JOURNAL OF RESEARCH IN SCIENCE TEACHING VOL. 37, NO. 6, PP. 602–614 (2000)

An Analysis of Question Asking on Scientific Texts Explaining Natural Phenomena

Jorge Costa,[1] Helena Caldeira,[2] Juan R. Gallástegui,[3] and José Otero[4]

[1]*Escola Secundária de Moimenta da Beira, 3620 Moimenta da Beira, Portugal*

[2]*Department of Physics, Universidade de Coimbra, 3000 Coimbra, Portugal*

[3]*IESP Negreira and Department of Science Education, Universidad de Santiago, 15704 Santiago de Compostela, Spain*

[4]*Department of Physics, Universidad de Alcalá, 28871 Alcalá de Henares, Madrid, Spain*

Received 16 December 1998; accepted 19 February 2000

Abstract: This article relays results of a study focused on questions invoked to correct declarative knowledge deficits while readers process science texts explaining natural phenomena. Firstly, the authors focused on finding out what kind of questions are asked by students who read these texts and, secondly how task demand influences quantity and quality of formulated questions. Two hundred and eighty nine Portuguese students from 8th, 10th, and 12th grade participated in the study. The students were instructed to ask questions on two short science paragraphs that explained natural phenomena. Three task conditions were chosen. Thus, in the ''Class'' condition, the task was introduced as an activity aimed at developing the capacity to ask questions. In the ''Examination'' condition, the task was presented as a test on question generation. Finally, in the ''Extra-academic'' condition the questioning task was camouflaged as a participation in a research project sponsored by the Ministry of Education and geared at the improvement of science textbooks. The results have shown that students are able to ask many questions when given an opportunity to do so. The study has also proven that students are capable of generating a large volume of causal-antecedent questions relative to this kind of texts. Finally, no clear effects were found between grade level and/or task demand as defined in the conducted study.© 2000 John Wiley & Sons, Inc. J Res Sci Teach 37: 602–614, 2000

Correspondence to: H. Caldeira or J. Otero
Contract grant sponsor: Acciones Integradas Hispano-Portuguesas (HC and JO)
Contract grant number: HP 1995-0039 and HP 1997-008
Contract grant sponsor: DGICYT, Ministry of Education, Spain (JO)
Contract grant number: PB93-0478

It is well-known that active self-regulation strategies such as elaboration, critical thinking, and organization of information influence learning (Weinstein & Mayer, 1985; Paris, Lipson, & Wixon, 1983). Monitoring one's own comprehension also forms a part of the strategies that play a key role in self-regulated learning (Corno, 1986; Zimmermann, 1990). Two phases may be distinguished in comprehension monitoring: evaluation and regulation of comprehension (Baker, 1985; Zabrucky & Ratner, 1986, 1989, 1992; Otero, 1996). Evaluation refers to the identification of a comprehension problem. Primary school students, for instance, frequently overlook apparently clear inconsistencies that exist between information found in a text and their knowledge (Markman, 1979). This may indicate an evaluation problem. Regulation consists of the actions undertaken to solve the problem, once it has been detected. Students are capable of detecting a comprehension problem but apparently fail to take appropriate action, when for example, they consciously ignore the difficulty or ''repair'' it inappropriately through unwarranted inferences (Baker, 1979; Otero, 1996). One of the appropriate regulation actions consists of asking questions about problematic information to be processed.

Questions may be formulated with several purposes in mind such as supervising common ground, socially coordinating action, or monitoring conversation and attention (Graesser, Person, & Huber, 1992). This study was not designed to focus on these kinds of questions but rather on those whose aim is the correction of declarative knowledge deficits—a topic specified by Graesser, Person, and Huber (1992) as one of the mechanisms responsible for the generation of questions.

Question asking is known to have positive effects on comprehension. Rosenshine, Meister, and Chapman (1996) reviewed 26 studies about the effects of questioning on comprehension, and were able to determine a 0.36 global effect size when standardized comprehension texts were used to measure the dependent variable. The effect size increased to 0.86 when teacher-made texts were used to measure comprehension. The improvement brought about by questioning was patent at all grade levels, from the third grade until college.

However, asking questions may not be an easy task for all students. Generating a knowledge deficit question is a process comprising three distinct stages: anomaly detection, question articulation, and social editing (Graesser, Person, & Huber, 1992). There are influential variables that operate on any of these three stages, which may prevent the generation of knowledge deficit questions. First, there are influences arising from cognitive and metacognitive variables. Shallow information processing, for example, may limit anomaly detection—the first stage in the generation of knowledge deficit questions. As indicated above, many results in the comprehension monitoring literature show the difficulties that students encounter in detecting apparently evident anomalies such as textual contradictions or inconsistencies with their own knowledge (Markman, 1979; Baker, 1979; Otero & Campanario, 1990). For instance, Otero and Campanario (1990) gave 2nd BUP students in Spain (equivalent to 10th grade level) six short paragraphs, 70–90 words each, for comprehensibility evaluation. Four of these paragraphs contained explicit contradictions inserted into the second and the sixth sentences. An example of one of these contradictions was the following: ''The interatomic bindings of ceramics are strong and rigid . . . Ceramics are fragile because their interatomic bindings are weak and not rigid.'' Subsequently, they found that 34% of the students did not detect *any* of the contradictions in the four paragraphs. A failure to detect such difficulties would surely prevent asking any question geared at solving the comprehension problem. Similar monitoring problems appear in studies on ''calibration of comprehension'', i.e. the assessment of one's own state of comprehension. Glenberg and Epstein (1985) found that subjects' performance in verifying inferences derived from a given text was very poor as compared with their confidence in relying on that text in order to draw correct inferences.

Second, there are personal variables such as achievement, motivation, and self-esteem that influence question asking, especially at the social editing stage. For example, Good, Slavings, Harel, and Emerson (1987) found that elementary grade pupils considered as average achievers ask more questions than low and high achievers. High self-esteem students seem to be very sensitive to the cost of asking questions, although interactions with other variables occasionally produce zero effect of this variable on question asking (Van der Meij, 1994).

Third, contextual variables and those related to social constraints may prevent students from asking questions as well. Thus, Karabenick (1996) examined social influences on learners' judgement of comprehension, a crucial prerequisite for question asking. He found that students' judgements of comprehension depended on the number of questions made by colearners. In other words, other learners' signals of confusion create the condition for the expression of one's own incomprehension. Likewise, García-Arista, Campanario, and Otero (1996) found that anomaly detection, which is the first stage in the creation of a question, is dependent on context: secondary school students detected more inconsistencies when relying on texts in a science class setting than when they used the same texts in a language class setting. Cuerva and Otero (1996) found additional evidence about the effect of contextual variables on comprehension monitoring: secondary school students detected more inconsistencies in scientific information when they thought that it was provided by an authoritative source (college level textbook), than when they thought it came from a source of lesser epistemic authority, such as a primary school level textbook.

In addition to these deficiencies at any of the mentioned question generation stages, previous studies have shown that evaluation in some educational systems, (for example, the Spanish and its Portuguese counterpart,) stresses only vaguely the importance of comprehension monitoring abilities (Otero, Campanario, & Hopkins, 1992; Campanario, García-Arista, Otero, Patricio, Costa, Prata Pina, Caldeira, & Thomaz, 1994). Correlation between academic achievement as measured by regularly administrated teacher-made tests, and measurements of comprehension monitoring ability were found to be positive and significant but low to moderate in magnitude. This may be interpreted as the result of the low weight of comprehension monitoring abilities in teachers' evaluation instruments. Comprehension monitoring ability includes asking questions when the subject cannot solve a problem. Consequently, some evaluation procedures may discourage students from actively regulating their comprehension by means of asking questions, due to the limited importance that comprehension monitoring is given when it comes to academic success.

As a consequence of these influences, it is not surprising to discover that students ask fewer questions in normal class environments. Graesser and Person (1994) report that estimated frequencies of students' questions per hour of class, from kindergarten to 12th grade, range from 1.3 to 4.0 with a median of 3.0. This corresponds to a mean of .11 questions per hour and per pupil for the average of 26.7 students per class reported by Dillon (1988), although it improves up to 0.7 in classes designed for discussion at the secondary level. In conclusion, these studies show that students ask fewer questions in the classroom environment and, in addition, the frequently asked questions have a low cognitive level (Dillon, 1988; Pedrosa de Jesus & Maskill, 1990).

However, research has also shown that students are able to ask questions on a scientific subject matter, provided that certain conditions are met. Examples of these conditions are those existing in tutoring settings (Graesser & Person, 1994) or those provided by computer-aided teaching systems such as ''Point and Query'' (Graesser, Langston, & Bagget, 1993). Point and Query provides an interface where students learn by pointing to words or pictures on the computer screen, then by asking questions from a menu that are relevant to the selected element, and finally by interpreting the supplied answers.

By identifying variables that facilitate question asking under these conditions, we may learn how to improve question asking in other settings. That would be the case of the study conducted by Graesser, Langston, and Bagget (1993), using Point and Query, where learners' goals was a variable that had a great impact on question asking. In that study, college students were told to acquire knowledge about woodwind instruments from the ''Point and Query'' system. The students' goals were manipulated by varying the specific tasks. In a ''Design Instrument'' condition, the students used Point and Query to learn how to design a new woodwind instrument that had a specific tone. However, in an ''Assemble Band'' condition, the task consisted of assembling a six-player woodwind band to play in a New Year's Eve party. The results showed that there was a significant and important increase in the frequency of causal questions asked in the Design Instrument task as compared with the Assemble Band task. Causal knowledge was sampled if the students' task was to solve a casual problem; otherwise students' asked relatively superficial questions. Thus, readers' goals, influenced by task demands, appears as an interesting variable to be explored in studies on question asking in educational settings.

Students' grade level may also influence question asking. For that matter, Otero, Campanario, and Hopkins (1992) found that the comprehension monitoring ability of Spanish secondary school students gradually increased from 9th grade on to the 12th grade. Consequently, one would expect a corresponding increase in the quantity and/or quality of the questions asked about problematic information.

This article summarizes results of research conducted on the number and kind of questions asked by primary and secondary level Portuguese students after reading science paragraphs that described natural phenomena. The effect of task and grade level, two of the variables that are likely to influence question asking, were examined. We asked the following questions:

1. What kind of questions are asked by students of different grade levels who read science paragraphs dealing with natural phenomena?
2. How do type of task and grade level influence the number and quality of questions formulated?

Two science paragraphs were presented to several classes, instructing the students to read them carefully and to ask in writing about everything that they did not understand. Different task conditions were created by providing different information on the purpose of the activity. In one scenario, the question asking activity was simply presented as a class exercise (''Class'' condition). In another one, the same activity was introduced, but students were informed that they would be graded depending on the quantity and quality of questions asked (''Examination'' condition). It is well-known that students' perception of evaluation demands influences their learning strategies (Alexander, Jetton, Kulikovich, & Woehler, 1994; Entwistle & Ramsden, 1983). Crooks (1988, p. 443) singles out as one of the effects of evaluation ''communicating and reinforcing . . . the instructor's or the curriculum's broad goals for students.'' Thus, presenting the task as an activity that would be graded was expected to influence students' goals. We expected that generating questions might be regarded as more important in this condition than in the Class condition and the quantity and/or quality of questions asked might improve, too. Finally, in a third arrangement, the task was camouflaged as participation in a research project on the improvement of science textbooks (''Extra-academic'' condition). Students were treated as ''consultants'' on the comprehension difficulties of the submitted paragraphs. This condition, without a direct relation to the students' instructional activities, was supposed to provide a baseline for comparison with the other two arrangements, that were more academic situations.

Table 1
Taxonomy of inquiries developed by Graesser, Person and Huber (1992; Graesser & Person,1994). DRQs categories are indicated in parentheses. The examples correspond to questions asked by students in this study. In the cases of void categories the examples (identified by an asterisk) are taken from Graesser, Person, & Huber (1992)

Question Category	Example
1. Verification	Are droplets drops?
2. Disjunctive	Are clouds made of water vapor or are they made of droplets?
3. Concept completion	In which liquids does nitrogen dissolve? Which are the other things that allow fish to breathe in water?
4. Feature specification	What kind of polluting process? The water contained in a cloud, if it is not in the liquid state . . . in what state is it?
5. Quantification	What is the average speed of a drop of water? What is, more or less, the quantity of oxygen the fish need in order to survive in water?
6. Definition	What is air friction? What is solubility?
7. Example	Could you give examples of the type of pollution the text refers to?
8. Comparison	What is the difference between water vapor and water droplets? What are other differences between oxygen and dissolved oxygen?
9. Interpretation	*Does that graph show a main effect for A?
10. Causal antecedent (DRQ)	Why does it rain sometimes more often than other? Why are these gases soluble in water?
11. Causal consequence (DRQ)	And if there is no water vapor in air, what happens?
12. Goal orientation (DRQ)	*What was the purpose of the city's cutting taxes?
13. Instrumental/Procedural & Enablement (DRQ)	How is a storm created? How will it be possible to break the water molecule?
15. Judgemental	*What do you think about the new taxes?
14. Expectational (DRQ)	Why don't droplets fall in summer? Why is the oxygen in contact with water not completely linked to hydrogen?
16. Assertion	I understood almost everything, only the last part was a little confusing.
17. Request/Directive	*Please tell me how to get a printout of this file.

Graesser, Person, and Huber (1992) define an inquiry as ''an expression (i.e., speech act, utterance) in which the speaker is genuinely seeking information from the listener (p. 169).'' We have used the taxonomy of inquiries developed by Graesser, Person, and Huber (1992) to analyze quantity and quality of summoned up questions (Table 1). However, the original taxonomy was slightly modified by collapsing the ''Instrumental/Procedural'' and ''Enablement'' categories, as explained below.

Some types of inquiries, namely antecedent, consequence, goal orientation, instrumental/procedural, enablement, and expectational, are termed by Graesser and his colleagues as ''deep reasoning questions'' (DRQs). DRQs are highly correlated with inmost levels of cognition (Graesser & Person, 1994) and consequently, they can be regarded as ''high quality questions.'' As such they played an important role in our analysis.

Method

Subjects

A total of 289 Portuguese students participated in the study: 94 from 8th grade, 123 from 10th grade, and 72 from 12th grade. The students belonged to 16 intact classes from four schools.

Three of these schools are located in Coimbra, and one in Moimenta da Beira, a town of 13,000 inhabitants located in northern Portugal. The participating students represented various socioeconomic levels: middle and middle-high classes in Coimbra schools and middle and low classes in Moimenta da Beira.

All of the students had some prior training in science. The Portuguese curriculum includes General Science courses for all the students in fifth and sixth grades and a course on Introductory Biology plus Geology in the seventh grade. Courses on Introductory Physics and Chemistry are also compulsory in eighth and ninth grades. All of the 10th grade students who participated in the study were taking elective courses in Physics and Chemistry as well as Earth and Life Sciences. The 12th grade participants were taking one or two of the following courses: Physics, Chemistry, Biology, or Geology.

With regard to the question-generating activity, it must be added that none of the participants had any previous specific training on question asking or analyzing scientific texts. As far as the discussed tasks and procedures were concerned, it was totally new experience for the polled students.

Materials

The students were instructed to ask questions on two short science paragraphs, which described natural phenomena: ''Clouds'' and ''Dissolved oxygen'' (Table 2). The paragraphs were selected so that, i) each included a content from the areas of physics and chemistry, and ii) each described natural phenomena unfamiliar to the students that prompted them to ask questions. The paragraphs, written in Portuguese, were printed on different pages with instructions to read the paragraphs carefully and to write all questions in the blank spaces provided.

The booklets had three different introductions, corresponding to each of the three mentioned task conditions. In the introduction corresponding to the Class condition, it was explained that the booklet included an activity aimed at developing the capacity to ask questions. It did not include allusions to any tests involving the specified activity. In the introduction corresponding to the Examination condition, the booklet was presented as a test on question generation. The students were informed that they would be graded according to the quantity and quality of questions asked about the included paragraphs. In the introduction relative to the Extra-academic condition there was a commentary stating that the questioning task formed a part of a research project of the Ministry of Education on the improvement of science textbooks. Therefore, students' help was requested to identify comprehension difficulties in the attached paragraphs.

Procedure

The experiment was conducted during one of the regular 50 min class periods. The researchers provided science teachers with the experimental materials and only they were allowed to communicate with the students during the testing session. The teachers were instructed to introduce the activity in a neutral way. The definition of the task was based on the instructions provided on the first page of the booklets. The booklets corresponding to the three conditions were randomly distributed among the students who worked at their own pace. The students were instructed to remain in their seats and read a class textbook after finishing the assignment. After all students handed in their booklets they were debriefed on the purpose of the study.

Table 2
Texts used in the study (translated from Portuguese)

Clouds	Most of the people think that clouds are made of water vapor. But this cannot be so because in this case clouds would be transparent, and we would not see them. Water contained in a cloud is in the form of minute droplets that scatter light and give clouds their characteristic white color. These droplets fall down slowly because of air friction and the turbulence existing in clouds. Their falling speed increases if there is enough water vapor for the water droplets to increase in size. As they fall, they merge with other droplets found in their path and continue to grow.
Dissolved oxygen	Some gases, like oxygen, nitrogen, carbon dioxide or ammonia, are soluble in liquids. The solubility of oxygen in water allows fish to breathe. When the quantity of oxygen dissolved in water decreases because of a polluting process, fish die because of asphyxia. This happens in spite of living in water, H_2O, that contains oxygen in its composition. The reason for this is the inability of fish (and also our inability) to break the water molecule, where the oxygen atom is linked to hydrogen atoms through strong bonds.

Measurements

The questions written by students in the booklets were independently classified by two of the authors using Graesser's et al., scheme as previously explained. First of all, two of the participating authors used the mentioned taxonomy to independently classify 100 questions from a pilot study. There was limited agreement in arranging questions in the categories: 75% agreement in category placement and 96% in classifying questions as DRQs. After this, the two authors jointly classified another set of 160 questions from the present study. Two main decisions were taken during this stage in order to improve the reliability of the instrument. First, the ''Instrumental/Procedural'' and ''Enablement'' categories were collapsed because distinguishing between them proved unreliable. This change did not affect the classification of questions as being DRQs, because both categories were included in the DRQ group. Second, it was found that some questions could be classified in more than one category, which is also an opinion of the taxonomy authors (Graesser & Person, 1994, p. 112). In the cases where one of the categories belonged to the DRQ group, the question was included in this category. In all other cases an agreement was reached through discussion. With these guidelines in mind, an agreement of 94% was reached in category assignment and 98% in DRQ group assignment over a sample of the definitive set of questions that were rated independently.

Results

Table 3 shows proportions of questions from every category made by students at different grade levels. Table 4 shows the average number of questions and average proportions of DRQs asked under each of the three conditions.

Three 3 (grade level) $\times$3 (task demand) ANOVAs were performed with total number of questions, number of DRQs, and proportion of DRQs asked by each student as dependent variables. No significant main effects or interaction were found for the first two dependent variables. With respect to the proportion of DRQs, there was concern about the assumption of normality, since the ANOVA is applied to proportions. In view of this, an arcsine transformation

Table 3
Percentage of questions according to grade level and category

Type of Inquiry	Grade Level			
	8	10	12	Total
1. Verification	15.3	16.2	26.6	18.3
2. Disjunctive	0.6	2.5	3.5	2.1
3. Concept completion	5.8	8.0	4.6	6.5
4. Feature specification	2.8	2.5	0.0	2.0
5. Quantification	0.0	0.6	1.2	0.5
6. Definition	5.6	7.4	7.0	6.7
7. Example	0.0	0.2	0.0	0.1
8. Comparison	0.0	0.8	0.0	0.4
9. Interpretation	0.0	0.0	0.0	0.0
10. Causal antecedent	40.6	35.4	32.4	36.4
11. Causal consequence	4.2	1.0	1.9	2.3
12. Goal orientation	0.0	0.0	0.0	0.0
13. Instrumental/Procedural & Enablement	3.3	3.7	5.0	3.9
14. Expectational	15.3	15.3	15.1	15.3
15. Judgemental	0.0	0.0	0.0	0.0
16. Assertion	6.7	6.5	2.7	5.7
17. Request/Directive	0.0	0.0	0.0	0.0

Table 4
Average number of questions asked ($\bar{n}$) and average proportion of DRQs according to grade level and task condition

Grade Level	Task Condition					
	Class		Examination		Extraacademic	
	$\bar{n}$	Prop DRQs	$\bar{n}$	Prop DRQs	$\bar{n}$	Prop DRQs
8th	3.45	0.54	4.13	0.66	3.91	0.52
10th	4.41	0.55	4.23	0.44	3.33	0.58
12th	3.28	0.67	3.75	0.56	3.78	0.45

was used. Also, an Fmax test (Keppel, Saufley, & Tokunaga, 1992, p. 119) showed that the equal variance assumption was met.

No significant main effects for grade level ($F(2, 280) = .623$, $p = .5$) or task ($F(2, 280) = .883$, $p = .4$) were found concerning this variable. However, a significant interaction of grade level $\times$ task was found ($F(4, 280) = 2.474$, $p = .04$). An analysis of simple effects revealed, first of all, a borderline significant difference among task conditions for the 12th level only ($F(2, 69) = 2.945$, $p = 0.059$). A post hoc LSD test (alpha $= 0.05$) showed that the mean proportion of DRQs for the Class condition (.67) was reliably higher than the mean proportion of DRQs for the Extra-academic condition (.45).

Discussion

Some of the results in the conducted study lend themselves to an easier interpretation than others, especially the negative results. The lack of main effects for each grade level and type of task point to the complex interactions that may exist among the variables, which influence the questioning behavior of students in the quasi-experimental situation as described above (in any case, considerably simpler than a normal classroom situation). We will discuss the positive, more easily interpretable results first. Then we shall suggest some possible explanations for the no-difference results.

The first, main result of the study concerns the quantity and quality of questions asked by the students. Our data support the conclusion that students are capable of asking many questions when given opportunity to do so. As proven by other studies, students are able to ask 26.5 questions per hour, per pupil under one to one tutoring conditions (Graesser & Person, 1994), and 135 when using ''Point and Query'' (Graesser, Langston, & Bagget, 1993). Students in the class condition of our study were able to ask more than three questions on the average, and at any level, from a relatively small amount of scientific information (in comparison to the amount of information usually dealt with in a class). More than 50% of these questions were DRQs. These figures are much higher than those reported in the literature for normal class situations.

The questions asked by students in this study had a different quality. Some of them were low level questions, for example, those triggered by unknown words, like ''What is ammonia? (6-15).'' Other questions corresponded to straightforward inconsistencies between reader's knowledge and text based information, as in the following example: ''I was always told that clouds are water vapor . . . why then do they say that these [clouds] are not water vapor? (23-20)''. Nonetheless, students were also able to produce high quality questions. These were generated when students noticed inconsistencies between their knowledge and inferences drawn from text: ''The text says that clouds have a characteristic white color. Why is it that clouds are darker sometimes? (3-2)'' or ''If water is transparent how is it possible to have white clouds? (21-6).'' These questions may be explained by readers' attempts to build a referential level representation of the text, i.e., the so called ''situation model'' (Van Dijk & Kintsch, 1983). The situation model represents ''what the text is about'' in contrast to the ''textbase,'' that represents the explicit meaning expressed in the text, or to the ''surface code'' that represents the exact syntactical structure and wording. The situation model is built by putting together information from the text and information retrieved from a reader's memory. The student in the mentioned example about clouds attempts to educe more than the explicit meaning from the text—''clouds made of water are white.'' She retrieves knowledge from memory (''water is transparent'') and attempts to link this knowledge to information in the text. The question is triggered by her attempt to build a coherent representation, i.e., a situation model, with these two pieces of information: white water vs. transparent water. However, not every reader attempts to create such a high level of representation. In fact, building an appropriate situation model has been taken as the criteria of successful comprehension (Graesser, Singer, & Trabasso, 1994). Because of this, the high quality questions are less frequent in our linguistic corpus than questions related to the other levels of representation, like the questions concerning unknown terms or incomprehensible statements.

Thus, the students in the current study were able to ask questions of variable quality but, in any case, vastly greater in number than under normal classroom conditions. Limited questioning in regular science classes may not be caused by incapability to detect anomalies, but probably because of an environment hardly suitable for questioning as a mechanism for comprehension regulation. The proven capability of asking relevant questions by our science students, when

compared to the low frequency of quality questions under regular science classroom conditions, indicates that there exists many anomalies that are not transformed into questions, and least of all into answered questions. This fact suggests that finding the effect which these unresolved anomalies have on students' science comprehension monitoring behavior may be an important research topic. Accumulated and unresolved inconsistencies may conceivably have a deleterious effect on students' comprehension standards, or in their disposition to meaningfully understand the world—a variable initially studied by Cohen, Stotland, and Wolfe (1955) under the name ''need for cognition'' (see also, Cacioppo & Petty, 1982).

The second main positive result of our study ensues from the categorization of questions using Graessers et al., scheme (1992): students ask an overwhelming number of causal antecedent questions compared to the other categories. Causal relations provide coherence to texts by linking propositions. There has been much research done on the role of causal relations in processing narratives (see for example, Trabasso, Secco, & Van den Broek, 1984) and on the psychological validity of causal links in readers' representations of these narratives. There is an evidence of increased recall of events causally linked and of the rated importance given by the readers to these events (Trabasso & Van den Broek, 1985; Trabasso & Sperry, 1985). Some results point out the importance of causal relations in understanding expository texts as well (Coté, Goldman, & Saul, 1998; Millis, Morgan, & Graesser, 1990). In particular, Graesser and Millis (1994) found that readers generate causal antecedent inferences when reading short scientific texts, but not causal consequence inferences. This result coincides with our finding that reflects higher occurrence of causal antecedent questions. Generating a question presupposes identifying an anomaly (Otero & Graesser, 1999). This can be done only by comparison between text information and expectations generated by the reader. The inference activity about causal antecedents may be responsible for the generation of these expectations and consequently, trigger causal antecedent questions. Thus, the students in our study appear as active processors who are ''searching after meaning'' (Graesser, Singer, & Trabasso, 1994), trying to explain why the events and phenomena mentioned in the text occur in the first place. However, extrapolations from our study to more naturalistic settings would require some circumspection. Asking causal questions was found to be strongly dependent on reader's goals, as clearly shown in the previously discussed study by Graesser, Langston, and Bagget (1993). Students' goals in many academic situations may be quite different from the goals of the students in our experimental setting. Fostering students' learning goals that produce more and better questioning remains one of the most important challenges in the teaching process.

This result has also clear implications on the design of science texts intended to explain natural phenomena. Students are particularly sensitive to causal connections in this type of explanation. Consequently, writers should be aware of the need to include causal explanations in a manner that can be assimilated by students. A model of causal relations like the one proposed by Einhorn and Hogarth (1986) may be useful in understanding the variables that influence the students' perception of causal strength and it may help in identifying factors which influence their dominion of a causal relation.

Our study generated some unexpected negative results as well. No main effects were found for each grade level or task, as defined in the current experiment. Henceforth, it suggests that these influences may be more complex than we had anticipated. An influence was found for tasks at the 12th grade level only: a classroom setting furthers asking more quality questions than in an extra-academic situation. This outcome coincides with other results obtained while verifying the effect of context in comprehension monitoring. A more formal context promotes better evaluation and regulation of comprehension in science texts (García-Arista, Campanario, & Otero, 1996). This in turn, should result in a significantly higher proportion of DRQs asked in the

classroom environment. However, the mentioned effect was found unreliable in our study. According to our expectations concerning the effect of evaluation demands, the Examination condition should have had a positive effect on questioning as well, when compared with the Extra-academic condition.

However, this was not the case: the difference between proportion of DRQs in the Examination condition (0.56) and the Extra-academic condition (0.45) was in the expected direction but it did not reach statistical significance. This may be simply due to a lack of power of the statistical text originated by the relatively small number of subjects in the two conditions (24 in one group and 23 in the other). There may be another related explanation: the lack of difference might have been caused by a weak manipulation of task demand. Students might not have felt compelled to regulate their comprehension differently just by finding the instructions in the provided booklets: the two conditions were possibly equally demanding. A stronger, more convincing manipulation than the one used here will be necessary in future experiments that examine the effect of students' goals and task demand on question asking.

Other confounding variables might have been responsible for the complex pattern of results which reflect the effect on grade level. For example, it is well known that comprehension monitoring ability, and anomaly detection in particular, improve with grade level (Otero & Campanario, 1990; Otero, Campanario & Hopkins, 1992). This would cause an increase of questions asked by the older students. However, readers' relevant knowledge of the information in the paragraphs would probably increase too, making the paragraphs more comprehensible. Such situations, according to a ''knowledge deficit'' hypothesis (Otero & Graesser, 1999), would lead to less questions asked by the older, more knowledgeable students. Thus, studying the variation of question asking depending on grade level may require controlling potential confoundings like these. An adequate analysis of this relation remains an interesting issue that requires more complex experimental designs than the one established in the current study.

The carried out research was supported by *Acciones Integradas Hispano-Portuguesas* HP 1995-0039 and HP1997-008 to M.H. Caldeira and J. Otero, and by Grant PB93-0478 of DGICYT, Ministry of Education, Spain, to J. Otero. We are indebted to the teachers and students from *Escolas Secundárias Infanta D. Maria, Jaime Cortesão* and *EB23 Alice Gouveia*, from Coimbra, and *Escola Secundária de Moimenta da Beira* for their valuable collaboration. Our thanks to Jerry Keller, from the ICE of the *Universidad de Alcalá*, for the linguistic revision of the manuscript.

References

Alexander, P.A., Jetton, T.L., Kulikovich, J.M., & Woehler, C. (1994). Contrasting structural importance: The seductive effects of teacher questions. Journal of Reading Behavior, 26, 19–45.

Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. Journal of Reading Behavior, 11, 363–374.

Baker, L. (1985). How do we know when we don't understand? Standards for evaluating text comprehension. In D.L. Forrest-Pressley, G.E. Mackinnon, & T.G. Waller (Eds.), Metacognition, cognition and human performance (pp. 155–205). New York: Academic Press.

Campanario, J.M., García-Arista, E., Otero, J., Patricio, A., Costa, E., Prata Pina, E.M., Caldeira, M.H., & Thomaz, M.F. (1994). En qué medida o controlo da compreensao ajuda a melhorar o rendimento académico? [To what extent comprehension monitoring helps in improving academic achievement?] 4° Encontro Ibérico para o Ensino da Física. Covilha, Portugal, 19–23 September.

Cacioppo, J.T., & Petty, R.E. (1982). The need for cognition. Journal of Personality and Social Psychology, 42, 116–131.

Cohen, A.R., Stotland, E., & Wolfe, D. (1955). An experimental investigation of need for cognition. Journal of Abnormal and Social Psychology, 51, 291–294.

Corno, L. (1986). The metacognitive control components of self-regulated learning. Contemporary Educational Psychology, 11, 333–346.

Coté, N., Goldman, S., & Saul, E.U. (1998). Students making sense of informational text: relations between procesing and representation. Discourse Processes, 25, 1–53.

Crooks, T.J. (1988) The impact of classroom evaluation practices on students. Review of Educational Research, 58, 438–481.

Cuerva, J., & Otero, J. (1996). Influence of epistemic authority and knowledge activation on the detection of inconsistencies in science texts. Paper presented at the International Seminar ''Using Complex Information Systems'', UCIS '96, Poitiers, France, 4–6 September.

Dillon, J.T. (1988). The remedial status of student questioning. Journal of Curriculum Studies, 20, 197–210.

Einhorn, H.J., & Hogarth, R.M. (1986). Judging probable cause. Psychological Bulletin, 99, 3–19.

Entwistle, N.J., & Ramsden, P. (1983). Understanding student learning. London: Croom Helm.

García-Arista, E., Campanario, J.M., & Otero, J. (1996). Influence of subject matter setting on comprehension monitoring. European Journal of Psychology of Education, XXI, 427–441.

Glenberg, A.M., & Epstein, W. (1985). Calibration of comprehension. Journal of Experimental Psychology, 11, 702–718.

Good, T.L., Slavings, R.L., Harel, K.H., & Emerson, H. (1987). Student passivity: a study of question asking in K-12 classrooms. Sociology of Education, 60, 181–199.

Graesser, A.C., Langston, M.C., & Bagget, W.B. (1993). Exploring information about concepts by asking questions. In G.V. Nakamura, R.M. Taraban, & D. Medin (Eds.), The psychology of learning and motivation: vol 29, Categorization by humans and machines (p. 411–436). Orlando, Fl: Academic Press.

Graesser, A.C., & Millis, K.K. (1994). The time-course of constructing knowledge-based inferences for scientific texts. Journal of Memory and Language, 33, 583–599.

Graesser, A.C., & Person, N.K. (1994). Question asking during tutoring. American Educational Research Journal, 31, 104–137.

Graesser, A.C., Person, N.K., & Huber, J.D. (1992). Mechanisms that generate questions. En T. Lauer, E. Peacock, & A.C. Graesser (Eds.), Questions and Information Systems (pp. 167–187). Hillsdale, N.J.: Erlbaum.

Graeser, A.C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychological Review, 3, 371–395.

Karabenick, S.A. (1996). Social influences on metacognition: effects of colearner questioning on comprehension monitoring. Journal of Educational Psychology, 88, 689–703.

Keppel, G., Saufley, W.H., & Tokunaga, H. (1992). Introduction to design and analysis (2nd ed.). New York: W.H. Freeman and Company.

Markman, E.M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. Child Development, 50, 643–655.

Millis, K.K., Morgan, D., & Graesser, A. (1990). The influence of knowledge based inferences on the reading time of expository text. In A.C. Graesser & G.H. Bower (Eds.), The Psychology of Learning and Motivation (pp. 197–212). New York: Academic Press.

Otero, J. (1996). Components of comprehension monitoring in the acquisition of knowledge from science texts. In K.M. Fisher & M.R. Kibby (Eds.), Knowledge Acquisition Organization and Use in Biology (pp. 36–43). Berlin: NATO-Springer Verlag.

Otero, J., & Campanario, J.M. (1990). Comprehension evaluation and regulation in learning from science texts. Journal of Research in Science Teaching, 27, 447–460.

Otero, J., Campanario, J.M., & Hopkins, K.D. (1992). The relationship between academic achievement and metacognitive comprehension monitoring ability of Spanish secondary school students. Educational and Psychological Measurement, 52, 419–430.

Otero, J., & Graesser, A. (1999). PREG: Elements of a model of question asking. Manuscript submitted for publication.

Paris, S.G., Lipson, M.Y., & Wixson, K.K. (1983). Becoming a strategic reader. Contemporary Educational Psychology, 8, 293–316.

Pedrosa de Jesús, M.H., & Maskill, R. (1990). Teachers' questioning practices in some portuguese science classes. Revista Portuguesa de Educaçao, 3, 37–56.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: a review of intervention studies. Review of Educational Research, 66, 181–221.

Trabasso, T., Secco, T., & van den Broek, P.W. (1984). Causal cohesion and story coherence. In H.L. Mandl, N.L. Stein, & T. Trabasso (Eds.), Learning and comprehension of text (pp. 83–111). Hillsdale, N.J.: Erlbaum.

Trabasso, T., & Sperry, L.L. (1985). The causal basis for deciding importance of story events. Journal of Memory and Language, 24, 595–611.

Trabasso, T., & Van den Broek, P. (1985). Causal thinking and the representation of narrative events. Journal of Memory and Language, 24, 612–630.

Van der Meij, H. (1994). Student questioning: a componential analysis. Learning and Individual Differences, 6, 137–161.

Van Dijk, T., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.

Weinstein, C.E., & Mayer, R.E. (1985). The teaching of learning strategies. In M.C. Wittrock (Ed.), Handbook of research in teaching (3rd ed.) (pp. 315–327). New York: McMillan.

Zabrucky, K., & Ratner, H.H. (1986). Children's comprehension monitoring and recall of inconsistent stories. Child Development, 57, 1401–1418.

Zabrucky, K., & Ratner, H.H. (1989). Effects of reading ability on children's comprehension evaluation and regulation. Journal of Reading Behavior, XXI, 69–83.

Zabrucky, K., & Ratner, H.H. (1992). Effects of passage type on comprehension monitoring and recall in good and poor readers. Journal of Reading Behavior, 24, 373–391.

Zimmerman, B.J. (1990). Self-regulated learning and academic achievement: an overview. Educational Psychologist, 25, 3–17.