



Examining question type and the timing of IRE pattern in elementary science classrooms

Sibel Kaya¹
Zeynel Kablan²
Diana Rice³

Abstract

The purpose of this study was to examine the relations among types of teacher questions, student responses, and the timing regarding questioning within science classroom discourse. Thirty one teachers consented to the study and their classrooms were videotaped during a 40-minute science lesson. Classroom discussions that followed an Initiation-Response-Evaluation (IRE) pattern were coded. It was found that about 65% of all teacher questions were short answer questions. Teacher wait-time, student response, and teacher evaluation time were significantly higher in long answer questions compared to short answer questions. Finally, there were positive correlations among these three variables. Detailed analysis results for short and long answer question types were also examined as a part of this study.

Keywords: IRE; questioning; wait time; response time; evaluation time; elementary science education

Introduction

In order to better understand interactions between teachers and students in science classrooms, recent research has focused on classroom discourse, specifically questioning (Chin, 2007; Erdogan & Campbell, 2008; Reinsvold & Cochran, 2012; Scott, Mortimer, & Aguiar, 2006; van Zee, Iwasyk, Kurose, Simpson, & Wild, 2001). Teacher questions are a frequent form of classroom interaction; therefore, they reveal rich information about classroom discourse (Chin, 2007). Research shows that a large majority of questions in a classroom are asked by teachers (96%). In terms of frequency, teachers typically ask 30 to 120 questions per hour in classrooms (Graesser & Person, 1994).

¹ Assistant Professor, Kocaeli University, Faculty of Education, Department of Primary Education, sibelkaya@gmail.com

² Assistant Professor, Kocaeli University, Faculty of Education, Department of Educational Sciences, zeynelkaban@gmail.com

³ Associate Professor, Florida State University, College of Education, Department of Elementary Education, drice@fsu.edu

Traditionally, the purpose of questioning is to evaluate what students know. However, it may also serve different purposes such as eliciting what students think and to help them construct conceptual knowledge (Chin, 2006).

Background

From a Vygotskian perspective, learning takes place through social interaction. According to his sociocultural theory, the learner and the expert negotiate meaning in a scaffolded and supported learning environment. Through the notion of “zone of proximal development” teachers can guide the discourse to support student learning (Vygotsky, 1986). One of the most common patterns of teacher-student interaction in classrooms is Initiation-Response-Evaluation (IRE) (Mehan, 1979). Lemke (1990) calls this three-part exchange as “triadic dialogue”. In *initiation*, the teacher usually asks a question; in *response*, a student (or students) responds to the question, and in *evaluation*, the teacher evaluates the student’s response (Mehan, 1979; van Zee & Minstrell, 1997). Molinari and colleagues (2012) reported that the IRE pattern could take various forms within the same classroom discourse. Moving beyond its basic knowledge transmission format, the pattern may sometimes be utilized to initiate sequences, encourage a variety of perspectives, or stimulate students’ reasoning skills (Molinari, et al., 2012; Nassaji & Wells, 2000). As noted, the first move in an IRE pattern is initiation. In whole-class instruction, it is usually the teacher who initiates the pattern with a question (Nystrand & Gamoran, 1991). The types and ways of questioning by the teacher influence how students construct scientific knowledge (Chin, 2007). Researchers have used different terms when classifying teacher questions. Some researchers grouped them as ‘closed ended’, ‘open ended’, and ‘task oriented’ (Erdogan & Campbell, 2008; Reinsvold & Cochran, 2012); some grouped them as ‘authentic’ and ‘inauthentic’ (Nystrand & Gamoran, 1991); and others described them as ‘known information questions’ and ‘negotiation questions’ (Nassaji & Wells, 2000). Graesser and Person (1994) classified teacher questions as ‘short answer’ and ‘long answer’.

Short answer questions typically require a single word or phrase and do not place much cognitive demand on students. For example, in a verification question such as: “Do liquids have a definite shape?” students are expected to say ‘Yes’, ‘No’, or ‘Maybe’. Long answer questions, on the other hand, usually involve several sentences and shed light on students’ reasoning and misconceptions. Therefore, in order to elicit student talk, teachers are expected to use long answer questions more frequently than short answer questions (Graesser & Person, 1994). Erdogan and Campbell (2008) emphasize that when open ended (long answer) questions are used more frequently in classrooms, students are given more opportunities to construct science knowledge. Asking this type of questions is a way of maintaining students’ interest and engagement in the topic (Nystrand &

Gamoran, 1991). Lemke (1990) points out that if the IRE pattern occurs as asking short answer, recall questions, it limits students' learning. However, through open ended questions that encourage multiple perspectives on a topic, teachers may elicit student contributions in the co-construction of scientific knowledge (Nassaji & Wells, 2000).

The second move in an IRE pattern is student response. Frequent and extended student responses are encouraged for the construction of meaning and understanding (Myhill, 2006). Mortimer and Scott (2003) described an I-R-E-R-E chain, where elaborative teacher feedback is followed by further student response. Through this interactive approach, a teacher is able to more thoroughly explore students' ideas. Teachers may also provide opportunities to respond and engage by allowing multiple responses to questions (Mayer & Patriarca, 2007). Sometimes I-R₁-R₂-R₃-E patterns occur during teacher questioning where multiple students are involved in the process (Scott et al., 2006).

The third move in an IRE pattern is evaluation. Sometimes IRE chains might occur when the teacher asks a new question after completing an evaluative remark (van Zee & Minstrell, 1997). Sinclair and Coulthard (1975) proposed that the evaluation could be a 'follow up' in which the teacher simply accepts or rejects the response, or comments on, exemplifies, expands or justifies it. Evaluation is a very critical part of the triadic dialogue, as it is during this step that teachers replace incorrect information with correct answers (Newman, Griffin, & Cole, 1989) and it leads to the next cycle of teaching and learning. For an effective IRE pattern, the teacher needs to provide students with high cognitive level, open-ended questions that stimulate critical thinking (Nystrand & Gamoran, 1991; Chin, 2007). Another aspect of effective teacher questioning is to provide students with necessary silent time to think and elaborate on their ideas (Tobin, 1987; Wilen, 2004). In the education literature, this silent time is called 'wait time' (Rowe, 1974). There are different types of wait-time described in the literature. We will be using the wait-time defined as "the pause following any teacher utterance and preceding any student utterance" (Tobin, 1987, p.90). Research has shown that when the average wait time was at least 3 seconds or more, student achievement was enhanced (Riley, 1986; Rowe, 1978; Tobin, 1986). Three seconds or more are considered optimal for students to formulate a well-thought out response (Rowe, 1974; Stahl, 1994; Tobin, 1986). Wilen (2004) indicates that students must be given time to comprehend the question, connect the ideas, formulate and express their response. If not given enough time to think, students often get frustrated (Wilen, 2001).

In addition to achievement, extended wait time also contributed to student participation, with an increase in the number of student responses (Honea, 1982; Swift & Gooding, 1983; Tobin, 1986),

in the complexity (DeTure & Miller, 1985; Fagan, Hassler, & Szabo, 1981) and length of student responses (Fagan et al., 1981; Honea, 1982; Rowe, 1974; Swift & Gooding, 1983; Tobin, 1986); and in student participation in problem solving and decision-making (Rowe, 1978). Studies involving the length of student response measured the response length either in the number of words used by students (Honea, 1982; Nassaji & Wells, 2000; Rowe, 1974) or the time students spoke (Baysen & Baysen, 2010; Fagan et al., 1981; Tobin, 1986). In Fagan and colleagues (1981) and Tobin's (1986) studies, increased response time was observed when wait time was extended.

In another line of investigation, wait time was examined in relation to types of teacher questions. Some studies looked at the impact of wait time on teacher question types (DeTure & Miller, 1984; Fagan et al., 1981; Swift & Gooding, 1983). In these studies, the number of high cognitive level questions increased through extended wait time. Others examined how wait time changed via different types of questions (Jones, 1980; Boeck & Hillenmeyer, 1973; Arnold, Atwood, & Rogers, 1974; Matthiesen, 2006). These studies reported that on average, teachers provided more wait time after asking high cognitive level questions.

Other research has examined the relationship between the types of questions and length of student responses. In her experimental study, Brock (1986) found that students' responses to questions that requested information not known by the questioner were longer and more complex than the responses to questions that requested information already known to the questioner. The author suggests that the type of questions plays an important role in enhancing the amount of student output in classrooms. In a more recent study Myhill (2006) reported that in whole class primary/middle school classroom settings, the average length of student utterance was four words. Even though the author had not conducted comparative statistics based on question types, she speculated that the short length of student responses might be due to the high percentage of factual questions used in the classroom.

The Purpose of and Significance of the Study

Type of teacher questioning was the focus of this study. Little, if any research has examined multiple factors related to type of teacher questioning in one study. This study is significant in that it examined several factors centered around teacher questioning within a single context and also considered possible relations among the factors under consideration. Thus, the study had three primary purposes.

One purpose was to determine the frequencies of short and long answer teacher questions used in elementary science classrooms. Previous research reported that teachers tended to use closed

ended, short answer questions, answers to which are already known by the teacher (Graesser & Person, 1994; Myhill, 2006; Reinsvold & Cochran, 2012). For the current study, Graesser and Person's (1994) classification of question types was used. Although the context for their study was a college setting, their scheme was still deemed to be applicable to questioning in the elementary school whole class instructional environment. Teacher questions were thus classified into one of fifteen different categories, five of which were short answer and ten of which were long answer questions.

Another aspect of this study was looking into the wait-response-evaluation time relationship. In previous studies, wait time was examined in relation to response time (Fagan et al., 1981; Tobin, 1986) or these two variables were examined independently (Baysen & Baysen, 2010). However, no studies were found that investigated the three-way relationship among teacher wait time, student response time and teacher evaluation time within the IRE pattern. It would be reasonable to expect to observe relationships among these three variables. That is, more wait time might lead to more student talk; subsequently, where more student talk occurred, more teacher talk would be likely to occur. For the current study in all cases, namely, the length of teacher wait, student response and teacher evaluation time, time was measured in seconds.

The final aspect of this study was to investigate wait, response, and evaluation times in relation to the types of teacher questions. The wait time-question type relationship (Jones, 1980; Boeck & Hillenmeyer, 1973; Arnold, Atwood, & Rogers, 1974) was previously examined in detail, and results of these studies have shown that wait time tended to increase when open ended questions were used. Although no studies have examined the additional time-question type relations, it was expected that students' response times and teachers' evaluation times would also increase when open ended questions were used. Therefore, in the present study, we sought to extend the research in this area by detailing how a number of factors including wait, response, and evaluation times differ based on different types of teacher questions in elementary science classrooms.

Research Questions

The following research questions guided the current investigation:

1. What are the frequencies of short answer and long answer teacher questions in elementary science classrooms?
2. Are there significant correlations among teacher wait time, student response time, and teacher evaluation time?

3. Are there significant differences between short and long answer question types in terms of:

- a) teacher wait time?
- b) student response time?
- c) teacher evaluation time?

Method

The current study relied upon observational methods to obtain data regarding teacher questioning and student responses during fourth grade science classes. Utilizing discourse analysis techniques, types of teacher questions, teacher wait time, student response time and teacher evaluation time were identified and/or quantified. The use of this methodology in educational research has been validated previously on theoretical and pragmatic grounds (Gee & Green, 1998).

Participants

The research was conducted in a Northwestern province of Turkey. The research permission was received from the Ministry of Education, and 20 public schools were randomly selected. After the individual school visits, teachers from seven schools volunteered to participate in the study. Thirty-one 4th grade classrooms from the seven schools were videotaped during a 40-minute science lesson. As seen in Table 1, there were a total of 775 students in the observed classrooms. Of the 31 participating teachers, 13 were male and 18 were female. Years of teaching experience ranged from one year to 30 years. In some schools, all of the 4th grade teachers agreed to participate in the study (i.e., School F), while only two teachers volunteered from School G.

Table 1
Distribution of Classrooms and Students Within Schools in Study

School	A	B	C	D	E	F	G	
Classroom	n	n	n	n	n	n	n	
1	22	32	29	27	28	26	31	
2	26	30	29	26	34	29	31	
3	26	31	30	22	28	26		
4	25	32	30	25	23	27		
5	27					23		
6						29		
7						17		
8						26		
Total n	126	125	118	100	113	131	62	775

n = number of children per classroom

Data Collection

Data were collected during the fall semester of 2012-2013 school year. Video recording dates were previously scheduled with teachers. Therefore, we suppose that teachers might have made special preparation for their lessons. This issue is further discussed in the limitations section.

In order to reduce the anxiety of teacher and students, the camera was introduced into each classroom two weeks before the actual recording. Lessons were recorded by two professionals with wide angle cameras, so that we were able to observe every student and the teacher in each classroom. Video recordings were completed in two weeks. The duration of the videos ranged from 35 to 40 minutes. Whole-class instruction was a common occurrence in classrooms. In terms of content, all teachers taught the *Properties of Solids, Liquids, and Gases* within the *Matter* unit specified by the national curriculum in this two-week period. The national science curriculum in Turkey, which took effect in 2005, focused on a constructivist student-centered instruction. The schools participated in the study use the same Science textbook for 4th grade.

Research Variables

This study examined five research variables based on video recordings in science classrooms: a) type of teacher questions, b) teacher wait time, c) student response time, and d) teacher evaluation time. The first and second variables are categorical and the others are continuous.

As seen in Figure 1, the categorical variable is the type of teacher questions used in elementary science classrooms. These questions were defined as ‘short answer’ and ‘long answer’ as described by Graesser and Person’s (1994). There are five types of short answer and ten types of long answer questions. All questions asked by teachers except for the ones that are not related to content were ultimately coded according to these 15 types (see Table 2 for details).

The continuous variables were: teacher wait time, student response time and teacher evaluation time. The teacher wait time in this study was described as “the pause following any teacher utterance and preceding any student utterance” (Tobin, 1987, p.90). The student response time is described as the duration between when a student starts his/her response and when the response is completed. Finally, the teacher evaluation time is described as the duration of the follow up time after the student response until the teacher accepts, rejects, or evaluates the response (Sinclair & Coulthard, 1975).

As noted, questions unrelated to content were not included in the data analysis for this study. Further, since the study aimed to investigate the IRE pattern in elementary science classrooms,

questions that did not follow the IRE pattern were also excluded. For example, questions that were not answered by students or self-answered by teacher were not included in the analyses.

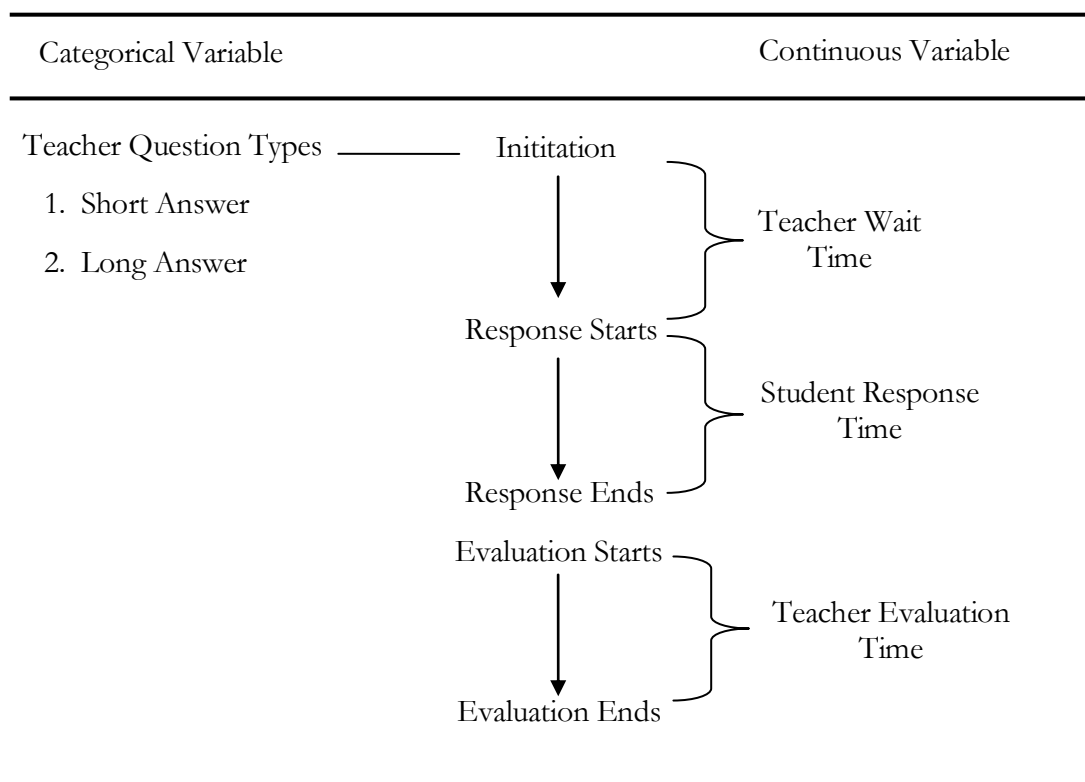


Figure 1. Research variables

Video Coding

Question types. To select the most appropriate scheme for video coding of questions, Graesser and Person's (1994) taxonomy of question types, as well as revised versions by Erdogan and Campbell (2008) and Reinsvold and Cochran (2012) were examined by two researchers. Graesser and Person's taxonomy was found to be most appropriate for the study; therefore, their categories were used (see Table 2). In interpreting Table 2, it should be noted that the question categories and descriptions were taken from the original study by Graesser and Person (1994) and the examples are from the current study. For the purpose of this study the 'instrumental/procedural' category in the original taxonomy was also coded as 'enablement' since these two types were too difficult to distinguish from each other.

Table 2

Question Categories [revised from Graesser and Person (1994)]

Question Categories	Descriptions	Examples from Current Study
1.Short Answer		
S1. Verification	Requires a yes or no response	Do liquids have a definite shape?
S2. Disjunctive	Requires a decision between two options	Is lemonade a pure substance or a mixture?
S3. Concept Completion	Requires completing a definition	When you mix sugar with lemon juice, what is it called?
S4. Feature Specification	Determines qualitative attributes of an object or situation	What are some of the properties of gases?
S5. Quantification	Determines quantitative attributes of an object or situation	How many grams is the empty cup?
2.Long Answer		
L1. Definition	Determines meaning of a concept	What does 'elastic' mean?
L2. Example	Determines examples of a category	Can you give me examples of solids which are also fluid?
L3. Comparison	Compares two different situations	How is sand different from gravel?
L4. Interpretation	Describes what can be inferred from a pattern of data	We can't hold air but we can feel it, how is that possible?
L5. Causal Antecedent	Asks for explanation of what state or event lead to the current state	Why did oil not sink in water?
L6. Causal Consequence	Asks for explanation of the consequence of an event	What happens when you pour melted chocolate into a cup?
L7. Goal Orientation	Asks about motives or goals behind an action	Why did you use filtering for this mixture?
L8. Enablement	Asks for an explanation of the object or resource that allows student to perform an action	How would you separate a mixture of iron filings and sand?
L9. Expectational	Asks about expectations or predictions	Will the magnet attract plastic?
L10. Judgemental	Asks about value placed on an idea or plan	I think I can pour the water on a scale and measure its mass. What do you think?

Graesser and Person (1994) specified a third category for questions and labeled it as 'other'. These are assertion and request/directive type questions. In assertion questions, the teacher indicates that he does not understand an idea and in request/directive questions, asks the student to perform an action. For the purpose of this study, only short answer and long answer questions were coded. Other types of questions were not included in the coding procedure. Each video was coded with a coding template as seen in Figure 2.

Before actual formal coding was implemented, four videos were coded and question types were discussed by researchers for training purposes. Previous research that used Graesser and Person's (1994) taxonomy was examined (i.e., Erdogan & Campbell, 2008; Reinsvold & Cochran, 2012). During this process, a science educator who previously used the Graesser and Person's taxonomy

guided the researchers and provided feedback on some coding issues. After both researchers gained experience in classifying questions, the two researchers shared these 31 videos and coded them independently as the process was quite time-consuming with coding of a single video taking approximately 90 minutes.

Each question type was coded as S1, S3, L4, etc. (see Figure 2). Since, coding of the questions can be subjective at times, inter-rater reliability was determined. To do so, five videos coded by each researcher were randomly selected and cross-coded. Therefore, a total of 10 videos (40 minutes each) were coded by both researchers independently. Total agreement on question types were computed in percentages and as Cohen's Kappa statistic. The coding consistency ranged between 84% and 91% while, Cohen's Kappa values ranged between 0.81 and 0.89 with significance values below 0.001 for 10 videos. Even though these values are considered sufficient for reliability, all 31 coding templates were double checked, without watching the whole 40-minute video, in order to resolve differences and to reach 100% agreement.

Question Number	Q1	Q2	Q3	Q4
Question Type	S3	L4	S1	...
First IRE Pattern				
Time Question Asked*	01:41:8			
Student 1 (Response Starts)	01:45:3			
Student 1 (Response Ends)	01:47:8			
Teacher (Evaluation Starts)	01:50:0			
Teacher (Evaluation Ends)	01:55:5			
Second IRE Pattern				
Time Question Asked	03:10:8			
Student 1 (Response Starts)	03:13:7			
Student 1 (Response Ends)	03:28:7			
Student 2 (Response Starts)	03:37:6			
Student 2 (Response Ends)	03:40:3			
Teacher (Evaluation Starts)	03:41:0			
Teacher (Evaluation Ends)	03:50:6			
Third IRE Pattern				
Time Question Asked	05:34:2			
Choral (Response Starts)	05:35:1			
Choral (Response Ends)	05:36:5			
Teacher (Evaluation Starts)	05:39:0			
Teacher (Evaluation Ends)	05:40:2			

*Time in mm/ss/ff format

Figure 2: Sample Coding Template

Wait, response and evaluation time. For wait time, the time when the teacher finished asking the question and the time when the first student started a response were coded and the difference was computed (see Figure 2). For response time, the time when a student started answering the question and the time when the student finished an answer were coded and, again, the difference was computed. In questions where multiple students responded, the total response time was divided by the number of students who spoke (see Figure 2). If the nature of teacher talk in between students' responses was evaluation, this was noted on the template. For evaluation time, the starting and ending times of the teacher's evaluative response were coded and the difference was computed.

For the accuracy of timing, all coding templates (see Figure 2) were double checked, without watching the whole 40-minute video. The starting and ending times for each IRE cycle were double checked by researchers together.

Limitations

Although this study contains rich observational data, it is limited in some respects. First of all, teachers and students in the classrooms might not have behaved naturally due to the observer effect. The teachers were informed that their classrooms would be evaluated in terms of teacher and student behaviors and interactions. It was observed that most teachers used hands-on activities during their instruction. However, it is not possible to know if teachers actually used these techniques when there was no observer in their classrooms. This is the main limitation in all observation studies (Daymon & Holloway, 2011). During a video recording, participants may be more anxious about the camera. This anxiety might be reduced by fixing the camera in one place rather than moving it around (Hancock, Ockleford, & Windridge, 2009). That procedure was used in present study.

The other limitation of the study was the sampling of each classroom only once. The goal of this study was to reach as many classrooms as possible in order to examine the trends in teacher questioning. This decision was predicated, in part, on the knowledge that types and numbers of questions observed in the classrooms might be content specific. In the current study, only science teaching was observed. Thus, to obtain the broadest range of teachers and greatest variation in questioning in the single subject area, repetition of sampling of a particular set of classrooms was sacrificed for sampling of increased numbers of different classrooms. This brings us to another rather obvious limitation which is that the study examined questioning only in the context of fourth grade science classrooms.

Data Analysis

For data analysis, descriptive statistics, chi-square tests, t-tests and bivariate correlational analysis were conducted using SPSS (18). The frequencies of short answer and long answer teacher questions were reported descriptively. For testing the differences between short and long answer questions in terms of teacher wait time, student response time, and teacher evaluation time, independent samples t-tests were conducted. Within group short answer question types and within group long answer question types could not have been compared in terms of wait time, response time and evaluation time since there were not enough questions from each question type. Descriptive statistics were reported for these variables. In order to determine the relations among teacher wait time, student response time, and teacher evaluation time, bivariate correlational analysis was conducted.

Results

Descriptives for Short Answer and Long Answer Teacher Questions

A total of 872 teacher questions from 31 fourth grade science classrooms were coded during the observed 40-minute science lessons (see Table 3).

Table 3
Frequency and Percentages of Question Types

		f	%
Short Answer	1.Verification	146	16.74
	2. Disjunctive	58	6.65
	3. Concept Completion	174	19.95
	4. Feature Specification	106	12.16
	5. Quantification	86	9.86
	Total	570	65.37
Long Answer	6.Definition	38	4.36
	7. Example	97	11.12
	8. Comparison	17	1.95
	9. Interpretion	34	3.90
	10.Causal Antecedent	21	2.41
	11.Causal Consequence	7	0.80
	12.Goal Orientation	6	0.69
	13.Enablement	58	6.65
	14.Expectational	16	1.83
	15.Judgmental	8	0.92
	Total	302	34.63
	Grand Total	872	100

Of these questions, 570 (65.37%) were short answer and 302 (34.63%) were long answer questions. Most frequently used question type was short answer, concept completion (19.95%), followed by

verification (16.74%) and feature specification (12.16%). Three types of questions, causal consequence, goal orientation, and judgemental were below 1%; comparison and expectational questions were below 2%.

Correlations Among Wait, Response and Evaluation Time

Table 4 displays the correlations among teacher wait time, student response time, and teacher evaluation time. In general, there are positive but weak correlations among the three time variables. In other words, as wait time increases, student response time and teacher evaluation time also increase when short and long answer questions are examined separately or together. The highest correlation was between wait time and evaluation time for all question types ($r = 0.262$, $p < 0.001$) and the lowest correlation was between wait time and response time for long answer questions ($r = 0.125$, $p < 0.05$).

Table 4
Correlations Among Wait Time, Response Time and Evaluation Time for Short and Long Answer Questions

		Response Time Evaluation Time		
All Questions (N=872)	Wait Time	r	0.222**	0.262**
		Sig.	0.000	0.000
	Response Time	r		0.228**
		Sig.		0.000
Short Answer Question (n=570)	Wait Time	r	0.176**	0.190**
		Sig.	0.000	0.000
	Response Time	r		0.201**
		Sig.		0.000
Long Answer Question (n=302)	Wait Time	r	0.125*	0.225**
		Sig.	0.029	0.000
	Response Time	r		0.130*
		Sig.		0.000

* $p < 0.05$, ** $p < 0.001$

Wait, Response and Evaluation Time by Type of Teacher Questions

Table 5 shows that the average wait time that the teachers used in 31 elementary science classrooms for short answer question types was 2.52 seconds. Their wait time for long answer questions was on average 4.56 seconds. The difference between these two wait times was significant according to the independent samples t-test ($t = 8.02$, $p < 0.001$). This result indicates that on average, teachers wait longer until they let students respond after asking a long answer question compared to a short answer question.

The average student response time for short answer question types was 2.72 seconds and the average response time for long answer questions was 6.26 seconds. The difference between these two response times was significant according to the independent samples t-test ($t = 10.94$, $p < 0.001$).

The average teacher evaluation time after a student finished a response for short answer question types was 2.46 seconds and the average teacher evaluation time for long answer questions was 5.74 seconds. The difference between these two evaluation times was significant according to the independent samples t-test ($t = 7.35$, $p < 0.001$).

Table 5

Independent Samples T-test Statistics with Wait Time, Response Time, Evaluation Time for Short and Long Answer Questions

	n	Mean	SD	df	t	p
Wait Time						
Short Answer	570	2.52	2.68	870	8.02	<0.000
Long Answer	302	4.56	4.81			
Response Time						
Short Answer	570	2.72	3.60	870	10.94	<0.000
Long Answer	302	6.26	5.93			
Evaluation Time						
Short Answer	570	2.46	4.75	870	7.35	<0.000
Long Answer	302	5.74	8.40			

Table 6 shows the descriptive statistics calculated for wait time, response time and evaluation time for the 15 different question types. A comparative statistical analysis could not be conducted due to the insufficient numbers from some question types. Among the short answer questions, the longest teacher wait time was for feature specification questions ($M = 3.89$, $SD = 3.89$) and the shortest was for verification questions ($M = 1.88$, $SD = 1.86$). The longest student response time was for feature specification questions ($M = 5.26$, $SD = 5.63$) and the shortest was for disjunctive questions ($M = 1.88$, $SD = 2.29$). The longest teacher evaluation time was for feature specification question ($M = 4.16$, $SD = 6.99$) and the shortest was for disjunctive questions ($M = 1.59$, $SD = 3.50$). Of note is the large variation in all of these time variables.

Table 6

Descriptive Statistics for Wait Time, Response Time and Evaluation Time across Short and Long Answer Questions

Short Answer	n	Wait Time		Response Time		Evaluation Time	
		X	SD	X	SD	X	SD
Verification	146	1.88	1.86	2.16	2.56	2.43	5.85
Disjunctive	58	2.00	2.07	1.88	2.29	1.59	3.50
Concept Completion	174	2.48	2.77	2.03	2.73	1.82	2.56
Feature Specification	106	3.89	3.89	5.26	5.63	4.16	6.99
Quantification	86	2.36	1.44	2.52	2.82	2.31	2.43
Long Answer							
Definition	38	4.61	3.66	5.65	3.96	4.74	4.04
Example	97	3.76	3.58	6.39	5.88	3.51	4.58
Comparison	17	6.00	8.93	5.62	6.29	7.53	9.23
Interpretation	34	5.41	6.21	6.68	6.08	7.59	10.99
Causal Antecedent	21	4.29	4.19	5.85	7.86	4.14	6.38
Causal Consequence	7	4.00	3.65	9.14	12.75	11.43	7.70
Goal Orientation	6	5.33	4.55	4.25	4.98	7.17	10.34
Enablement	58	4.83	5.32	7.14	5.96	7.87	12.97
Expectational	16	5.31	4.99	4.35	3.88	8.12	6.20
Judgmental	8	4.50	1.77	4.81	1.51	3.62	0.52

Among the long answer questions, the longest teacher wait time was for comparison questions ($M= 6.00$, $SD= 8.93$) and the shortest was for example questions ($M= 3.76$, $SD= 3.58$). The longest student response time was for causal consequence questions ($M= 9.14$, $SD= 12.75$) and the shortest was for goal orientation questions ($M= 4.25$, $SD= 4.98$). The longest teacher evaluation time was for causal consequence questions ($M= 11.43$, $SD= 7.70$) while the shortest was for example questions ($M= 3.51$, $SD= 4.58$).

Discussion

The current study served four purposes. One purpose was to determine the frequencies of short and long answer teacher questions used in elementary science classrooms. The results indicated that about 65% of all teacher questions observed in 31 elementary science classrooms were short answer questions. These findings of low frequencies for long answer or open ended questions support and extend those of previous studies that examined questioning across a variety of settings (Graesser & Person, 1994; Matthiesen, 2006; Myhill, 2006; Reinsvold & Cochran, 2012).

In the Graesser and Person study (1994) question types were examined during tutoring sessions at the college level where tutors were graduate students. Similar findings were noted in studies in involving whole class settings. For example, in Literacy and Numeracy Strategies classrooms, (Myhill, 2006) 60% of all teacher questions were factual with an acceptable response known by teacher and in three middle school mathematics classrooms (Matthiesen, 2006) a 78% to 22% ratio

of low-order to high-order questions was found. Even in classrooms where an inquiry focused curriculum was used, short answer or closed-ended questions were more prevalent (Reinsvold & Cochran, 2012). In Myhill's (2006) study, teachers reported that they preferred asking more close ended questions due to the necessity of covering the curriculum objectives, even though they were aware of the value of open ended, higher level, long answer questions.

According to Graesser and Person's (1994) taxonomy of teacher questions used in this study, some types of long answer questions such as antecedent, consequence, goal orientation, expectation, and enablement questions, stimulate students' deep-reasoning skills more than other types. However, in the current study, teachers' use of these types of questions was quite low (see Table 3). Instead, teachers in our study preferred using short answer concept completion and verification questions most frequently. Similar findings were reported in Graesser and Person's (1994) study with college students and Reinsvold and Cochran's (2012) study with third grade elementary students.

Among the long answer questions, teachers in our study used more example and enablement questions compared to other types. Graesser and Person (1994) observed more interpretation questions, while Reinsvold and Cochran (2012) identified interpretation and enablement questions more often. In their study Reinsvold and Cochran used a modified version of Graesser and Person's (1994) taxonomy and did not include example type questions. Therefore, they did not report any results related to this question type.

Another finding of our study was that the teacher wait times, student response times and teacher evaluation times were longer for long answer questions compared to short answer questions. Previous research on wait time identified a 3-second threshold after asking higher cognitive level questions (Boeck & Hillenmayer, 1973; Edwin, 1999; Jones, 1980; Matthiesen, 2006; Swift & Gooding, 1983; Riley, 1986; Tobin, 1986). Jones (1980) reported 2.8 seconds of wait time for convergent questions, whereas the average time for divergent questions was 6.9 seconds in a middle school science classroom. Similarly, in her study with middle school mathematics teachers, Matthiesen (2006) found that the wait time after a high-order question was 4.54 seconds and the wait time after a low-order question was 2.56 seconds. Arnold and colleagues (1974) reported longer wait time for analysis questions (4.6 seconds) compared to questions at other levels of Bloom's taxonomy. These findings point out a 3-second threshold for wait time in high order questions.

We observed a similar trend in our study. A 3-second threshold was also found for average student response and teacher evaluation time for long answer questions. When examined separately, the

wait time, response time and evaluation time for all short answer questions except feature specification were all below 3 seconds. This finding was expected since the verification, disjunction, concept completion, and quantification questions usually require a single word or a yes/no answer. For feature specification questions, wait, response and evaluation times were all over 3 seconds. Although a single feature of an object may require shorter time to utter, teachers might have allowed more think time since there might be several features of an object or situation. For the same reason perhaps, these questions yielded more response time and more evaluation time compared to the other types of short answer questions (see Table 9). When long answer questions were examined separately, making generalizations for causal consequence, goal orientation, and judgemental questions is difficult as the frequencies of these questions are too low. However, when descriptive statistics were examined (see Table 9), causal consequence type questions provided longer response time for students, therefore, these questions might be helpful in eliciting more complex students' ideas.

Another aspect of this study was focused on the wait-response-evaluation time relationship. The positive relationship between wait and response time was already known (i.e., Fagan et al., 1981; Tobin, 1986). Increased wait time might positively influence communication in the classroom as it provides opportunities for students with extended response time; thus, it elicits their ideas. The present study provided further evidence of the three-way positive relationship among wait-response-evaluation times. However, as the other findings of our study indicated, the type of questions used by teachers plays an important role in determining the length of student responses. Therefore, the type of teacher questions might be a better predictor of the length of student response than wait time, as indicated in Fagan and colleagues' (1981) study. In fact, in addition to independent effects, there might be interaction effects of wait time and type of teacher questions on response and evaluation time. Future research might focus on causal relationships among these variables through experimental studies. Another recommendation for researchers might be to examine the quality of student response and teacher evaluation. The current study used the time that student and teacher spoke in analyses. However, longer student response or teacher evaluation might not necessarily indicate high-quality interactions. The quality of classroom interaction might be investigated in relation to wait time and type of teacher questions.

In terms of frequency, teachers in our study asked nearly twice as many short answer questions compared to long answer questions (see Table 3). However, in terms of duration the case is just the opposite (see Table 8). That is to say, teachers spent about twice as much time on long answer questions as short answer questions on average. Therefore, when we considered both the number

of questions and the amount of time spent on the questions for a given type concomitantly and compared the totals for the two types of questions, we found that teachers spent nearly equal amounts of time engaging in the two types during a 40-minute lesson. That is, even though the teachers in our study used more short answer questions, they spent nearly as much time on these questions as long answer questions. Hence, the teachers' concern regarding time constraints about high cognitive level questions (i.e. Myhill, 2006) seems unwarranted. In their experimental study Fagan and colleagues (1981) investigated the effects of wait time and higher level questions on different variables. They concluded that the use of higher level questions reduced the number of total questions. Therefore, rather than focusing on asking low cognitive level questions, a strong suggestion would be that teachers focus on asking high level questions even if it reduces the number of total questions.

As established by this and previous research, IRE reflects a common form of classroom interaction and it has several components to be considered by educators, such as question types, wait time, student response and teacher evaluation. As Haneda (2005) stated, IRE cannot be labeled as 'good' or 'bad'; it is how it is implemented that makes it more or less effective in promoting active student participation. Our study showed that providing more long answer questions as well as extended wait time during an IRE pattern are likely to increase student participation in elementary science classrooms.

Conclusions

While the implementation of this study in Turkey might be viewed by some as a limitation, the opportunity to expand the research in this area in another context adds in substantive ways to the literature, particularly in elementary science education. As a participant in the TIMSS 2011, Turkey is an active participant in efforts to improve science and mathematics education internationally. Of greater significance would be the limitation resulting from the lack of random selection that prevents generalizability of the study results. This issue would be resolved by future experimental studies as described below.

Overall, the current study makes a number of contributions to the study of questioning, particularly in science classrooms. Not only did the results confirm findings of a number of previous studies regarding relations among question types and student responses and wait times, the current study revealed a number of avenues for further research. Among these are investigations to explore the role of the quality of student and teacher responses, not just the time frame for these variables; studies to look at interaction effects involving wait time and type of questions with response and

evaluation times; and research to examine the impact of type of science content on a number of the variables examined in the study. Experimental designs are highly recommended so that causal relations might be examined in all of these cases.

The study strongly confirms a number of implications for classroom teaching. Of particular note is that longer, interpreted here as 'better', questions in general elicit longer wait times and multiple or 'better' answers from students. Findings regarding the relationship between the types of questions, types of answers and the corresponding types of content addressed in the classrooms in the study also suggest that the manner in which content is presented might be examined more closely as we work to provide 'meaningful learning' for all students.

References

- Arnold, D. S., Atwood, R. K., & Roger, V. M. (1974). Question and response levels and lapse time intervals. *Journal of Experimental Education*, 43(1), 11-15.
- Baysen, E., & Baysen, F. (2010). Prospective teachers' wait-times. *Procedia-Social and Behavioral Sciences*, 2(2), 5172-5176.
- Boeck, M. A., & Hillenmeyer, G. P. (1973). *Classroom interaction patterns during microteaching: Wait time as an instructional variable*. Paper presented at the Annual Meeting of the American Research Association, New Orleans.
- Brock, C. (1986). The effects of referential questions on ESL classroom discourse. *TESOL Quarterly*, 20(1), 47-59.
- Chin, C. (2007). Teacher questioning in science classrooms: Approaches that stimulate productive thinking. *Journal of Research in Science Teaching*, 44(6), 815-843.
- DeTure, L. R., & Miller, A. P. (1985). *The effects of a written protocol model on teacher acquisition of extended wait-time*. Paper presented at the annual meeting of the National Science Teachers Association, Cincinnati, OH.
- Daymon, C., & Holloway, I. (2011). *Qualitative research methods in public relations and marketing communications* (2nd ed.). London: Routledge.
- Edwin, R. G. (1999). Developing novice teachers' oral-questioning skills. *McGill Journal of Education*, 34, 29-47.
- Erdogan, I., & Campbell, T. (2008). Teacher questioning and interaction patterns in classrooms facilitated with differing levels of constructivist teaching practices. *International Journal of Science Education*, 30(14), 1891-1914.
- Fagan, E. R., Hassler, D. M., & Szabo, M. (1981). Evaluation of questioning strategies in language arts instruction. *Research in the Teaching of English*, 15, 267-273.
- Gee, J. P., & Green, J. L. (1998). Discourse analysis, learning and social practice: A methodological study. *Review of Research in Education*, 23, 119-169.
- Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, 31(1), 104-137.

- Kaya, S., Kablan, Z., & Rice, D. (2014). Examining question type and the timing of IRE pattern in elementary science classrooms. *International Journal of Human Sciences*, 11(1), 621-641. doi: [10.14687/ijhs.v11i1.2730](https://doi.org/10.14687/ijhs.v11i1.2730)
- Hancock, B., Ockleford, E., & Windridge, K. (2009). *An introduction to qualitative research*. United Kingdom: The NIHR Research Design Service for the East Midlands/Yorkshire & the Humber
- Haneda, M. (2005). Some functions of triadic dialogue in the classroom: Examples from L2 research. *Canadian Modern Language Review*, 62(2), 313-333.
- Honea, M. J. (1982). Wait time as an instructional variable: An influence on teacher and student. *Clearinghouse*, 56(4), 167-170.
- Jones, N. A. (1980). The effect of type and complexity of teacher questions. Doctoral Dissertation, University of Pittsburgh. *Dissertation Abstract International*, 41(2), 529-A.
- Lemke, J. L. (1990). *Talking science. Language, learning and values*. Norwood, NJ: Ablex.
- Liu, Y. B. (2008). Teacher-student talk in Singapore Chinese language classroom: A case study of IRF. *Asia-Pacific Journal of Education*, 28(1), 87-102.
- Matthiesen, E. A. (2006). *Teacher questioning: Effect on student communication in middle school algebra mathematics classrooms*. (Master's thesis). Texas A&M University.
- Mayer, M. J., & Patriarca, L. A. (2007). Behavioral scripts and instructional procedures for students with learning and behavioral problems. *Preventing School Failure*, 52(1), 3-12.
- Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.
- Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction*, 6(4), 359-377.
- Molinari, L., Mamei, C., & Gnisci, A. (2012). A sequential analysis of classroom discourse in Italian primary schools: The many faces of the IRF pattern. *British Journal of Educational Psychology*, doi:10.1111/j.2044-8279.2012.02071.x
- Mortimer, E. F., & Scott, P. H. (2003). *Meaning making in secondary science classrooms*. Maidenhead: Open University Press.
- Myhill, D. (2006). Talk, talk, talk: Teaching and learning in whole class discourse. *Research Papers in Education*, 21, 19-41.
- Nassaji, H., & Wells, G. (2000). What's the use of triadic dialogue? An investigation of teacher-student interaction. *Applied Linguistics*, 21(3), 376-406.
- Newman, D., Griffin, P., & Cole, M. (1989). *The construction zone: Working for cognitive change in school*. Cambridge, MA: Cambridge University Press.
- Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research in the Teaching of English*, 25(3), 261-290.
- Reinsvold, L.A., & Cochran, K. F. (2012). Power dynamics and questioning in elementary science classrooms. *Journal of Science Teacher Education*, 23, 745-768.
- Riley, J. P. (1986). The effects of teachers' wait-time and knowledge comprehension questioning on science achievement. *Journal of Research in Science Teaching*, 23(4), 335-342.
- Rowe, M. B. (1974). Wait time and rewards as instructional variables, their influence in language, logic, and fate control: Part 1. Wait time. *Journal of Research in Science Teaching*, 11(2), 81-94.

Rowe, M.B. (1978). Wait, wait, wait. . .*School Science and Mathematics*, 78(3), 207-216.

Scott, P. H., Mortimer, E. F., & Aguiar, O. G. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons. *Science Education*, 90(4), 605-631.

Sinclair, J., & Coulthard, R. M. (1975). *Toward an analysis of discourse. The English used by teachers and pupils*. Oxford: Oxford University Press.

Stahl, R. J. (1994). *Using think-time and wait-time skillfully in the classroom*. ERIC Digest. ERIC Clearinghouse for Social Studies/Social Science Education. Bloomington, IN. ED370885.

Swift, J. N., & Gooding, C. T. (1983). Interaction of wait time feedback and questioning instruction on middle school science teaching. *Journal of Research in Science Teaching*, 20(8), 721-730.

Tobin, K. (1986). Effects of teacher wait time on discourse characteristics in mathematics and language arts classes. *American Educational Research Journal*, 23(2), 191-200.

Tobin, K. (1987). The role of wait time in higher cognitive level learning. *Review of Educational Research*, 5(1), 69-95.

van Zee, E. H., Iwasyk, M., Kurose, A., Simpson, D., & Wild, J. (2001). Student and teacher questioning during conversations about science. *Journal of Research in Science Teaching*, 38(2), 159-190.

van Zee, E.H., & Minstrell, J. (1997). Using questioning to guide student thinking. *The Journal of the Learning Sciences*, 6, 229-271.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press

Wilén, W. (2001). Exploring myths about teacher questioning in the social studies classroom. *Social Studies*, 92(1), 26-32.

Wilén, W. (2004). Refuting misconceptions about classroom discussion. *The Social Studies*, 95(1), 33-40.