience, Faql= quality of guiding question, Faqn= quantity of guiding questions, Fbql= quality of probing questions, Fbqn= quantity of probing questions.

4.1.1. Model Fit

Mplus analysis produced the sample correlations and the chi-square test of the model for the sample data. As seen in the results in Table 5, the chi-square test was statistically significant, and the null hypothesis says that if the factor fitting the data is rejected, more factors are required to obtain a non-significant chi-square. Because the chi-square test is sensitive to sample size, large samples often return statistically significant chi-square values and non-normality in the input variables. Mplus also provides the *Root Mean Square Error of Approximation* (*RMSEA*) statistic. The RMSEA is not as sensitive to large sample sizes. According to Hu and Bentler (1999), RMSEA values below .06 indicate satisfactory model fit. Mplus displays the sample statistics for each group separately.

Table 5. Tests of Model Fit for Imputed Data 1, 2, and 3

	1	2	3	Average
P value	0.00	0.00	0.00	0.00
Chi-square	486.29	453.54	490.04	476.62

	RMSEA	0.00	0.00		0.00		0.00	
	Df	10	10		10		10	

From Table 5, it is seen that the obtained chi-square model fit statistic (476.62) and is larger than its degrees of freedom (10). But the RMSEA is well below the cutoff value of .06, leading to the conclusion that the model fits the data very well.

4.2. Research Question 2: What is the relationship between the quality of questions, quantity of questions, and types of questions?

The relationship between the types of teacher questions was investigated by computing inter-correlations between the variables. Since there were no preassumptions about the relationship between quality, types, and quantity of questioning, a two-tailed t-test was accepted for significance level. As shown in Table 6, the quality and quantity of guiding questions and probing questions significantly correlated. The quality of probing questions negatively correlated with the quantity of probing questions, indicating that the higher the quality of probing questions, the shorter the length of class time available for teachers to use for probing questions. Moreover, teachers' quality of probing questions positively correlated to the quality and quantity of guiding questions. So, it can be said that the quality of teachers' probing questions was associated with using higher quality and more class time on guiding questions.

There was a positive, significant relationship between the quantity of probing questions and the quality and quantity of guiding questions. Finally, there was a positive, significant relationship between the quality of guiding questions and the quantity of guiding questions.

Table 6. The Relationships Between Quality, Quantity, and Types of Teacher Questioning: Correlations for Imputed Data 1, 2, and 3.

	Faql	Faqn	Fbql	Fbqn	Faql	Faqn	Fbql	Fbqn
Faql	1	-.059**	.376**	.152**				
	1	-.066**	.365**	.148**	1	-.069**	.365**	.146**
	1	-.082**	.355**	.137**				
Faqn	-.059**	1	.138**	.355**				
	-.066**	1	.137**	.355**	-.069**	1	.138**	.355**
	-.082**	1	.140**	.354**				
Fbql	.376**	.138**	1	.387**				
	.365**	.137**	1	.387**	.365**	.138**	1	.386**
	.355**	.140**	1	.385**				
Fbqn	.152**	.355**	.387**	1				
	.148**	.355**	.387**	1	.146**	.355**	.386**	1
	.137**	.354**	.385**	1				

Note. ** $p < 0.01$ (2-tailed), Faql = quality of guiding question, Faqn = quantity of guiding questions, Fbql = quality of probing questions, Fbqn = quantity of probing questions.

5. DISCUSSION

This study investigated the effects of teachers' quality and quantity of probing and guiding questioning on students' posttest scores by using HLM. Bivariate correlations were then applied to determine if there was a relationship between teachers' quantity, quality, and types of questioning. The results of the study showed that the quality of teachers' probing questions affected student performance when teachers' mathematics preparation, textbook, and teaching experience were controlled. We also found that the quality and quantity of

guiding questions and probing questions significantly correlated.

5.1. Higher Quality Probing Questions Yield Higher Mathematics Learning

The first question revealed that well-prepared and good quality probing questions affected student's learning when other teacher variables including their math preparation, the textbook they use, and their years of teaching experience were controlled. The possible reason could be that teachers with higher-quality probing questions may have provided a richer learning environment for their students because probing questions let students not only express their knowledge and understanding but also justify and interpret their understanding (Sahin & Kulm, 2008). When students had an opportunity to elaborate their thinking and understanding, they may have been more involved with the lesson and contributed to the discussion; thus, they were more likely to understand the topic being discussed (Smith & Stein, 2011). Another possible explanation for this finding could be that since the textbook and teachers' knowledge came out as significant, we may say that the partial effect of probing questions on student performance may have been due to better teacher knowledge and higher-rated textbooks. The findings of this study support the findings of Redfield and Rousseau (1981), who found that teachers' higher-order questions have positive effects on student achievement. In contrast to the previous research, this study considered teachers' experience, teachers' mathematics preparation, and textbooks as independent variables.

DeBoer et al. (2004) provided a model showing a linear relationship between the following aspects: professional development; curriculum materials; teacher knowledge, skills, and attitudes; teaching behavior; and student learning. The present study supports the hypothesis of DeBoer et al., indicating that teachers' mathematics preparation and textbooks positively affected student learning. For example, teachers who don't have a content knowledge problem may be able ask more and different open-ended questions, including probing questions, without the stress of being unable to answer student questions. Also, research found that knowing the subject matter well is a crucial component of good teaching (e.g., Ball, 1991; Ma, 1999; Monk & King, 1994).

The study showed that the particular textbooks used have a significant effect on student achievement. This result confirms the findings of previous research that higher-rated textbooks positively affected students' performance (e.g., Kulm & Capraro, 2004; Reys et al., 2003; Schmidt et al, 1997). This result occurred despite the variation of enacted curriculum delivered by the teachers. This might be so because even though some teachers did not teach from these textbooks, they used them as the main source for giving assignments and quizzes.

5.2. The Relationship between Quality, Quantity, and Types of Questioning

The findings from the second research question focused on how the quality and quantity of guiding and probing questions were related to each other. It was found that teachers' quality of probing questions negatively correlated with the quantity of probing questions. A possible reason for this finding could be that teachers who plan probing questions carefully don't need to ask as many of them to be effective. Another reason could be related to "wait time" (Rowe, 1974): Teachers who ask many open-ended questions are less likely to wait a sufficient time for answers and thus may lower the quality of questions.

Teachers' quality and quantity of probing questions were positively correlated with quality and quantity of guiding questions. When

teachers ask many high-quality probing questions, some students may need guidance and help in order to answer these higher-level questions. This results in incorporating many teacher questions in classes. This finding indicates that teachers who ask higher-quality probing questions should also ask higher-quality guiding questions in order to help students provide answers and become part of the class discussion. This might happen through better training or planning. The last finding from this research question was a positive association between the quality and the quantity of guiding questions. In order to ask better guiding questions, teachers need to spend more time asking guiding questions. Indeed, research has found that guiding questions are usually a set of factual or open-ended questions (Sahin & Kulm, 2008). So asking good quality guiding questions may require a set of or series of guiding questions.

6. IMPLICATIONS AND CONCLUSION

The importance of teacher questioning is immense. The implication from these findings could be that teachers need to spend more time on planning in order ask higher-quality probing questions rather than asking many random open-ended questions. In order to get the maximum effect of good quality probing questions, teachers need to provide enough wait time for students to attend the question and generate a covert and overt answer (Cotton, 1988). This will also provide sufficient time to ask better and longer guiding questions. Another important suggestion or implication of this study would be about the need for teacher training that will prepare teachers as better question askers, because learning how to ask the right good, quality questions does not happen overnight. Attending questioning workshops, watching and observing other teachers, and watching their own teaching are some of the common and working techniques (Sahin, 2013) that teachers can use to develop and improve their questioning skills.

References

Agodini, R., Harris, B., Atkins-Burnett, S., Heaviside, S., Novak, T., & Murphy, R. (2009). *Achievement effects of four early elementary school math curricula.* Retrieved from http://ies.ed.gov/ncee/pubs/20094052/pdf/20094052.pdf

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy.* New York: Oxford.

American Association for the Advancement of Science. (2000). *Middle grades mathematics textbooks: A benchmarks-based evaluation.* Washington, DC: American Association for the Advancement of Science.

American Association for the Advancement of Science. (2002). *Number: Diagram to aid assessment task design*. Project 2061. Unpublished document.

Andrew, M., & Schwab, R. L. (1995). Has reform in teacher education influenced teacher performance? An outcome assessment of graduates of eleven teacher education programs. *Action in Teacher Education, 17*, 43–53.

Ball, D. (1991). Research on teaching mathematics: Making subject matter knowledge part of the equation. In J. Brophy (Ed.), *Advances in research on teaching* (pp. 1–48). Greenwich, CT: JAI Press.

Billstein, R., Williamson, J., Montoya, P., Lowery, J., Williams, D., Buck, M., Winston, B. (1999). *Middle grades MathThematics*. Evanston, IL: McDougal Littell.

Bloom, B., Englaehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *A taxonomy of educational objectives, handbook 1: The cognitive domain*. New York: David McKay.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Capraro, M. M., Capraro, R. M., & Henson, R. K. (2001). Measurement error of scores on the mathematics anxiety rating scale across studies. *Educational and Psychological Measurement, 61*, 373–386.

Carlsen, W. S. (1991). Questioning in classrooms: A sociolinguistic perspective. *Review of Educational Research, 61*(2), 157–178.

Chin, C. (2007). Teacher questioning in science classrooms: Approaches that stimulate productive thinking. *Journal of Research in Science Teaching*, *44*(6), 815–843.

Collins, W., Dritsas, L., Frey-Mason, P., Howard, A. C., McClain, K., Molina, D. D., Wilson, P. (1999). *Mathematics: Applications and connections*. Columbus, OH: Glencoe/McGraw-Hill.

Cotton, K. (1988). *Classroom Questioning*. *Close-Up No. 5*. Portland, OR: Northwest Regional Educational Laboratory.

Darling-Hammond, L. (2000). Authentic assessment of teaching in context. *Teaching and Teacher Education, 16*, 523–545.

DeBoer, G., Morris, K., Roseman, J. E., Wilson, L., Capraro, M. M., Capraro, R.,et al. (2004 April). *Research issues in the improvement of mathematics teaching and learning through professional development.* Paper presented at American Educational Research Association annual meeting, San Diego, CA.

Denton, J. J., & Peters, W. H. (1988). *Program assessment report: Curriculum evaluation of a non-traditional program for certifying teachers*. College Station, TX: Texas A&M University Press.

Dewey, J. (1939). The individual in the new society. In J. R. Editor (Ed.), *Intelligence in the modern world: John Dewey's philosophy*. New York: The Modern Library.

Dillon, J. T. (1982). The effects of questioning in education and other enterprises. *Journal of Curriculum Studies*, *14*, 127–152.

Gall, M. D. (1970). The use of questions in teaching, *Review of Educational Research, 40*, 707–721.

Gall, M. (1984). Synthesis of research on teachers' questioning. *Educational Leadership, 42*(3), 40–47.

Gallagher, J., & J. Parker. (1995). *Secondary science teacher analysis matrix - (SSTAM).* East Lansing, MI: Michigan State University Press.

Ginsburg, H. P. (2009). The challenge of formative assessment in mathematics education: Children's minds, teachers' minds. *Human Development*, *52*(2), 109–128.

Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*, 361–396.

Harrop, A., & Swinson, J. (2003). Teachers' questions in the infant, junior and secondary school, *Educational Studies, 29*(1). 49–57.

Howe, R., (1998). The AMS and mathematics education: Revision of the NCTM standards, *Notices American Mathematics Society, 45*, 243–247.

Ilaria, D. R. (2002). Questions that engage students in mathematical thinking. *Proceedings of the annual meeting (of the) North American Chapter of the International Group for the Psychology of Mathematics Education* (24th, Athens, GA, October 26–29, 2002). Vol. 1–4; SE 066 887.

Jordan, H. R., Mendro, R. L., & Weersinghe, D. (1997). *Teacher effects on longitudinal student achievement: A preliminary report on research on teacher effectiveness*. Paper presented at the National Evaluation Institute, Indianapolis, IN. Kalamazoo, MI: CREATE, Western Michigan University.

Kawanaka, T., & Stigler, J. W. (1999). Teachers' use of questions by eight-grade mathematics classrooms in Germany, Japan, and the United States, *Mathematical Thinking & Learning, 1,* 255–278.

Klitgaard, R. E., & Hall, G. R. (1974). Are there unusually effective schools? *Journal of Human Resources, 10*(3), 90–106.

Kulm, G., & Capraro, R. (2004, April). *Relationship between textbook use and student learning of number and algebra ideas in middle grades.* Paper presented at the research pre-session of the National Council of Teachers of Mathematics Annual Meeting, Philadelphia, PA.

Lappan, G., Fey, J. T., Fitzgerald, W. M., Friel, S. N., & Phillip, E. D. (1998). *Connected mathematics.* Menlo Park, CA: Dale Seymour Publications.

Ma, L. (1999). *Knowing and teaching elementary mathematics.* Mahwah, NJ: Erlbaum.

Martino, A. M., & Maher, C. A. (1999). Teacher questioning to promote justification and generalization in mathematics: What research practice has taught us. *Journal of Mathematical Behavior 18* (1), 53–78.

Mills, S. R., Rice, C. T., Berliner, D. C., & Rosseau, E. W. (1980). The correspondence between teacher questions and student answers in classroom discourse. *Journal of Experimental Education, 48,* 194–204.

Monk, D. H. (1999). Subject area preparation of secondary mathematics teachers and science teachers and student achievement. *Economics of Education Review, 13,* 125–145.

Monk, D. H., & King, J. A. (1994). Multilevel teacher resource effects on pupil performance in secondary mathematics and science. In R. G. Ehrenberg (Ed.), *Choices and consequence* (p. 29–58). Ithaca, NY: ILR Press.

Murnane, R. J., & Phillips, B. R. (1981). Learning by doing, vintage, and selection: Three pieces of the puzzle relating teaching experience and teaching performance. *Economics of Education Review, 11*(4), 691–693.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

Nelson, G. D., Kulm, G., & Manon, J. R. (2000). *Improving mathematics teacher practice and student learning through professional development.* Proposal to the Interagency Educational Research Initiative, American Association for the Advancement of Science, Washington, DC.

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

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

Reys, R., Reys, B., Lapan, R., Holliday, G., & Wasman, D. (2003). Assessing the impact of standard-based middle grades mathematics curriculum materials on student achievement. *Journal for Research in Mathematics Education, 34,* 74–95.

Riley, J. P. (1986).The effects of teachers' wait-time and knowledge comprehension questioning on science achievement. *Journal of Research in Science Teaching, 23,* 335–342.

Riordan, J. E., & Noyce, P. E. (2001). The impacts of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal of Research in Mathematics Education, 32* (4), 368–398.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (1998, April). *Teachers, schools and academic achievement.* Cambridge, MA: National Bureau of Economic Research.

Romberg, T. A., Burrill, G., Fix, M. A., Middleton, J. A., Meyer, M., Pligge, M., et al. (1998). *Mathematics in context.* Chicago: Encyclopedia Britannica.

Rosenholtz, S. J. (1986). The organizational context of teaching. *Learning to Teach*. University of Illinois at Champaign, Urbana: IL

Rosenshine, B. (1971). *Teaching behaviours and student achievement*. London: National Foundation for Educational Achievement.

Rowe, M. B. (1974). Wait-time and rewards as instructional variables, their influence in language, logic, and fate control: Part one-wait time. *Journal of Research in Science Teaching, 11*(2), 81–94.

Sahin, A. (2013). Teachers' Awareness and Acquisition of Questioning Strategies. *Sakarya] University Journal of Education, 3*(3), 17-36.

Sahin, A., & Kulm, G. (2008). Sixth grade mathematics teachers' intentions and use of probing, guiding, and factual questions. *Journal of Mathematics Teacher Education, 11*(3), 221-242.

Samson, G. E., Strykowski, B., Weinstein, T., & Walberg, H. J. (1987).The effects of teacher questioning levels on student achievement. *Journal of Educational Research 80*, 290–295.

Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Schmidt, W. H., McKnight, C., & Raizen, S. (1997). *A splintered vision: An investigation of U.S. science and mathematics education*. Dordrecht, Netherlands: Kluwer.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws, (ed.). *Handbook for research on mathematics teaching and learning*(pp. 334–370).New York: MacMillan.

Shafer, J. (1997). *Analysis of incomplete multivariate data.* London: Chapman and Hall.

Smith, M., & Stein, M. K. (2011). *5 practices for orchestrating productive mathematics discussions*. Reston, VA: NCTM.

Spencer, H. (1910). *Education: Intellectual, moral, and physical*. New York: Appleton.

Strudler, N. B., Mckinney, M. O., Jones, W. P., & Quinn, L. F. (1999). First year teachers' use of technology: Preparations, expectations, and realities. *Journal of Technology and Teacher Education*, 7(2), 115–129

Thompson, B. (2003). *Score reliability: Contemporary thinking on reliability issues.* Thousand Oaks, CA: Sage.

Trafton, P. R., Reys, B. J., & Wasman, D. G. (2001). Standards-based mathematics curriculum materials: A phrase in search of a definition. *Phi Delta Kappan, 83*(3), 259–263.

Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educational and Psychological Measurement, 58,* 6–20.

Winne, P. H. (1979). Experiments relating teachers' use of higher cognitive questions to student achievement. *Review of Educational Research, 49,* 13–50.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 57–67.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*, 458–477.

Genişletilmiş Özet

Soru sorma ve cevaplama; öğretmen ve öğrenci arasında sıklıkla meydana gelmektedir. Ama yapılan araştırmalar henüz hangi soru tipinin, hangi kalitede ve miktarda sorulan soruların öğrencinin başarısına daha çok etkilediğini tam olarak cevaplamamaktadır(Carlsen, 1991; Chin, 2007; Dillon, 1982; Kawanaka & Stigler, 1999; Samson, Strykowski, Weinstein, & Walberg, 1987).

Literatür incelendiğinde çalışmaları çeşitli gruplara ayırmak mümkündür. Mesela, Winne (1979) 18 deneysel ve yarı deneysel çalışmayı incelemiş ve soru tiplerini yeniden tanımlamıştır. Yüksek düşünce gerektiren açık uçlu sorular ve sadece hafızaya dayalı çok derin düşünmeyi gerektirmeyen sorular diye öğretmenin sorduğu soruları ikiye ayırmıştır. Çalışmasında bu tür gruplamalardan sonra öğretmenin sorduğu soru tipinin öğrencilerin başarısı üzerinde anlamlı fark oluşturacak bir etki oluşturmadığını bulmuştur.

Aynı şekilde, Redfield ve Rousseau (1981), benzer bir meta-analiz çalışması yapmış ve bu konuda yapılan 20 çalışmayı analiz etmiştir. Bu çalışmaya göre öğrencilerin yüksek düşünce gerektiren açık uçlu sorularının öğrencilerin öğrenmesinde daha etkili olduğunu bulunmuştur (etki katsayısı: .729).

Son olarak, Samson ve arkadaşları (1987)'de benzer bir calışma yapmış ve daha büyük bir grubu incelemişlerdir (44 deneysel çalışma). Çalışma diğerlerinde olduğu gibi öğrencinin başarısı üzerinde öğretmenin kullandığı soru tiplerinin etkisini belirlemeyi hedeflemiştir. Samson ve arkadaşları da bir önceki çalışma gibi yüksek düşünce gerektiren açık uçlu soruların az da olsa öğrencilerin öğrenmesinde etkili olduğunu bulmuştur. Görüldüğü üzere, soru tipleri ve öğrencilerin başarısı arasındaki ilişki hala gizemini korumaktadır. Aradan yaklaşık kırk yıl geçmesine rağmen hala benzer ve daha kapsamlı bir çalışma yapılmamıştır (Chin, 2007).

Bu çalışmada öğretmenin sorduğu soru tiplerinin (açık uçlu sorgulayıcı ve yönlendirici), soru kalitesinin ve miktarının öğrencinin öğrenme başarısı üzerindeki etkisi araştırılmıştır. En uygun ve yüksek kalitede soruların etkisini ölçebilmek için, öğrencilerin öğrenmesine etkisi olabilecek öğretmenin tecrübesi, kullandığı ders kitabı ve aldığı matematik derslerinin sayısı kontrol edilmiştir.

Çalışmanın araştırma soruları şunladır:

1. Öğretmenin sorduğu soruların türünün, kalitesinin ve miktarının öğrencinin başarısı üzerindeki etkisi nedir?

2. Öğretmenin sorduğu soruların türü, kalitesi ve miktarı arasındaki ilişkiler nelerdir?

Çalışmada, öğretmen videoları ve öğrenci puanlarını analiz etmek için birinci araştırma sorusunun analizinde Hiyerarşik Doğrusal Modelleme (HLM) iki seviyede kullanılmıştır (öğrenci ve öğretmen). İkinci araştırma sorusu için ise değişkenler arası korelasyonlar hesaplanmıştır.

Çalışma sonucunda, diğer değişkenler kontrol edildiğinde, öğretmenlerin sordukları açık uçlu sorgulayıcı soruların öğrencilerin başarısını etkilediği ortaya çıkmıştır. Ayrıca, öğretmenin sorduğu sorgulayıcı sorularla yönlendirici sorular birbiri ile anlamlı şekilde ilişkili bulunmuştur.

Çalışmadan çıkaracağımız uygulamaya dönük sonuçlar ise öncelikle öğretmenlerin soru sorduktan sonra öğrencinin soru üzerinde düşünüp tüm bildiklerini toparlayıp cevap vermesi için yeterli zamanın verilmesi gerekliliğidir. Diğer bir çıkarım ise; madem öğretmenin soracağı açık uçlu sorgulayıcı sorular öğrencinin öğrenmesinde etkili oluyor, o zaman öğretmenleri bu tür soruları nasıl ve ne zaman sorucakları noktasında eğitime tabii tutmak gereklidir. Bu konuda seminerler vermek, öğretmenlerin stajyerlikleri döneminde soru sorma kabiliyetlerini test etmek ve tecrübeli öğretmenleri bu konuda dikkatle gözlemlemek soru sorma kabiliyetlerini geliştirmek için yapılabilecek çalışmalar olabilir.