The influence of the length of causal chains on question asking and on the comprehensibility of scientific texts

José Otero,[a,*] Helena Caldeira,[b,1] and Carlos João Gomes[c]

[a] *Departamento de Física, Universidad de Alcalá, 28871, Alcalá de Henares Madrid, Spain*
[b] *Departamento de Física, Universidade de Coimbra, 3000 Coimbra, Portugal*
[c] *Departamento de Ciências da Educação, Universidade dos Açores, 9500 Ponta Delgada, Açores, Portugal*

Abstract

This study examines question asking about *causal* relations when students read scientific texts. We examine the influence of the length of causal chains and the knowledge of readers. Students from grades 8 and 12 read two short paragraphs that described natural phenomena. Length of the causal chain linking cause and effect in two key sentences was manipulated. The students were instructed to ask on anything that they did not understand. The results showed, in the first place, that students ask a majority of causal antecedent questions. Second, a longer causal chain resulted in less causal questions asked by the 12th grade students, but did not reduce the number of causal questions asked by 8th grade students. Third, there was no relation between comprehensibility ratings given to the causal relations and causal question asking. According to this, an explicit comprehensibility score that taps on the metacognitive awareness of students may not be a good indicator of the level of understanding.
© 2003 Elsevier Science (USA). All rights reserved.

1. Introduction

Question asking is an important way by which learners attempt to solve their comprehension problems. It is one of the possible actions taken by learners to regulate comprehension, one of the two components of the metacognitive strategy

---

* Corresponding author. Fax: 34-91-885-4942.
E-mail addresses: jose.otero@uah.es (J. Otero), helena@teor.fis.uc.pt (H. Caldeira), cgomes@notes.uac.pt (C.J. Gomes).
[1] Fax: +351-239-829158.

of comprehension monitoring (Baker, 1985; Otero, 1996; Zabrucky & Ratner, 1986). One way to regulate comprehension is to explicitly state a difficulty as a problem and to ask a question on it.

Some of the studies on question asking have shown that there are few questions asked by students in normal classroom situations, and that most of the questions asked are shallow (Dillon, 1988; Graesser & Person, 1994; Van der Meij, 1988). Why don't students ask deep questions? Where does the question asking mechanism fail? Asking a question addressed to correct a knowledge deficit has several steps: anomaly detection, question articulation and social editing (Graesser & McMahen, 1993). Failing to produce a question may be caused by a problem in any of these steps. This study examines difficulties in the first step, anomaly detection.

Anomalies that may be found in provided information and the questions that may be asked on this information can be of different kinds. Otero and Graesser (2001) proposed a model of question asking that identifies different types of anomalies and corresponding questions that may be asked when learners read problematic texts. In this study we focus on the questions asked on *causal* relations between statements, and on the comprehensibility ratings that students of two grade levels give to these causal relations.

Causal relations play a central role in discourse comprehension models like the constructionist framework of Graesser, Singer, and Trabasso (1994). They are considered important to create coherent text representations (Keenan, Baillet, & Brown, 1984) and to understand both narrative (Trabassso & Sperry, 1985) and expository texts (Coté, Goldman, & Saul, 1998). Regarding question asking, previous research has shown that causal relations are the most important source of questions on scientific texts that describe natural phenomena (Costa, Caldeira, Gallástegui, & Otero, 2000). This implies that students are sensitive to the comprehensibility of causal relations, and are able to detect difficulties in the causal relations existing in science texts that describe natural phenomena.

Given the salience of causal relations, it is important to identify variables that influence the detection of problems in these relations and, consequently, variables that influence the generation of questions on these relations: Why do readers ask a question on a causal relation? What makes a causal relation more or less comprehensible? Research on causal inferences is relevant in order to answer the previous questions. Studies in this area have shown that readers are sensitive to the degree of causal relation between two sentences. In a study by Keenan et al. (1984), subjects read pairs of sentences having different levels of causal relatedness, according to previous norms. Some pairs were seen as strongly related, such as "Joey's big brother punched him again and again. The next day his body was covered with bruises." This condition was compared to another where the first sentence was substituted resulting in a weaker causal relation: "Joey went to a neighbor's house to play. The next day his body was covered with bruises." The researchers found that subjects needed more time to read the second sentence of the pair when it was weakly related to the first one. Myers, Shinjo, and Duffy (1987), replicated Keenan et al.'s (1984) study under more controlled conditions and found the same results. This supports the assumption of readers trying to relate causally both sentences. The increase in

reading time is interpreted as a consequence of the increased difficulty to relate the sentences in the condition of weak causal relation.

Other studies have analyzed the mechanism responsible for the detection of difficulties in causal relations. According to the "inference validation model" (Singer, Halldorson, Lear, & Andrusiak, 1992) readers try to validate tentative bridging inferences, in particular causal inferences, while processing a text. Singer et al. (1992) showed that readers activate relevant knowledge to validate causal bridging inferences that link sentences such as "Dorothy poured the bucket of water on the fire. The fire went out." Also, readers were able to notice inconsistencies in sentences like "Dorothy poured the gasoline on the bonfire. The fire went out" (Singer, 1993). This is interpreted to reflect a validation process that consists, first, in deriving a mediating proposition that could causally link both sentences (Gasoline is not combustible). Then, this idea is checked against reader's knowledge so that it is either validated or rejected.

Singer, Harkness, and Stewart (1997) claim that readers may try to validate bridging causal inferences even where content is unfamiliar, as it is frequently the case with scientific texts. For example, a student with little knowledge of chemistry that reads: "Ron dipped the litmus paper in the acid. The paper turned red" may generate the mediating idea "Acid turns litmus paper red." Although no knowledge will be available to validate this causal bridging inference, the reader might accept its truth on the assumption that the writer is cooperative and is telling the truth (Grice, 1975).

However research on question asking, like the one mentioned above, suggests that not all causal relations in scientific texts are equally accepted on the cooperative assumption. For example, the following phrase was included in a paragraph used in the study on question asking by Costa et al. (2000): "When the quantity of oxygen dissolved in water decreases because of a polluting process..." Twelfth grade students who did not have specific knowledge of this relation read the paragraph. They had instructions to ask on anything that they did not understand. The causal relation expressed by the previous sentence was the second most frequent source of questions in the paragraph: "Why does pollution cause a decrease in oxygen?" (Otero & Graesser, 2001) This poses the problem of explaining why readers lacking knowledge needed to validate a causal bridging inference ask a question on the causal relation, and do not accept it on the cooperative assumption.

Keenan et al.'s (1984) work mentioned above suggests that one factor which influences question asking on causal relations is the level of causal relatedness. The level of causal relatedness was calculated in Keenan et al.'s (1984) study through a norm study. Einhorn and Hogarth's (1986) model of causality includes the similar concept of "strength" of a causal relation. In this model, a "generative" force that links causes and effects is included as one of the factors that determine the strength of a causal relation: "events need to be linked through a causal chain so that the force can be transmitted from one link to the next" (Einhorn & Hogarth, 1986, p. 10). For example, two events may be separated by a temporal and spatial gap: sunspots and changes in the stock market, in Einhorn and Hogarth's (1986) example. They may be perceived as causally related only insofar as it is possible to construct a causal chain that bridges this gap. For example, one may imagine the following

causal chain: sunspots cause changes in weather, weather causes changes in crops, and changes in crops cause changes in the stock market.

The strength of the relation between cause and effect depends on the strength of the causal chain that connects both. Two causally related events described in a text might appear disconnected to a particular reader. However, another reader with sufficient knowledge may create an appropriate causal chain by generating links between the two events in the form of bridging causal inferences, as in the examples above. Alternatively, a writer may explicitly provide links between two causally related events in an attempt to increase the strength of the causal relation and to facilitate its understanding. For example, one may present the relation mentioned above between pollution in rivers and the disappearance of oxygen dissolved in water using causal chains that have different length, i.e., a different number of component links. The resulting causal relations would presumably have different strength for a particular audience. For example, a relatively weak causal relation exists between "pollution" and "diminution of oxygen dissolved in water" for Portuguese primary and secondary school students, according to the number of questions asked in the study by Costa et al. (2000) mentioned above. A more comprehensible causal relation, for some readers, could be obtained by introducing a link relating pollution to the appearance of micro-organisms that consume oxygen, which leads to its diminution. An even longer chain could be created by introducing an additional link: micro-organisms feed on the organic materials present in the pollutants; this metabolic activity consumes oxygen in water, causing its diminution. Will the addition of new links indefinitely add to the perceived strength of the chain and to the comprehensibility of the causal relation? According to Einhorn and Hogarth's (1986) model, strength of a causal relation depends through a multiplicative function on the strength of each of the component links—a strength that ranges from 0 to 1. Some of the new, more precise links may actually be weak for a particular reader, diminishing the strength of the chain.

According to this, there is probably a nonmonotonic function relating length and comprehensibility. Some causal relations are difficult to understand, as in the example of sunspots and the stock market. One may increase the comprehensibility of a causal relation by adding causal links. However, adding many links may actually decrease the comprehensibility of the whole chain when one or several of these links are weak for a particular subject. This may be the case when microscopic causal explanations of natural phenomena are provided to the non-expert. In the example above, the causal link between metabolic activity and consumption of oxygen may be strong for 12th grade students and possibly weak for the less knowledgeable, younger students.

Thus, strength should not be considered in absolute terms but relative to a particular knower. The variable that is actually manipulated in these examples is the length of the causal chain. This is also related to the depth of the causal explanation because adding elements to a causal chain in a scientific text, like the ones used in this study, frequently implies using elements at a deeper, more microscopic level.

To summarize the reasoning above, we hypothesize that the comprehensibility and the number of questions asked on a causal relation depend nonmonotonically on the causal chain's length. An explanation of a natural phenomenon by a causal chain where very few elements are made explicit, as in the relation between pollution

and diminution of oxygen, is expected to trigger more questions than the same phenomenon explained by a more explicit, longer, causal chain. However, increasing the number of links of a causal chain may have a different effect for readers with different knowledge: readers will take profit of additional links only when these are individually comprehensible. For less knowledgeable readers these links may be weak, and useless for increasing the perceived causal strength of the relation.

In this study we examine the influence of these two variables: the length of a causal chain, defined in terms of number of links made explicit, and the knowledge of readers. We want to discover their effect on comprehensibility and question asking on the causal relations existing in two short scientific paragraphs that describe natural phenomena. According to this, the questions that this study attempts to answer are: Is comprehensibility of and questioning on a sentence containing a causal relation related to the length of the causal chain linking cause and effect? Do readers with different knowledge ask different questions on sentences containing short or long causal chains?

2. Method

2.1. Subjects

One hundred and sixty three Portuguese students from grades 8 and 12 participated in the study. All of the students had some prior training in science.

The two experimental paragraphs that were used in this study respectively deal with the solution of oxygen in water and the scattering of light in clouds. This specific information had not been taught to the students, according to the information provided by the teachers. Although no pretest was made to probe students' knowledge related to these topics (and that remains as one limitation of this study), we assumed considerable differences in this knowledge, given the important difference in grade level. There are general science courses for all the students in 5th, 6th, and 7th grades of the Portuguese curriculum. Courses on Introductory Physics and Chemistry are also compulsory in 8th and 9th grades. However, although water and light phenomena are included in science curricula of grades 5th to 9th, the topics of the two experimental paragraphs are not included at any of these levels.

In contrast, the 12th grade participants were taking one or two of the following courses: Physics, Chemistry, Biology or Geology. In 10th grade and 11th grade, they all had taken Physics and Chemistry that are compulsory for students in a scientific branch. Solution of gases in liquids, and the concepts of reflection and refraction are included in the physics and chemistry curricula at these levels.

Students from each grade level were randomly distributed in two conditions, as explained below, corresponding to length of the causal chain: short vs. long. An adjustment in the number of students was made after having gathered the responses. In order to obtain an equal number of 40 students in every condition at each grade level, results from one student randomly chosen were discarded from one of the groups, and those corresponding to two students in another group.

2.2. Materials

Two short paragraphs, "Clouds" and "Dissolved Oxygen," that had already been used in a previous experiment on question asking (Costa et al., 2000) were adapted to be used in this study. The "Clouds" paragraph is shown as an example in Table 1. Sentences are numbered for clarity, although they were not numbered in the paragraph read by students. Sentences, #4 and #7, were manipulated as follows: in a "Short" version (S), the causal relations were presented in a way that was not obvious to students, according to previous talks with experienced teachers; in a "Long" version (L) the causal relations were explained by adding a link between cause and effect.

The two key sentences manipulated in each paragraph were the most questioned sentences in the previous experiment (Costa et al., 2000). The paragraphs in the two conditions, S and L, differed only in these two key sentences. In addition, each paragraph included one control sentence that was the same in the two conditions. It was used to adjust for biases in judging comprehensibility, as explained below.

2.3. Procedure

The experiment was carried out during a 50 min regular class period. One of the researchers randomly (according to position in the classroom) provided students with booklets that included the two paragraphs, either in the short or long versions, and two blank pages where students could write questions. These booklets had instructions to carry out the task on the first page. The students were instructed to read the paragraph, one at a time, ask in writing any question that they may have on the paragraph, and to rate the comprehensibility of each key sentence on a five point scale, from Very Poor (1) to Very Good (5). The students had to rate the comprehensibility of the control sentence too. Ratings of the key sentences were calculated relative to the comprehensibility rating of the control sentence. Thus, any bias toward high or low ratings caused by personal factors was corrected.

Table 1
Versions of the "Clouds" paragraph used in the study (translated from Portuguese)

Clouds
1. Most of the people think that clouds are made of water vapor
2. But this cannot be so because in this case clouds would be transparent, and we would not see them
3. Water contained in a cloud is in the form of minute droplets (control)
4. S. Clouds' white color is caused by sunlight that impinges on the droplets (key sentence, S version)
4. L. Clouds' white color is due to droplets reflecting and refracting sun's rays in various directions (key sentence, L version)
5. These droplets fall down slowly because of air friction and the turbulence existing in clouds
6. When there is enough water in air the droplets increase in size
7. S. When they increase in size their falling speed increases (key sentence, S version)
7. L. When they increase in size droplets' weight increases more than air resistance and as a consequence their falling speed increases (key sentence, L version)
8. As they fall, they merge with other droplets found in their path and continue to grow

3. Results

3.1. Questions asked

The questions asked by students on any sentence of the two texts (although the examples used here are taken from the Clouds text only) were classified using a slightly modified version of Graesser, Person, and Huber's (1992) categories, as explained in Costa et al. (2000). According to this taxonomy, questions can be classified in 17 types. The four categories more frequent in our data, i.e., each of them amounting to more than 10% of total number, were "Verification" (for example, "Are clouds made of gases?), "Causal Antecedent" ("Why do droplets refract sun rays?"), "Expectational" ("If droplets in clouds are made of water, and water is colourless, why are clouds white or grey and not transparent?"), and "Assertion" ("I do not understand why droplets fall down with more or less speed due to friction"). Percentages of questions belonging to these categories are reported in Fig. 1. These percentages are compared to those found in the study of Costa et al. (2000) where the original, non-manipulated paragraphs were used. The high percentage of questions in the Causal Antecedent category found in that study is clearly replicated here.

We were especially interested in questions addressing the causal relation in the key sentences. However, these questions may be classified in several of Graesser et al.'s (1992) categories, for example Causal Antecedent ("Why do droplets send sun rays in different directions?"), or Causal Consequence ("What happens [to the colour of clouds] when there is no sun?"). Questions in any category that were addressed to the causal relation were classified as "Causal Questions."

The number of questions asked on terms, "What are droplets?" for example, were computed also. Noticing an incomprehensible term is the simplest situation that may
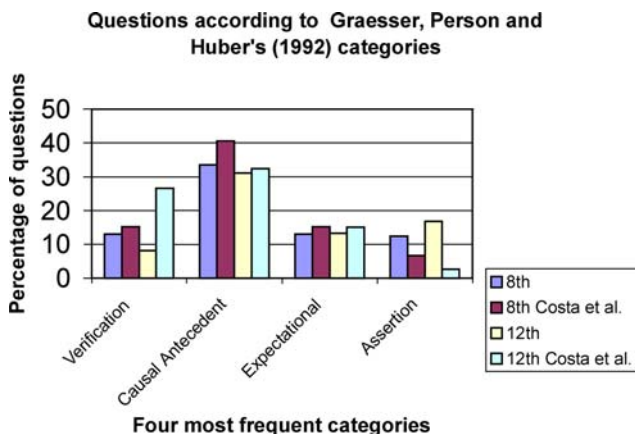

Fig. 1. A comparison of the distribution of the four most frequent categories of questions in this study and in Costa et al.'s study (2000).

Table 2
Frequency of questions on the key sentences

Type of question	Causal		Terms		Other		Total	
Length of causal chain	Short	Long	Short	Long	Short	Long	Short	Long
8th Grade	49	50	5	19	14	12	68	81
12th Grade	27	13	4	2	11	8	42	23

trigger a question (Otero & Graesser, 2001). Questions on terms belong most frequently to the "Definition" category, but they can be found in other categories also like Assertion ("I do not understand what refract means") or Verification ("Droplets are what we usually call rain?"). All of these were computed as questions on terms.

Table 2 presents the total number of questions, the number of causal questions, the number of questions on terms, and the number of other questions asked on the four key sentences. As the number of subjects is the same in every condition, frequencies are directly comparable. The distribution of questions in the two levels of length of the causal chain is significantly different in 8th and 12th grade, both for the total number of questions ($\chi^2 = 6.52$, $p = .01$) and for causal questions ($\chi^2 = 3.72$, $p = .05$). Fisher's Exact test indicates that there is also a significant difference for the distribution of questions on terms ($p = .049$).

Also, a χ^2 goodness of fit test shows that the number of causal questions asked by 12th grade students is significantly smaller for the long version of the causal relation than for the short version ($\chi^2 = 4.90$, $p = .03$). This is not observed in 8th grade. However 8th grade students ask significantly more questions on terms in the long condition than in the short condition ($\chi^2 = 8.17$, $p < .01$).

In order to examine a possible interference of knowledge of terms on causal question asking, a conditional probability analysis was carried out. The key sentences include more technical terms in the long version than in the short version. Terms like "refracting" or "micro-organism," can be found in the long versions only. Although both terms are used in 8th grade Portuguese curricular materials, they may not be well known by all of the students. The analysis showed that the probabilities of 8th grade students asking a causal question given that they have not asked a term question were .91 and 1.00 for the S and L versions respectively. These probabilities were .64 and .29 for the S and L versions respectively in 12th grade. The difference between versions is not statistically significant in 8th grade ($p = .23$), but it is statistically significant in 12th grade ($p < .01$).[2]

---

[2] As the samples do not allow for classical approximations to normal distributions, we tested for differences using one sided exact tests, i.e., tests where the probability distributions of proportions were exactly characterized. Distributions of proportions belong to a binomial family where the critical region has a probability that depends on the second parameter of these distributions. To calculate *p* values we used a bayesian approach assuming an uniform distribution of this second parameter in the interval [0, 1].

3.2. Comprehensibility ratings

Fig. 2 represents mean relative comprehensibility ratings of key sentences for the two explanation levels in 8th grade and 12th grade. There are consistently lower relative comprehensibility ratings of the long version for all target sentences in 8th grade. No such difference is observed in 12th grade. χ^2 Tests were used to test the differences between length conditions. Relative comprehensibility ratings range from 4 (a rating of 5 for the key sentence minus 1 for the control sentence) to -4 (the opposite situation). In order to avoid very small or zero number of cases in some cells, 3 (rating categories) $\times$ 2 (length conditions) contingency tables were created by collapsing the relative comprehensibility ratings into three categories: (1) those ratings of the key sentence that were below the rating of the control sentence, i.e., $a-1$ relative comprehensibility score or less, (2) those ratings of the key sentences that were like the rating of the control sentence, i.e., a 0 relative comprehensibility score, and (3) those ratings of the key sentence that were above the rating of the control sentence, i.e., $a+1$ relative comprehensibility score or higher. The results for 8th grade students show that the difference in the distribution of comprehensibility ratings between strength conditions is significant for the two key sentences in the oxygen paragraph ($p < .01$ for oxygen 1 and $p = .02$ for oxygen 2) and not significant for the key sentences in the Clouds paragraph ($p = .17$ for Clouds 1, and $p = .77$ for Clouds 2), although the decreasing trend between the short and long condition remains.

No significant differences ($p > .05$) between length conditions are found in the distribution of comprehensibility for any of the key sentences in 12th grade.

4. Discussion

Three main results can be synthesized in our study. In the first place we have replicated the results obtained by Costa et al. (2000) using a slightly different version of the paragraphs: students ask a majority of causal antecedent questions when trying
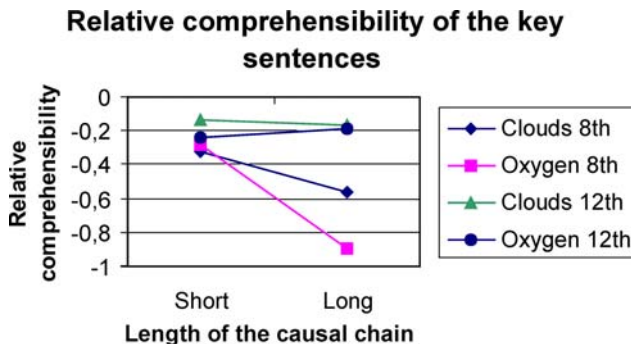

Fig. 2. Relative comprehensibility of the four key sentences.

to understand these texts. This points to the important role played by causal relations in texts that describe natural phenomena. Readers are especially sensitive to why events occur, and textbook writers and teachers should address this concern providing appropriate explanations of these occurrences. As already pointed out for narratives (Graesser et al., 1994), readers of scientific texts describing natural phenomena also ''search after meaning'' and try to explain why the events and states mentioned in the text occur.

The second result contributes to clarifying the way in which these explanations should be presented: it shows the effect of providing causal chains of different length to students of quite different age levels that presumably have different topic knowledge. Length of causal chains was found to influence the causal questions asked by the 12th grade students in our study. They asked significantly fewer causal questions when provided with an explicit causal chain between cause and effect, than when provided with cause and effect without any other explanation.

Eighth grade results were different. These students did not ask more causal questions on the short version compared to the long version. Before concluding that 8th grade students' questions are not affected by the length of the causal chain, however, one needs to rule out the increased lexical difficulty associated with the long version as a possible explanation. Not understanding the terms could conceivably prevent students from asking causal questions on the relation involving these terms. However, would 8th grade students who *do not have lexical problems* ask less questions on the long version than on the short version, as 12th grade students do? We partialled out the effect of difficult terms by calculating the conditional probabilities of asking a causal question given that a term question had not been asked. These probabilities correspond to the students who do not have lexical problems. As shown above, no significant difference was found in these probabilities: .91 and 1.00. So, when judged by the number of causal questions asked, the 8th grade students who understood the terms found equally incomprehensible also the short version and the long version of the causal relation.

Thus, the different results obtained for 8th grade students and 12th grade students indicate a different reaction to the length of causal relations, corresponding to different perceived strengths. Twelfth grade students ask more questions on shorter causal chains. This may be interpreted in terms of the validation model proposed by Singer et al. (1992) as reflecting an attempt to generate the bridging inferences that link cause and effect. These students are unable to create these links and ask questions to find them. The links are explicitly provided in the long versions. They are found to be appropriate by 12th grade students, and the need to ask a question is diminished. In contrast, 8th grade students' asking of causal questions does not differ significantly between versions. After having discarded an explanation in terms of lexical problems, this no difference result may be explained by a difficulty to understand the new links added to the causal chains. Adding a new link did not improve comprehensibility of the causal relation for 8th grade readers, i.e., did not increase the perceived strength. This seems to place a limit on the comprehensibility of certain causal relations, like the ones used in the study, by younger students. This limit should

probably be attributed to differences in students' content knowledge, more than to a direct effect of developmental differences.

In the previous discussion comprehensibility has been related to the number of questions asked. However, this study also included a request to evaluate comprehensibility by directly giving a score on a 5 point scale. A third important result of the study is the absence of a relation between this comprehensibility rating and causal question asking. Eighth grade students rated as less comprehensible the long version although there was not a corresponding increase in the number of causal questions asked. In contrast, 12th grade students asked more causal questions on the short version than on the long version. However this was not translated into lower comprehensibility ratings of the key sentences in the short version. These mismatches indicate that the lower comprehensibility ratings that 8th grade students give to the long version seem not to be caused by increased comprehension difficulties of the causal relation itself but by some other factor. The greater number of questions asked on terms in the long condition suggests that unknown words may be such a factor. These students may find the long version less comprehensible because of the existence of difficulties that are related to terminology rather than to the causal relation itself. This is in agreement with the ready use of the lexical criteria to monitor comprehension by young students, as pointed out by Baker (1979, 1985).

Twelfth grade students ask significantly more causal questions on the short version. However this is not translated into a lower comprehensibility score. According to this, an explicit comprehensibility score that taps on the metacognitive awareness of students as requested in this study, may not be a good indicator of the level of understanding of the causal relation. A less explicit, indirect index, like the number of questions asked, would be more appropriate.

To summarize, the results above show that 12th grade students' questioning is sensitive to the length of causal chains linking cause and effect in scientific texts. Twelfth grade students are able to detect weak causal relations and ask questions on them. This is in agreement with results of the studies reported above on the validation of causal relations and on the capability of students to ask pertinent questions, when placed in certain situations. However, a different pattern was found for 8th grade students. Causal relations of different lengths were equally incomprehensible, when judged by the number of question asked. Adding links to the causal chains did not make the relations more comprehensible, i.e., did not increase the strength perceived by these students, even for those understanding the terms.

The results obtained in this study suggest new problems also. As it was pointed out before, it is unreasonable to assume that question asking and causal chain length are related through a decreasing function for the whole range of chain length. An extremely detailed causal chain would trigger more questions than one less explicit, because the former would involve mediating ideas that may be difficult to understand for particular readers (for example, ideas needed to account in detail for the use of oxygen in the metabolism of micro-organisms). Thus, one may speculate about the existence of an inverted U-shaped function, relating chain length and comprehensibility having a maximum that corresponds to optimal comprehensibility for particular readers. Interestingly, non-monotonic functions have been suggested also for

the relation between causal strength and memory (Myers & Duffy, 1990; Myers et al., 1987). Characterizing this optimal chain length for particular readers is a worthwhile challenge.

Acknowledgments

References

Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. *Journal of Reading Behavior, XXI*, 365–374.

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

Costa, J., Caldeira, M. H., Gallástegui, J. R., & Otero, J. (2000). An analysis of question asking on scientific texts explaining natural phenomena. *Journal of Research in Science Teaching, 37*, 602–614.

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

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

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

Graesser, A., & McMahen, C. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Educational Psychology, 85*, 136–151.

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

Graesser, A. C., Person, N. K., & Huber, J. (1992). Mechanisms that generate questions. In T. W. Lauer, & E. Peacock (Eds.), *Questions and information systems* (pp. 2–23). Hillsdale, NJ: Erlbaum.

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

Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. Morgan (Eds.), *Syntax and semantics 3: Pragmatics*. New York: Academic Press.

Keenan, J. M., Baillet, S. D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior, 23*, 115–126.

Myers, J. L., & Duffy, S. A. (1990). Causal inferences and text memory. In A. C. Graesser, & G. H. Bower (Eds.), *The psychology of learning and motivation: Vol. 25* (pp. 159–173). San Diego: Academic Press.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26*, 453–465.

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.

Otero, J., & Graesser, A. (2001). PREG: elements of a model of question asking. *Cognition and Instruction, 19*, 143–175.

Singer, M. (1993). Causal bridging inferences: Validating consistent and inconsistent sequences. *Canadian Journal of Experimental Psychology, 47*, 340–359.

Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language, 31*, 507–524.

Singer, N., Harkness, D., & Stewart, S. T. (1997). Constructing inferences in expository text comprehension. *Discourse Processes, 24*, 199–228.

Trabassso, T., & Sperry, L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language, 24*, 595–611.

Van der Meij, H. (1988). Constraints on question asking in classrooms. *Journal of Educational Psychology, 80*, 401–405.

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