Running Head: AUTOMATED FEEDBACK FOR SUMMARIZING

Summary Street: Interactive Computer Support for Writing

David Wade-Stein and Eileen Kintsch

University of Colorado – Boulder

Correspondence about this article should be addressed to

Eileen Kintsch

Institute of Cognitive Science

University of Colorado

Boulder, CO 80309-0344

Email: ekintsch@psych.colorado.edu

Tel: 303-492-0736

Abstract

Summary Street is educational software that incorporates cognitive research on the development of summarization skills with the meaning representation method, called Latent Semantic Analysis. *Summary Street* provides students automatic feedback on the content of their summaries. The feedback is presented in an easy-to-grasp, graphic display that motivates students to improve their writing across multiple cycles of writing and revision on their own before handing it in to the teacher for final evaluation. The software thus has the potential to provide students with extensive writing practice without increasing the teacher's workload. In classroom trials sixth-grade students not only wrote better summaries when receiving content-based feedback from *Summary Street*, but also spent more than twice as long engaged in the writing task. Improvement in content scores was greater when students were summarizing more difficult texts. The authors suggest that *Summary Street* could be adapted to a wide variety of instructional goals beyond summary writing.

Summary Street is an interactive computer tutor designed to support a key learning strategy: constructing a written summary of instructional text on relatively unfamiliar topics. The purpose of summarizing is to help learners integrate new material with what they already know and in so doing to understand the material at a deeper level than they would by simply reading or even rereading a text. *Summary Street* supports learners by providing nearly instantaneous feedback on the content adequacy of their summary. It thus enables learners and their teachers to quickly assess comprehension in an ongoing fashion and to pin point problem areas to focus on. At the same time, summarizing also reinforces writing skills. Since the computer rather than the teacher can perform the onerous chore of assessing and grading summaries, students can receive the extensive practice needed to acquire this important skill.

In this paper we describe how *Summary Street* works, why such a tool is effective, and when and for whom it works best. The How and When questions can be answered in a straightforward manner on the basis of the experimental study described below. A review of the theoretical framework offered by the Kintsch (Kintsch, 1998; van Dijk & Kintsch, 1983) model of discourse comprehension, together with a brief review of the extensive empirical literature on the relationship between reading, writing and learning provides an answer to the Why question, with which we begin.

Why Should Children Learn to Summarize?

An important series of studies by Ann Brown and her colleagues in the 1980s demonstrated the importance to comprehension and learning of the ability to succinctly summarize what one has read. Their research (e.g., Brown, Bransford, Ferrara, & Campione, 1983; Brown & Day, 1983; Brown, Day & Jones, 1983) and subsequent studies also documented a developmental progression in learning this skill, with persistent weaknesses that extend to college students. It is simply assumed that students will acquire the skill on their own as they read instructional text. However, as Moore (1995) points out, NAEP (1990, 1993, 1998) reviews of educational progress show that school-age students continue to have trouble understanding expository text. Moore attributes this to the fact that students lack familiarity with the more complex and varied structures of expository text. Knowledge of diverse text structures helps the reader to organize information during comprehension and facilitates efficient retrieval later. Although the trend among educational researchers has been to emphasize direct instructional interventions, we believe the problem could be addressed simply by giving students intensive practice in working through many different examples of expository text at a deep level, for example, by summarizing them. Lack of opportunities to practice actively comprehending text is a major reason why students remain stuck in a passive reading mode (cf. Paris, Wasik & Turner, 1991).

There are a number of additional reasons why summarizing is an important learning tool, not only for academic success, but in the world beyond. Our argument is theoretically based, but is broadly supported by educational research conducted in the 1980s- to present (e.g., see review by Bransford, Brown, and Cocking, 2000) that emphasizes learning with deep understanding

and summarizing tasks as a key method for assessing whether or not this is occurring. Newer assessment methods – those that test student progress against a set of graded standards – also reflect this direction in educational research, by emphasizing open-ended essay and summary writing tasks over traditional multiple-choice questions.

A theoretical account of how summarizing information fosters deep comprehension and learning is provided by the Kintsch (1998) and van Dijk & Kintsch (1983) model of discourse comprehension. Namely, summarizing contributes to the goal of constructing a solid foundation of factual and conceptual knowledge because it serves to reinforce the memory representation of the content beyond that achieved through rereading. Writing a summary requires much more conscious thought, judgment and effort, whereas rereading is generally a more passive activity. Ideally, the summary writer not only selects the important ideas from the text, but also reconstructs the meaning in a more succinct, generalized form (van Dijk & Kintsch, 1983). The payoff is in terms of a better integrated, well connected memory representation that enables a learner to do something more with the material than simply reproduce it: with multiple links in the knowledge base, newly acquired knowledge becomes more accessible when needed to solve a new problem, to support an argument, or to share one's understanding with colleagues or classmates (Mannes & Kintsch, 1987; Kintsch & Kintsch, 1997). Teachers have noted marked differences in depth of understanding of topics that students have summarized as opposed to merely read, as demonstrated by their more detailed and thoughtful contributions to classroom discussions. Not surprisingly, this observation is reinforced by empirical evidence showing better comprehension (e.g., Bean & Stenwyk, 1984;

Casazza, 1992; Moore, 1995) and better retention by students of material that has been summarized (e.g., Rinehart, Stahl, & Erickson, 1986; Ross & DiVesta, 1976; Taylor, 1982; Taylor and Beach, 1984). Brown and Palincsar (1983; Palincsar & Brown) have moreover shown that guided practice in summarizing text does transfer to other classes and other subject domains.

As an assessment tool, a summary written in one's own words is more revealing of a student's understanding than the commonly used end-of-chapter questions. For one, it is difficult to ask the right questions: multiple-choice questions mainly test recognition memory; the answers to factual queries can be lifted right out of the text with minimal effort; inference questions are easily misunderstood and often require background knowledge that students may not have. Simply having students summarize a chapter is more likely to uncover specific gaps in their understanding. It can help students formulate the right questions to ask, and can show teachers where more instruction is needed.

For these reasons summarizing plays a central role in a number of programs seeking to train students in better comprehension and learning strategies (e.g., Baumann, 1984; Bean and Steenwyk, 1984; McNeil & Donant, 1982; Palinscar & Brown, 1984; Rinehart, Stahl & Erikson, 1986; Uttero, 1988). Summarization training and practice is especially beneficial to the comprehension and recall of lower achieving students and those with learning disabilities (e.g., Gajria & Salvia, 1992; Jitendra, Cole, Hoppes, & Wilson, 1998; Malone & Mastropieri, 1992; Palincsar & Brown, 1984; Reis & Leone, 1987; Weisberg & Balajthy, 1990), and is one of the core strategies taught in the highly effective reciprocal teaching method (Brown & Palincsar, 1989; Palincsar & Brown, 1984). According to Brown and Palincsar, metacognitive strategies such as summarization and paraphrasing

function as important self-testing mechanisms for assessing whether one is really comprehending what one reads. Few tasks require coordinating comprehension and writing activities to the degree that summarizing does (Tierney & Shanahan, 1991), as both depend on very similar sets of skills. Like writing, deep comprehension relies on sorting out the importance relationships in a text, on finding links between ideas and organizing them into a coherent textual or mental representation of the meaning. Summarization tasks further a student's writing skills by providing practice in a situation that is more structured than open-ended, creative writing, thus forming a bridge for students to learn how to write informational texts of their own.

Summarization clearly has a great deal of potential for improving students' learning and writing. Hence, it is unfortunate that it is largely neglected throughout children's academic training. There may be several reasons for this: for example, some teachers may feel uncertain about how to instruct summarizing, what operations to focus on, how to model the process. However, the major reason is no doubt that it is an overwhelming amount of work for teachers to evaluate and provide individual feedback on successive drafts of summaries, or on other kinds of extended writing tasks. The advantage of a computer-based system such as *Summary Street* is that students can receive the practice they need in summary writing and revising without increasing the teacher's work load. Guided by feedback on the content of their summaries, students can do much of the preliminary revising on their own. The fact that the automatic feedback is delivered almost instantly and can be summoned as often as desired, engages students in a cycle of writing, evaluating feedback and rewriting as long as they are motivated to improve.

Of course, it takes more than getting the correct content to write a good summary, but in general, formulating that content is a good starting point, given that the ability to isolate the main ideas is a difficult problem for most young learners, as is clearly revealed in their recall and summaries (e.g., Brown & Smiley, 1977; Brown, Day, & Jones, 1983). *Summary Street* by no means would replace a teacher's input, which provides the kind of feedback that an automatic system cannot supply: a critique of writing style, helping students resolve problems with organization, misconceptions, and mechanics. But although a computer tutor cannot mimic a human teacher, it can nevertheless greatly extend the effectiveness of a teacher's instruction. The following sections provide a detailed look at how Summary Street works and describe how users interact with the system.

LSA: The Basis for Summary Evaluation

Summary Street is a web-based system that provides an environment in which students can prepare multiple drafts of a summary with feedback about the adequacy of the content of what they have written. The feedback is generated by Latent Semantic Analysis, or LSA, which is a machine learning method that constructs semantic representations that in many ways mirror human semantics. Informative introductions to LSA are available (Landauer & Dumais, 1997; Landauer, Foltz & Laham, 1998; Landauer, 1998), and we can only briefly and informally describe here how LSA works and what its strengths and limitations are.

LSA represents the meaning of words, sentences and whole texts mathematically, as vectors in a high-dimensional semantic space. The semantic

space is created by analyzing the use of words in a very large set of written documents. A huge matrix is formed of millions of words and thousands of documents in which they occur. This matrix makes it possible to perform calculations on meanings, for instance, to evaluate the semantic distance between the meanings of two different words or texts. Texts with similar meanings are located near to one another in the space, and likewise words that occur in similar contexts are grouped together as semantically related (e.g., *doctor* and *physician*, or *run* and *running*).

To construct the semantic space, LSA needs as input a large amount of text (a corpus based on 11M words is used in *Summary Street*); it then reduces the dimensionality of this corpus by standard mathematical methods to about 300 dimensions. The effect of this dimension reduction is that the idiosyncrasies and irrelevant aspects (particular word use and grammatical variants)) of the input texts are de-emphasized while essential semantic relations are preserved. Once LSA has constructed this semantic space, it is very easy to compute the semantic distance between words, as well as between whole texts. The measure of semantic distance that we use is the cosine, which can be interpreted as a kind of correlation. Words (or texts) that are highly related have a high cosine (identity is 1; *lamp* and *light* have a cosine of .61); words that are unrelated have a cosine close to 0 (e.g., *lamp* and *bear* have a cosine of .03). As applied to *Summary Street*, this means we can readily compute the cosine between the summary a student has written and the to-be-summarized text; if the cosine is below a certain empirically determined threshold value, we tell the student that his or her summary does not yet cover the content sufficiently well.

LSA allows us to calculate with the meanings of words and texts without relying on the presence or absence of special keywords or terms. This capability makes it a powerful tool in a variety of applications. Most relevant in the present context is LSA's ability to grade essays as well as trained human graders (Foltz, Gilliam, & Kendall, 2000; Landauer, Laham, & Foltz, 2000). *Summary Street*, however, has a different purpose: rather than assigning grades to essays, it is designed to help students write better essays. LSA is also of theoretical interest, in that it can serve as the basis for a theory of how humans process meaning (e.g., Landauer, 1998; Kintsch, 2001). To help the reader understand what LSA can and cannot achieve, we briefly discuss two common misconceptions, and two important limitations of LSA:

• LSA is not a keyword method. LSA induces the meaning of words and represents it in a semantic space. Meanings matter, not words per se. As an illustration, consider the sentence *The airplane rescued the hiker.* A sentence that is close in meaning (but by no means synonymous) is *The climber was evacuated by helicopter*. The cosine in the LSA space between these to sentences is .43, reflecting their semantic relatedness - in the absence of any word overlap. In contrast, a sentence with word overlap but a rather different meaning, such as, *The hiker saw the bear*, has a cosine of only .09 with the first sentence. This is an important feature of LSA, for it allows an application such as *Summary Street* to detect when students say more or less the same thing as the source text, but with different words.

• LSA does not reflect the co-occurrence among words but rather their semantic relatedness. LSA detects semantic relations even between two words that never co-occur. It is enough if two words are used similarly in different

contexts that LSA knows about for it to infer that they are semantically related. As a simple illustration, consider the singular and plural of words, which in actual texts are not highly correlated (if we talk about one object, we don't usually talk about many objects at the same time). LSA, however, infers that singulars and plurals are closely related, because they fulfill similar functions in different contexts. Thus, *computer* and *computers* have a cosine of .92 – they mean almost the same thing. This is not always the case, however, for sometimes singulars and plurals are used in distinct contexts: *sun* and *suns* have a cosine of only .50, presumably because we talk about our *sun* in an everyday context and *suns* in an astronomy context.

• LSA does not take into account syntax or word order. This precludes the use of LSA in some cases, especially when we are concerned with very brief texts. Word order and syntactic relations are crucial to understanding a sentence such as *Blood flows from the lungs to the left atrium* and LSA will not be able to distinguish this sentence from *Blood flows from the left atrium to the lungs.* Nevertheless, LSA can score an essay on the functioning of the circulatory system quite well, for writers who do not have a good understanding of the system use different combinations of words in their essays than writers who understand it correctly.

• LSA is not a reasoning system – it merely computes semantic similarity. LSA does not model precise, logical thinking and it cannot be used where analytic clarity is required. For instance, LSA cannot distinguish antonyms – the cosine between *up* and *down* is .87; thus for LSA *up* and *down* mean more or less the same thing. Which is of course true: the words are closely related and occur in similar contexts, but they differ in one crucial respect. Whenever meaning

evaluation depends exclusively on such a difference, LSA fails. This makes LSA a poor choice for scoring word problems in mathematics or physics (for instance, it cannot distinguish *four divided by two* from *two divided by four*). But LSA performs quite well on evaluating an essay about the heart and circulatory system in spite of this limitation. This is because LSA is highly sensitive to how people use words, even though its analytic powers are limited. LSA cannot tell per se that *Blood flows from the lungs to the right ventricle* is incorrect, but the writer who makes such an error will betray himself through his inexpert choice of words in the rest of his essay, and LSA will assign it a low score. LSA, when used judiciously, can be a very powerful educational tool. Designers of educational software as well as researchers who understand its power as well as its limitations can make good use of it, for both practical and theoretical purposes.

Using *Summary Street*

Summary Street is the culmination of three years' research of the LSA Research Group[1] in collaboration with two sixth-grade teachers, Cindy Matthews and Ronald Lamb, at Platt Middle School in Boulder. The history of its development from an earlier version, called *State the Essence*[2], is described in E. Kintsch et al. (2000). The current version was completely rewritten by Wade-Stein to address various technical problems as well as to present the feedback in a graphic form that would be easier to grasp than the verbally presented suggestions in *State the Essence*. The goals of *Summary Street* are the same as in its earlier incarnation: (1) to provide students extended, guided practice in expository writing and revision without adding to their teachers' workload; (2) to

integrate seamlessly into teachers' curricular plans and instruction; and (3) to make students responsible for improving their own writing and revision skills.

The introductory page of *Summary Street* not only provides information on how to navigate through the system, but also guidelines for writing a summary, together with links to more elaborated explanations of summarization strategies and examples thereof. These guidelines are based on cognitive research on the development of summarization strategies (Brown & Day, 1983; Rinehart, Stahl & Erickson, 1986). Users may access this information at any time, but for younger students it functions more effectively as a framework for classroom instruction.[3]

To use *Summary Street*, students first log in, choose the topic they will be summarizing, and request either a new text box or the last version of their summary previously stored in the database. Students may compose or copy-paste a summary into the text window (see Fig. 1). Students are instructed to first press the *Spell Check* button, which automatically saves their writing. Misspelled words, which are not recognized by LSA, are flagged so that students can correct them before proceeding. By pressing the *Feedback* button on the same screen students can request feedback on the content of their writing, which is displayed in the form shown in Figure 2. The students' texts are sent for analysis via a web connection to Knowledge Analysis Technologies, a company organized to support software applications using the LSA methodology (http://www.knowledge-technologies.com).

Figure 1 & Figure 2

A Java applet graphically displays the length of the summary and its upper and lower length limits as a vertical bar on the left side of the screen. The bar is green if the

length is appropriate or red otherwise. Content coverage for each section (i.e., major subtopics) of the source text is represented by horizontal green bars that either surpass or fall short of a vertical threshold line. In addition, *Summary Street* provides indicators of changes in content coverage between the current and previous round of feedback in the form of a dashed blue line bisecting each green bar. This allows students to see at a glance whether or not progress has been made. Students can access sections of the source text that are judged to be inadequately covered via hyperlinks provided by *Summary Street*.

Figure 3 shows a summary that has satisfied the content criteria, but is too long. At this point two additional tools become available. The *Redundancy Check* (Fig. 4) performs a sentence-by-sentence comparison to flag sentences that appear to have too much overlapping content. The *Relevance Check* (Fig. 5) compares each sentence in the summary to the source text to alert the user to sentences that have little or no relevance to the topic (both off-the-wall sentences and those that describe some low-level detail).

Figures 3, 4, 5

Finally, when students feel that they have improved their writing as much as they can, a *Format for Printing* button (see Fig.1*)* provides a double-spaced version of their text that can be used for proofreading and then shown to the teacher for evaluation.

The graphical display of feedback in *Summary Street* addresses several problems inherent in the earlier *State the Essence* system. First, because the actual number of words in the summary is not displayed, students are less fixated on single word additions or deletions to change the length of their summary. Second, the section feedback is finer-grained, enabling students to quickly gauge the effect of their latest change. By

discarding the overall point score used initially, we succeeded in focusing students' attention more on constructing the meaning of their texts, in order to balance content coverage and appropriate length, rather than minor differences in their point scores. In general, the graphic display of content feedback in *Summary Street* has been a marked success: spurred by the game-like aspect of getting the green bars past the black line and guided by the feedback that is sensitive to where a user is in the process, students willingly engage in repeated cycles of rewriting and revision.

A Look inside *Summary Street*

In order to implement *Summary Street*, the source text is first divided into sections, more or less equivalent in length, corresponding to its major subtopics. The texts used in our work consist of four to seven topic sections, and authors' subheadings were generally used to label each section. Next, an appropriate semantic space is chosen, or created. Currently, this choice can only be determined by testing the results with previously graded or trial summaries to see if appropriate cosines are returned.

Section coverage is determined by examining the cosine between the entire summary and the text of each section of the source text and by then comparing the cosine to a threshold for that section: if the threshold is met or exceeded, the section is considered adequately covered. An automated method is used to compute the thresholds, as follows:

First, each section of a source text is divided into sentences. Next, the cosine between each sentence and every other sentence in the section is computed. The sentence with the highest average cosine to all the other sentences in the section is designated as

the “typical sentence” for that section. After this process has been repeated for each section of the text, the typical sentences are combined to form a “typical summary”, which is used to generate the thresholds - the cosines between it and each section of the text. The typical summary represents the minimal content needed to pass the content criterion of all the sections. However, given the minimum length constraint assigned by the teacher, students’ summaries should be more elaborated, and in fact, often err in the direction of too much length. The rather strict maximum length, on the other hand, curtails the tendency of some students to compose summaries of sentences copied directly from the text. Thus, students are guided towards using their own words, to avoid writing summaries that are much too long. Although it appears to the user that the threshold, represented by a vertical black line, is the same for each section of the text, in fact the cosines vary across sections. The green bars represent the percentage of the threshold that has been covered by the student’s summary (i.e., the cosine between the student’s summary and the corresponding section of the text, divided by the threshold value).

As previously described, once the student has covered the content adequately, two other options become available: The *Redundancy* and *Relevance Checks* can be used to help reduce the length of a too-long summary or simply to improve it. Both are useful for identifying possible problematic sentences that can be eliminated or rewritten. The *Redundancy Check*, shown in Figure 4, opens up a new window and then computes the cosine between each pair of sentences in a student’s summary. Those sentence pairs whose cosine exceeds a redundancy threshold are flagged with similar colors and numbers. Note that changes to a summary can only be made in the text box.

The redundancy threshold was defined to be two standard deviations above the average sentence-to-sentence cosine in the source text. Accordingly, each text may have its own redundancy threshold, but in practice one threshold has been used so far for all of the texts. This threshold is adequate, but far from perfect, with the result that too many sentence pairs are flagged as redundant. Some of the erroneous flagging is unavoidable, as many pairs of words have high cosines with one another but are not redundant. For example, antonyms, typically occur in very similar contexts and are thus closely related, but are hardly redundant (e.g., the cosine for *when you are good you go to heaven* and *when you are bad you go to hell* is *.84*).

The *Relevance Check*, displayed in Figure 5, operates on a sentence-by-sentence basis, computing a relevance index and flagging (in red color) any sentence whose index falls below the threshold. The relevance index consists of the sum of the maximum cosine between the sentence and each of the sections of the text, together with the vector length of the sentence. Both the cosine and the vector length are important to assess relevance, because the cosine only measures the semantic relatedness of two documents, while the vector length shows how much information the documents contain. Hence, in extreme situations a passage of text may contain a lot of information, but most of it is off topic; or conversely, the writing may be closely related to the topic but conveys little information.

The threshold for the *Relevance Check* currently used in the system was determined in an ad hoc fashion by examining summaries from previous studies and adjusting the threshold that flags the majority of irrelevant sentences. The *Redundancy* and the *Relevance Checks* are useful tools for dealing with problems in summary writing,

however neither of these checks returns completely optimal results. The disadvantage is that a user may disagree about some of the flagged sentences: for example, repeating an idea may be necessary to maintain coherence, or the writer may wish to keep a salient, though detailed piece of information (e.g., that Mayan pyramids were composed of narrow, steep steps in the example of the *Relevance Check* shown in Fig. 6); or the sentence may indeed be an important idea (e.g., the problems between present-day Mayan villages and the government). However, the advantage of these tools is that users must judge for themselves whether a flagged sentence belongs in the summary. In so doing they must reconsider what they have written and often compare it to the source text. If the user wishes to retain some or all of the information, he/she may seek briefer ways to express it, perhaps through sentence combining or generalizations. For example, it would be a good idea to combine the two sentences flagged as redundant in Figure 5 (about Mayan books) into a single sentence or to delete the second one entirely. Thus, the sentence flags, even when inappropriate, can promote prolonged deliberation with the meaning the user is trying to construct, more so than would happen with feedback that tells the user exactly what to do. Indeed, this feature is a pedagogical strength of a system such as *Summary Street*. Instead of having one's responses judged by an authority figure (whether human or computer), as right or wrong, the student engages in a non-threatening dialog with the computer, and is invited to challenge, even disagree with the advice given.

Summary Street in the Classroom

Method

In order to obtain formative evaluation with external validity, *Summary Street* was incorporated into the curriculum of two team-taught sixth-grade classes at Platt Middle School (Boulder, CO). The classes, taught by CM and RL, consisted of 60 students altogether. The trial took place during a curricular unit on Sources of Energy. Students chose two texts to summarize, one on a renewable and one on a nonrenewable energy source. The texts were composed of four to seven sections and were about 1,900 words long on average (range = 1,148 and 2,367 words). The purpose of summary writing in this case was to gain sufficient background knowledge about a particular topic in order to present the topic orally to other classmates in a small group teaching format. The two summaries were written on two successive weeks as homework. During a one-and-a-half-hour session in the computer lab one class of 30 students revised their summaries using the regular *Summary Street* program (feedback condition); the other 30 students used a visually similar interface that provided only information about spelling errors and the vertical length indicator but no content feedback (no-feedback condition). Each student wrote one summary with feedback and one without, with the order counterbalanced. The computer kept track of the amount of time each student spent working on a summary in both the feedback and the no-feedback condition.

The students' summaries were graded by their teachers, who were blind to the experimental condition in which they were written. First, they provided a score for each section of the summary, taking only the content adequacy into consideration, using a 3-point scale (0 = no information about this section topic, 1 = some information but not

enough, 3 = adequate section coverage). Next, the teachers provided a holistic score of the quality of the summary, evaluating other factors in addition to the content alone (e.g., style, coherence, organization, and mechanics). This scale consisted of 5 - 0 points, corresponding to the conventionally used A – F grades.

Results

Time on Task

Figure 6 confirms our informal observations during the trial: students who received full content feedback from *Summary Street* spent more than twice as long working on their summaries as students who received no content feedback (M =72 vs. 33 minutes, $t_{43} = 5.88$, $p < .0001$)[4]. In general, when only minimal (length and spelling) feedback was provided (no feedback condition), students were not motivated to continue revising and tended to work only until their summaries were of the appropriate length. In contrast, getting the green section bars to pass the threshold line proved to be an important motivator for students to keep revising until the content of all sections was adequately covered.

Figure 6

Content and Quality Scores

This extra time on task resulted in improved content scores for students in the feedback condition compared to those receiving no feedback. The results of t-tests are shown in Figure 7. Summaries composed with full *Summary Street* feedback received significantly higher content scores than those for which only feedback on length and spelling was provided, (M = 1.29 vs. 2.02, $t_{45} = 3.91$, $p = .003$).

Figure 7

The overall quality scores likewise showed a small but significant difference between the feedback and no-feedback conditions (M = 3.19 vs. 2.87, $t_{45} = 2.16$, p = .036). A repeated-measures ANCOVA was performed on the quality and content difference scores in the two conditions to determine whether the quality difference was distinct from the content difference. After controlling for content, the quality difference between feedback and no-feedback conditions was no longer significant ($t_{44} = 0.36$, p = .72). This result reflects the difficulty teachers had in trying to consider content separately from other quality characteristics as they graded the summaries. Problems with coherence, organization and sentence structure are generally symptomatic of a writer's inadequate grasp of the material, of "fuzzy thinking", whereas well articulated ideas reflect good comprehension .

Interaction of Text Difficulty and Content Feedback

Receiving feedback on the content of a summary apparently does help students write better summaries. But is it uniformly helpful for all the texts they summarized? To address this question we used two analyses that examined the interaction of feedback and text difficulty. The first analysis examined the improvement in teachers' content scores that resulted from receiving feedback (feedback scores minus no-feedback scores), we call this difference score the *feedback effect*. To calculate text difficulty we used students' scores in the condition with no content feedback. In Figure 8 the feedback effect is plotted against this measure of text difficulty. As seen in Figure 8, there is a significant negative correlation (r = -0.83, p = .003) between the difficulty of the text and the feedback effect - the improvement shown when content feedback was provided. Thus, for easy topics (summaries that received higher grades without feedback), *Summary Street*

feedback did not improve the summary content. In contrast, there was a substantial feedback effect for difficult topics, with moderate improvement for topics between the two extremes.

Figure 8

This result is not limited to the texts on energy sources used in this trial. Two earlier trials of *Summary Street* in the classroom yielded quite similar patterns of results. Both trials took place with sixth-grade students working on different curricular topics: Ancient Civilizations (Maya, Aztec & Inca), and the Human Circulatory System (Heart & Lungs). In both studies, using *Summary Street* resulted in a significant advantage when students summarized the more difficult texts. However, there were no differences in their summaries of the easier texts regardless of whether the students received feedback or not. We can thus conclude with some confidence that the kind of support that *Summary Street* provides is most effective when students are faced with a challenging text.

For the second analysis, LSA cosines in the no-feedback condition were used instead of teachers' grades to provide a measure of text difficulty. The correlation is highly significant ($r = -.748$, $p = .013$), and the pattern is similar to that found in the previous analysis of content scores provided by the teachers: the effectiveness of feedback on revising increases as a function of the difficulty of the text. Moreover, the similar trend exhibited in both analyses, and the fact that the correlation between content grades and cosines was high ($r = 0.70$), $p = .02$) supports our conclusion that teacher judgments of the content adequacy of students' summaries and the average no-feedback cosine work equally well as measures of text difficulty. Thus, either measure may be helpful for determining which texts are appropriate for the students.

Individual Differences

If content feedback is more beneficial when summarizing a difficult text, might its effectiveness also vary across individual learners? To pursue this question we asked teachers to classify their students according to their verbal ability on a scale of 1 – 3 (low, average, or high). The pattern shown in Figure 9 suggests that the feedback was most useful for learners of moderate ability: high ability students probably did not need the support to write competent summaries, while students with low verbal ability may have found the task of summarizing rather difficult texts too challenging. A repeated measures ANCOVA on these results failed to reach significance ($F_{2,33} = 1.99$, $p = .15$), however, it is a question that should be pursued in future research.

Figure 9

Discussion and Conclusions

The results of classroom trials with *Summary Street* seem quite encouraging and suggest that the tool could be beneficial to a broader population of learners than the sixth graders who participated in these trials. The simple graphic display of the content feedback is easy to grasp and motivates students to work longer on their summaries than would otherwise be the case. The extra time spent on the task resulted in better summaries, according to the teachers' judgment of their content and their overall quality. And as in studies of LSA essay grading (Foltz, Gilliam & Kendall, 2000; Landauer, Laham, & Foltz, 2000), LSA cosines in this study correlated well with teachers' grades of content adequacy for summaries written in the no-feedback condition.

The results also suggest answers to two of the questions posed at the beginning of this paper: when is Summary Street effective and for whom? The classroom trial clearly demonstrated that the usefulness of the feedback is tied to the difficulty of the text. This result supports the findings of earlier trials with *Summary Street* (E. Kintsch et al., 2000), showing that for easy topics students can write an adequate summary without the support from *Summary Street*. The feedback is most effective when a text presents a challenge, but a challenge for whom?

Although differences among lower and higher verbal ability students were not reliable in this experiment, the patterns seen in Figure 8 and Figure 9 suggest an interaction between text difficulty and learner ability. That is, the kind of feedback provided by *Summary Street* may be most beneficial when a text is moderately challenging for a particular learner. As in many situations, learners may become overly frustrated with a text or task that is too far beyond their current level of knowledge and ability and learn little; on the other hand, the support from *Summary Street* is not needed for a text that learners can easily summarize on their own. However, we believe that the needs of many different learners can be met by manipulating the difficulty of the instructional texts to the appropriate level of learner ability (cf. also McNamara, Kintsch, Kintsch & Songer, 1996; and Wolfe et al., 1998). The real value of a tool like *Summary Street* may be that it supplies the kind of guidance students need to work successfully with more challenging materials, texts that are somewhat beyond their current level of comfort. *Summary Street* may also be helpful in diagnosing what that level is for a particular student.

We should note that several students with mild learning disabilities participated in the experimental session. Although they did not succeed in completing the task of covering all the topics in the text, they nevertheless continued working with full engagement throughout the hour-and a-half-long session and greatly improved their original summaries. These observations further support our hope that *Summary Street* may benefit a broad range of students.

An issue that should be considered in more detail is what aspects of the sixth graders' summaries were improved by *Summary Street* feedback, and what kinds of writing problems are not addressed. Primarily, the feedback focuses attention on getting the right information from the source text and on stating it more briefly. Trying to balance the conflicting demands of content adequacy and conciseness is a difficult task for these students, who have little experience to draw on. These sixth graders are just beginning to deal with more difficult reading on unfamiliar topics, texts that cannot be understood well by depending on passive, fairly automatic processing. Instead, higher order, quasi-problem-solving strategies, such as summarizing, are required to comprehend and learn from such materials. Most of the students in the present study appeared to write summaries characteristic of copy-delete writing, described by Brown and her colleagues as the earliest strategy in the developmental trajectory (e.g., Brown, Bransford, Campione & Ferrara, 1983; Brown & Day, 1983). A significant problem for novice writers is to identify the main ideas (e.g., Winograd, 1984), and this is where *Summary Street* feedback is especially helpful. In addition, it also teaches them to avoid excessive repetition, low-level details and totally off-topic statements. Although it takes more than proper content selection to write a good summary, we agree with the teachers

of these sixth-grade classes, that this is as a good starting point not only for summarizing but for learning novel content in general. Moreover, *Summary Street* feedback does not label particular sentences as right or wrong, but requires the user's judgment and reflectivity: the redundancy and relevance checks merely make suggestions about potential problem sentences. In some cases the user may want to delete a sentence or find a more concise way to express the meaning; but some of the suggestions are not appropriate, to the delight of students who love it when they can show that the computer "is wrong".

Future Directions

Summary Street does not include a plagiarism check at this point, although it would be easy enough to implement one. Instead, together with the teachers, we decided to use a fairly strict length constraint to encourage students to write in their own words. Students soon learn that they cannot cover all the text topics by simply lifting a lot of sentences directly from the text. This insight, we believe, serves to make them more aware of the need for better summarization strategies, techniques like combining ideas into a single sentence and ultimately generalizing across a host of details. *Summary Street* offers guidelines on how to write a summary on its home page, but the task of instructing higher order comprehension and summarization strategies is taught more effectively by teachers through direct instruction, modeling and discussion of specific problems. Although students can use *Summary Street* on their own, it will work best if grounded in classroom instruction and used to supplement and enhance the teacher's curricular goals: it's primary purpose is to give students more practice, with guidance, in attentive, active reading and extended writing.

It should be noted that LSA does not directly assess whether statements in a student's summary are correct or incorrect. However, as noted earlier, writers who make false statements tend to betray their lack of knowledge by their choice of words – something LSA is quite sensitive to. Thus, factual errors are usually indicated by low cosines on the essay as a whole. A 200-word summary or essay is generally sufficient to reveal the writer's knowledge or lack thereof.

Many aspects of good writing, such as organization, coherence, voice and style are also not addressed by the system, likewise problems with sentence structure, spelling advice, punctuation, and the like. Some of these features may be desirable and entirely feasible additions to the system, according to our post-trial interviews with teachers and students (especially, spelling and grammar advice). Organization and coherence problems are often resolved when one has fully comprehended what one is writing about, yet LSA-based feedback may be helpful here as well. In general, we believe that the role of the tutor is to provide users with guidance that will enable them to use their own intelligence. Once students have done all the revising they can, the teacher's comments remain as important as always.

In improving the system, we shall continue to avoid providing students with correct answers. *Summary Street* occupies an intermediate position along the continuum from fully intelligent tutoring systems to completely non-intelligent, skill-and-drill tutors, though certainly both types of tutors can assist learning in important ways. Unlike intelligent tutors, *Summary Street* does not rely on models of student understanding to guide performance. And our goal is quite different as well: rather than trying to keep the learner on an error-free learning path, *Summary Street* aims to provide just enough

guidance to help learners debug and revise problems with their comprehension and writing on their own. On the other hand, the intelligence in *Summary Street* - its ability to evaluate meaning - places it far beyond the capability of game-like tutors for practicing isolated math or reading skills.

Support for practicing summary writing is only one of several potential uses for *Summary Street*. Indeed, teachers should be encouraged to adapt the technology to fit their own goals. For example, instead of having students hand in a finished summary, teachers may use the cosine scores for subtopics of a text to assess how well students are comprehending new material, and to pinpoint the gaps in their understanding where more explanation is needed. Of course students may also use the system to assess their own understanding, as an alternative to the often misleading end-of-chapter "comprehension questions". Instead of summarizing a single text, teachers may ask students to summarize selectively across two or more texts on a particular topic or to compare and contrast conflicting viewpoints on an issue (e.g., Wolfe & Goldman, in prep; Wolfe, M. B., Goldman, S. R., Mayfield, C., Meyerson, P. M., & Bloome, D. M., 2000). *Summary Street* or a comparable system need not be limited to the task of writing a summary. Potentially it can be extended to evaluate short or longer essays written by students of any age, as, for example, in The Intelligent Essay Grader (Foltz, Gilliam, & Kendall, 2000; Landauer, Laham, & Foltz, 2000). Going beyond individual classrooms, *Summary Street* could be used more broadly to assess literacy skills at a school or even district-wide level. In sum, we regard *Summary Street* not as a single successful application of LSA, but rather as general purpose educational tool which will be useful in a wide variety of tutoring and self-assessment applications.

References

Baumann, J. F. (1984). The effectiveness of a direct instruction paradigm for teaching main idea comprehension. *Reading Research Quarterly*, *20*, 93-115.

Bean, T. W., and Steenwyk, F. L. (1984). The effect of three forms of summarization instruction on sixth-graders' summary writing and comprehension. *Journal of Reading Behavior*, *16*, 287-306.

Bransford, J. D., Brown, A. L., and Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*. National Research Council Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Brown, A. L., Bransford, J. D., Ferrara, R. A., and Campione, J. C. (1983). Learning, remembering, and understanding. In J. Flavell and E. M. Markman (Eds.), *Handbook of child psychology* (4th ed.). *Cognitive development* (Vol. 3, pp. 515-629). New York: Wiley.

Brown, A. L., and Day, J. D., (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior*, *22*, 1-14.

Brown, A. L., Day, J. D., and Jones, R. S. (1983). The development of plans for summarizing texts. *Child Development*, *54*, 968-979.

Brown, A. L., and Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, learning, and*

instruction: Essays in honor of Robert Glaser (pp. 393-451). Hillsdale, NJ: Erlbaum.

Brown, A. L., and Smiley, S. S. (1977). Rating the importance of structural units of prose passages: A problem of metacognitive development. *Child Development, 48,* 1-8.

Casazza, M. E. (19192). Teaching summary writing to enhance comprehension. *Reading Today, 9,* 28.

Foltz, P. W., Gilliam, S., and Kendall, S. (2000). Supporting content-based feedback in on-line writing evaluation with LSA. *Interactive Learning Environments, 8 (20),* 111-127.

Gajria, M., and Salvia, J. (1992). The effects of summarization instruction on text comprehension of students with learning disabilities. *Exceptional Children, 58,* 508-516.

Jitendra, A. K., Cole, C. L., Hoppes, M. K., and Wilson, B. (1998). Effects of a direct instruction main idea summarization program and self-monitoring on reading comprehension of middle school students with learning disabilities. *Reading and Writing Quarterly, 14,* 379-396.

Kintsch, E., and Kintsch, W. (1997). Learning from text. In E. D. DeCorte, and F. E. Weinert (Eds.): International encyclopedia of developmental and instructional psychology (pp. 519-524). Amsterdam: Elsevier.

Kintsch, E., Steinhart, D., Stahl, G., LSA Research Group, Matthews, C., and Lamb, R. (2000). Developing summarization skills through the use of LSA-based feedback. *Interactive Learning Environments 8,* 87-109.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. NY: Cambridge University Press.

Landauer, T. K (1998). Learning and representing verbal meaning: The latent semantic analysis theory. *Current Directions in Psychological Science, 7,* 161-164

Landauer, T. K, and Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211-240.

Landauer, T. K., Foltz, P. W., and Laham, D. (1998). An introduction to Latent Semantic analysis [Special Issue]. *Discourse Processes, 25,* 259-284.

Landauer, T. K., Laham, D., and Foltz, P. W. (Sept/Oct. 2000). The intelligent essay assessor. *IEEE Intelligent systems: Trends and Controversies,* 27-31.

Malone, L. D., and Mastropieri, M. A. (1992). Reading comprehension instruction: Summarization and self-monitoring training for students with learning disabillities. *Exceptional Children, 58,* 270-279.

Mannes, S. M., and Kintsch, W. (1987). Knowledge organization and text organization. *Cognition and Instruction, 4,* 91-115.

McNamara, D. S., Kintsch, E., Kintsch, W., and Songer, N. B. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14,* 1-43.

McNeil, J., and Donant, L. (1982). Summarization strategy for improving reading comprehension. In J. A. Niles and L. A. Harris (Eds.*), New inquiries in reading research and instruction* (pp. 215-219). Rochester, NY: National Reading Conference.

Moore, S. R. (1995). Focus on research questions for research into reading-writing relationships and text structure knowledge. *Language Arts, 72,* 598-606.

National Energy Education Development Project. (1999*). Secondary Energy Fact Sheet: Solar energy*. [Brochure]. Reston, VA: Author.

Palincsar, A. S., and Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition and Instruction, 12,* 117-175.

Paris, S. G., Wasik, B. A., and Turner, J. C. (1991). The development of strategic readers. In R. Barr, M. L. Kamil, P. B. Mosenthal, P. D. Pearson (Eds.): *Handbook of reading research,* Vol. 2 (pp. 609-640). New York, NY: Longman.

Reis, R., and Leone, P. E. (1987). Teaching reading and study skills to mildly handicapped learners. *The Pointer, 31,* 41-43.

Rinehart, S. D., Stahl, S. A., and Erickson, L. G. (1986). Some effects of summarization training on reading and studying. *Reading Research Quarterly, 21,* 422-437.

Ross, S. M., and DiVesta, F. J. (1976). Oral summary as a review strategy for enhancing recall of textual material. *Journal of Educational Psychology, 68,* 689-695.

Taylor, B. M. (1982). Text structure and children's comprehension and memory for expository material. *Journal of Educational Psychology, 74,* 323-340.

Taylor, B. M., and Beach, R. W. (1984). The effect of text structure instruction on middle-grade students' comprehension and production of expository text. *Reading Research Quarterly, 19,* 132-146.

Tierney, R. J., and Shanahan, T. (1991). Research on the reading-writing relationship: Interactions, transactions, and outcomes. In R. Barr, M. L. Kamil, P. B. Mosenthal, P.

D. Pearson (Eds.): *Handbook of reading research*, Vol. 2 (pp. 246-280). New York, NY: Longman.

Uttero, D. A. (1988). Activating comprehension through cooperative learning. *The Reading Teacher*, *41*, 390-395.

van Dijk, T. A., and Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

Winograd, P. N. (1984). Strategic difficulties in summarizing texts. *Reading Research Quarterly*, *19*, 404-425.

Weisberg, R., and Balajthy, E. (1990). Development of disabled readers' metacomprehension ability through summarization training using expository text: Results of three studies. *Reading, Writing, and Learning Disabilities*, *6*, 117-136.

Wolfe, M. B., Schreiner, M. E., Rehder, B., Laham, D., Foltz, P. W., Kintsch, W., and Landauer, T. K. (1998). Learning from text: Matching readers and texts by Latent Semantic Analysis. *Discourse Processes*, *25*, 309-336.

Wolfe, M. B. W., Goldman, S. R., Mayfield, C., Meyerson, P. M., & Bloome, D. M. (2000, April). Latent Semantic Analysis as a tool for assessing information integration and synthesis across multiple texts. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

Wolfe, M. B. W., Goldman, S. R., Mayfield, C., Meyerson, P. M., and Bloome, D. M. (2000, July). Middle School students' processing of multiple accounts of an historical event. Paper presented at the meeting of the Society for Text and Discourse, Lyon, France.

Author Note

The research described here was funded by a grant from the James S. McDonnell Foundation and is based on the doctoral dissertation of David Steinhart (now David Wade-Stein). The authors gratefully acknowledge the assistance of Walter Kintsch, Thomas Landauer, Richard Olson, Gerhard Fischer, Gerry Stahl, Cindy Matthews, Ronald Lamb, and students of several classes of sixth-grade students at Platt Middle School in Boulder, Colorado.

Footnotes

[1]Members of the LSA Research Group are: W. Kintsch, T. Landauer, R. A. de Paula, E. Kintsch, D. Laham, M. E. Schreiner, G. Stahl, & D. Wade-Stein.

[2]*State the Essence* was written and maintained by G. Stahl, R. A. de Paula, D. Laham, M. Jones, M. Schreiner, & D. Wade-Stein.

[3]Users may try out the system by accessing the project's homepage at http://colit.org and choosing the "Demonstrations" link or by logging on as "Guest". The website also provides a set of texts and sample summaries.

[4]The degrees of freedom here reflect missing data for several students.

Figure Captions

Figure 1. *Summary Street* textbox window showing a saved summary about the Maya and the various option buttons.

Figure 2. Feedback window showing length and section coverage feedback.

Figure 3. Summary street textbox window showing revised Maya summary.

Figure 4. Window showing successful content coverage feedback.

Figure 5. Window showing feedback from *Redundancy Check*.

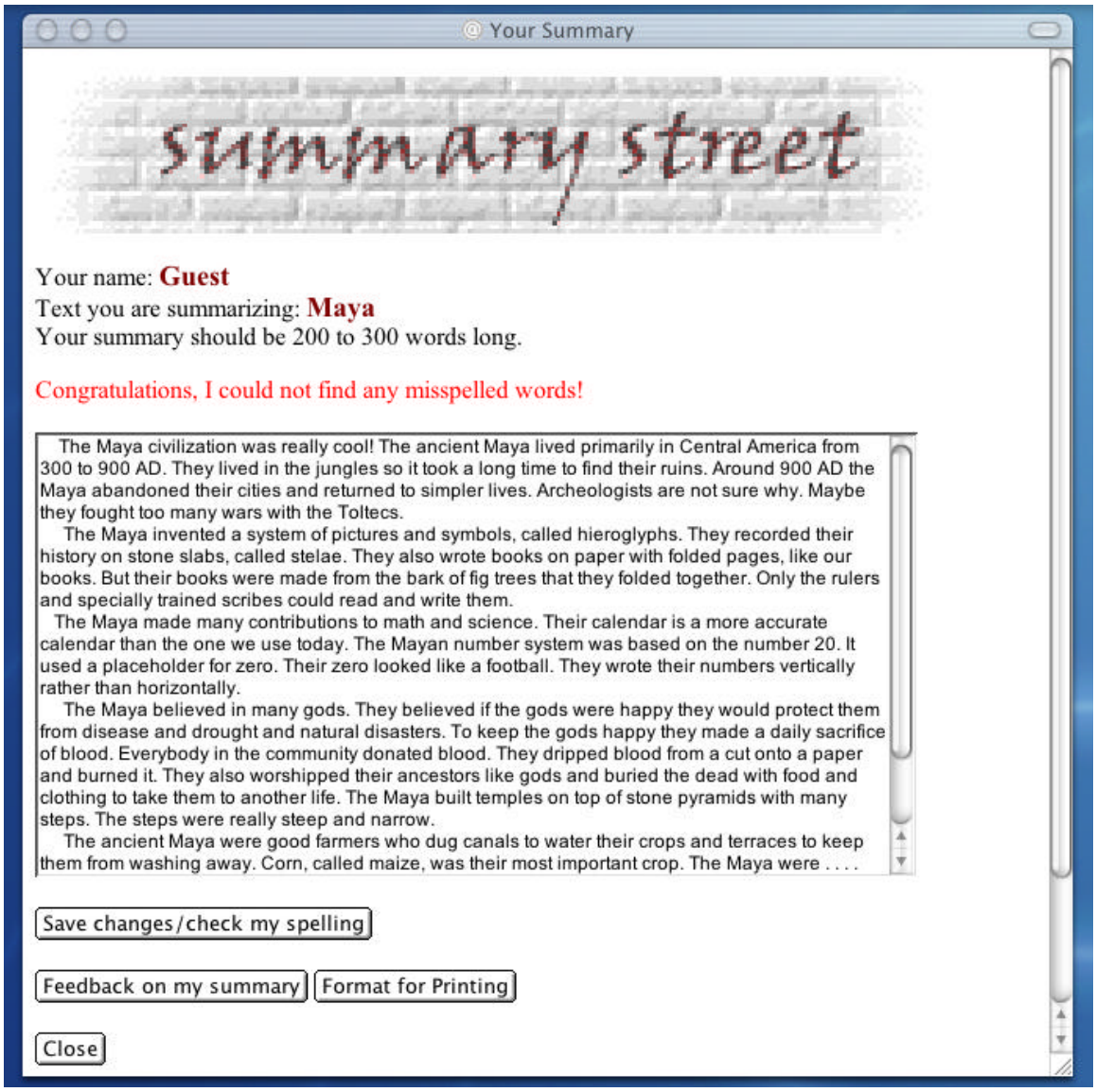
Figure 6. Window showing feedback from *Relevance Check*.

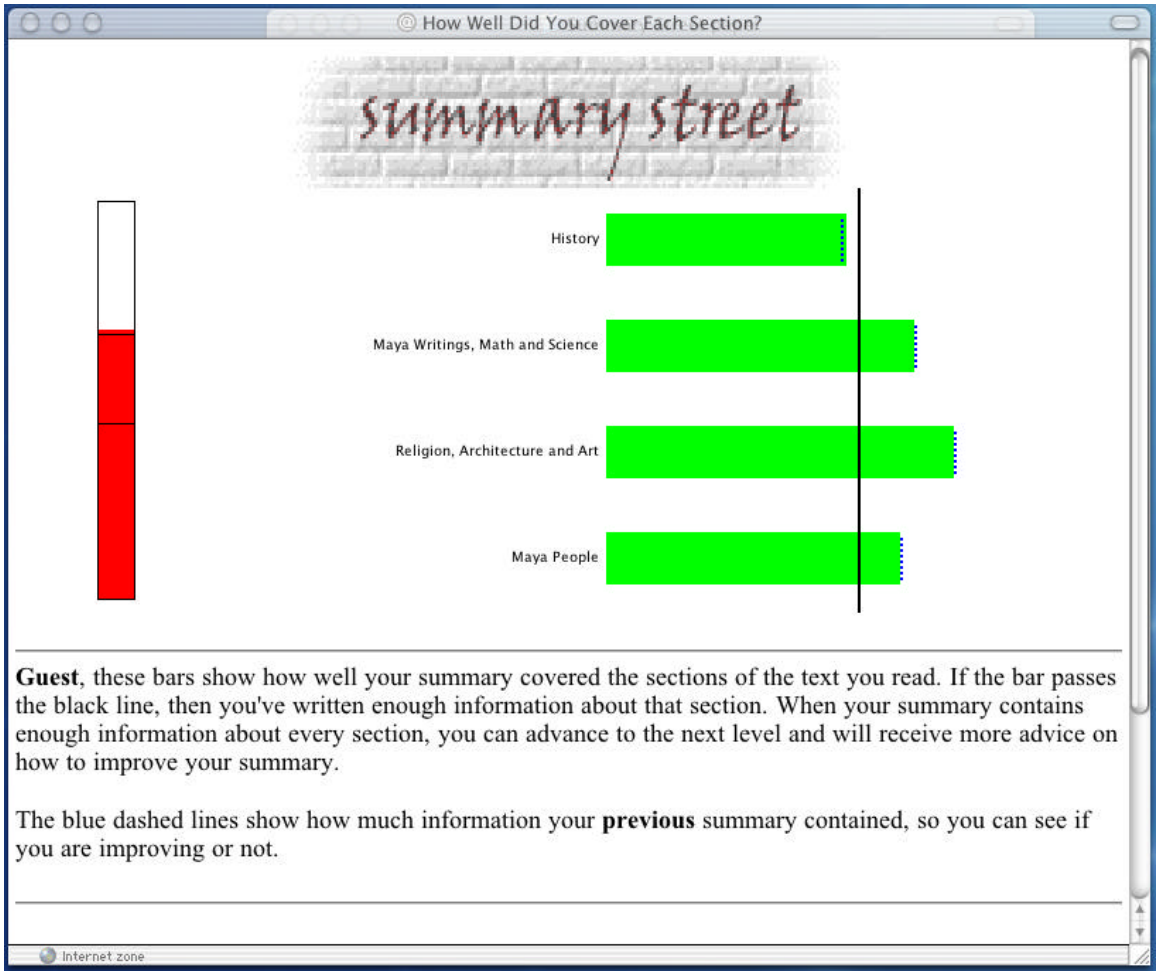
Figure 7. Mean time on task by condition.

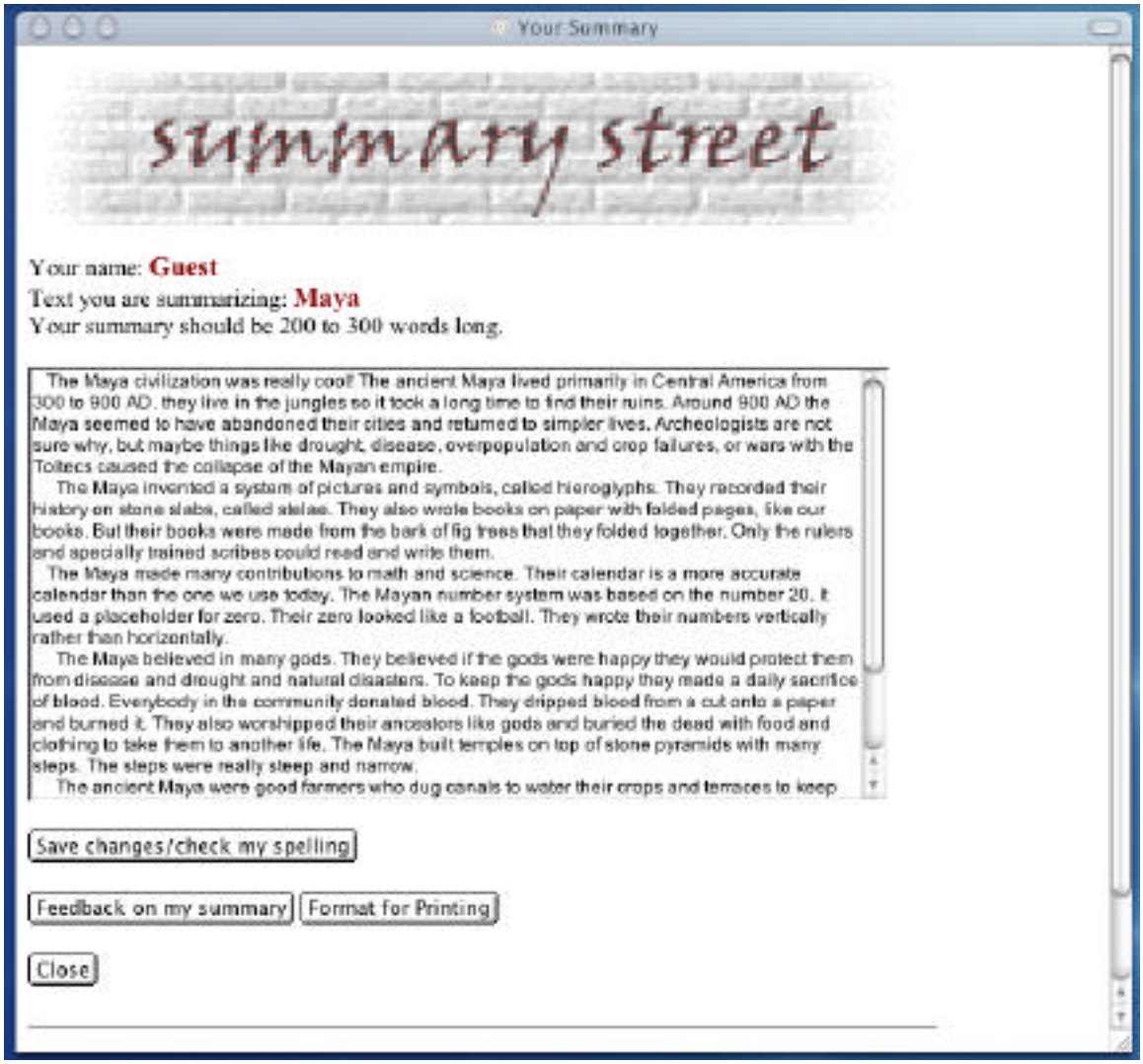
Figure 8. Mean content score by condition.

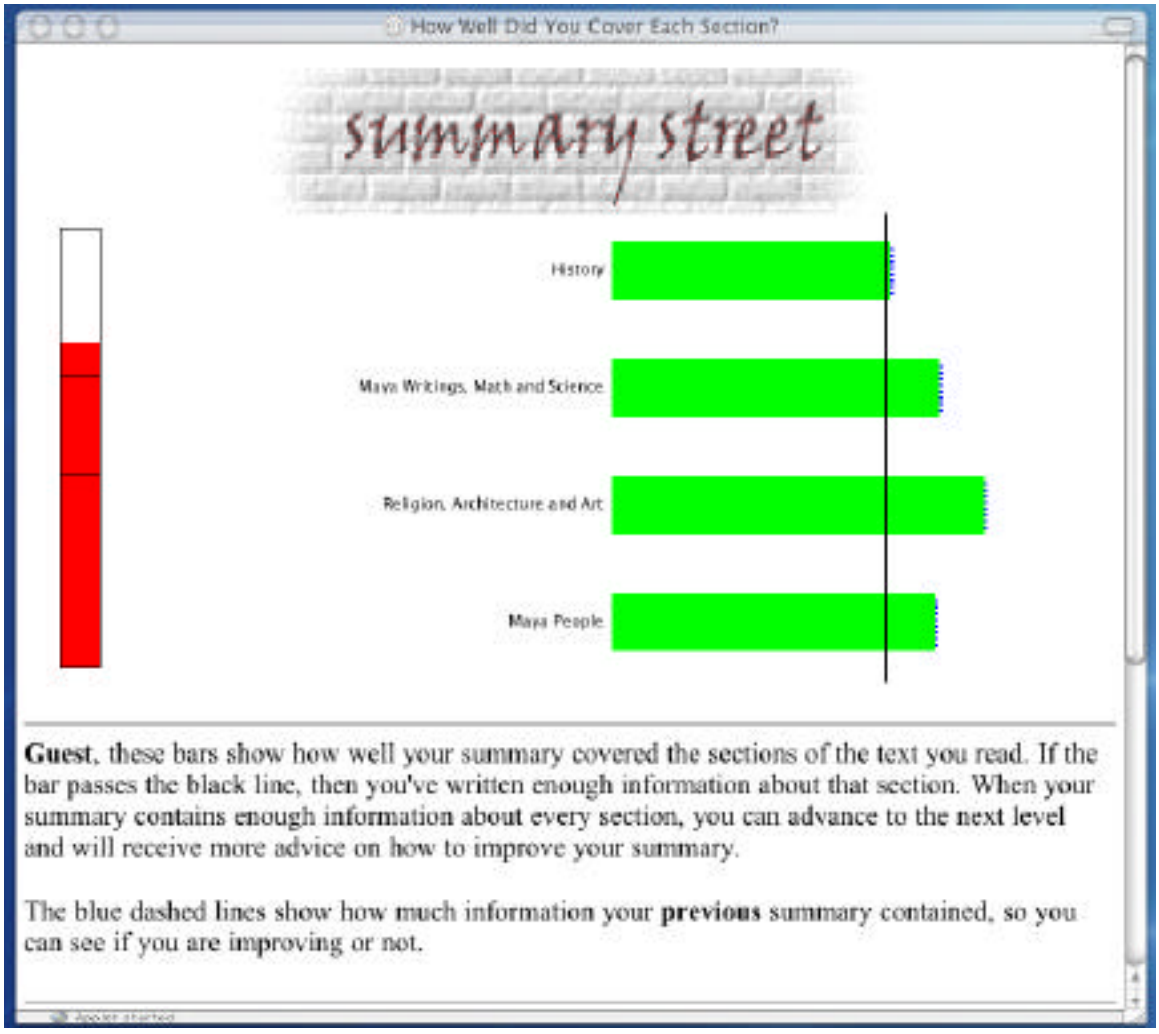
Figure 9. Feedback effect (improvement in content score for students receiving feedback) vs. text difficulty (average content score of students summarizing this topic without feedback); $r = -0.83$, $p = .003$. The negative effect in the case of one topic was not significant.

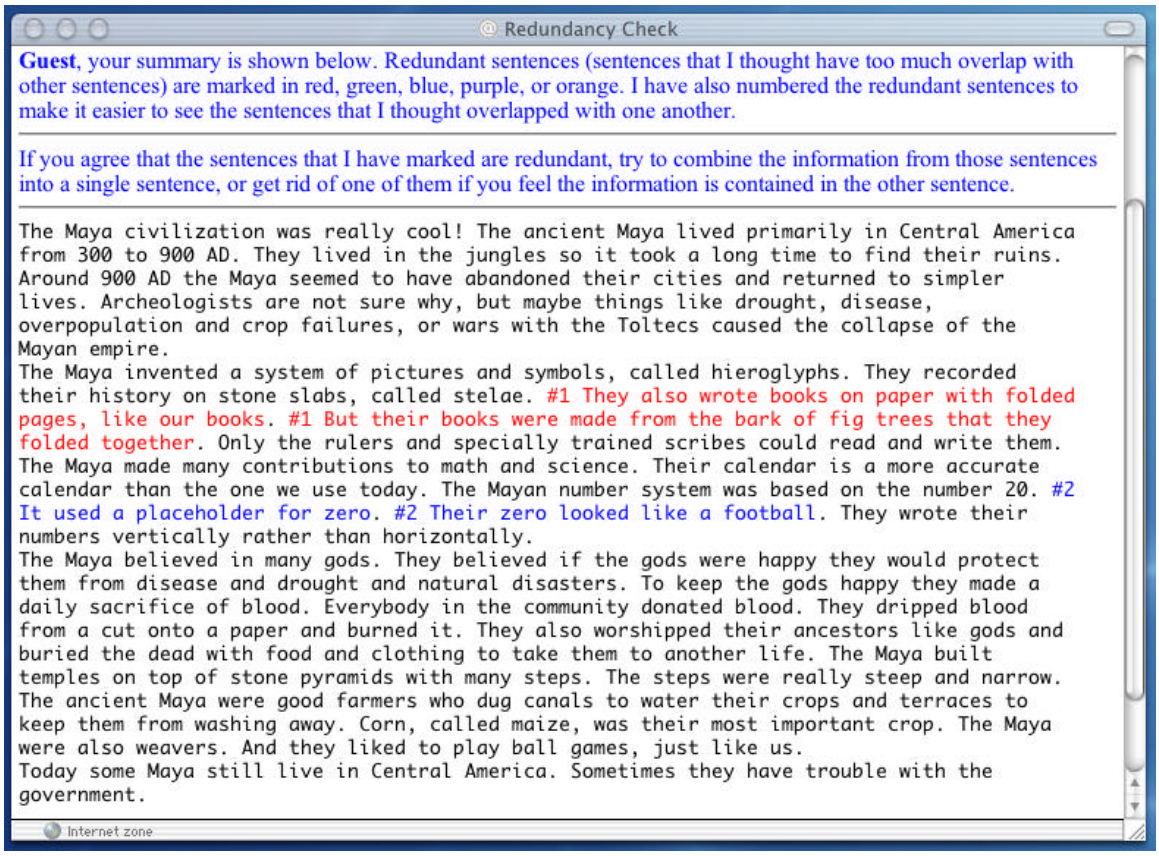
Figure 10. Effect of *Summary Street* feedback as a function of students' verbal ability.

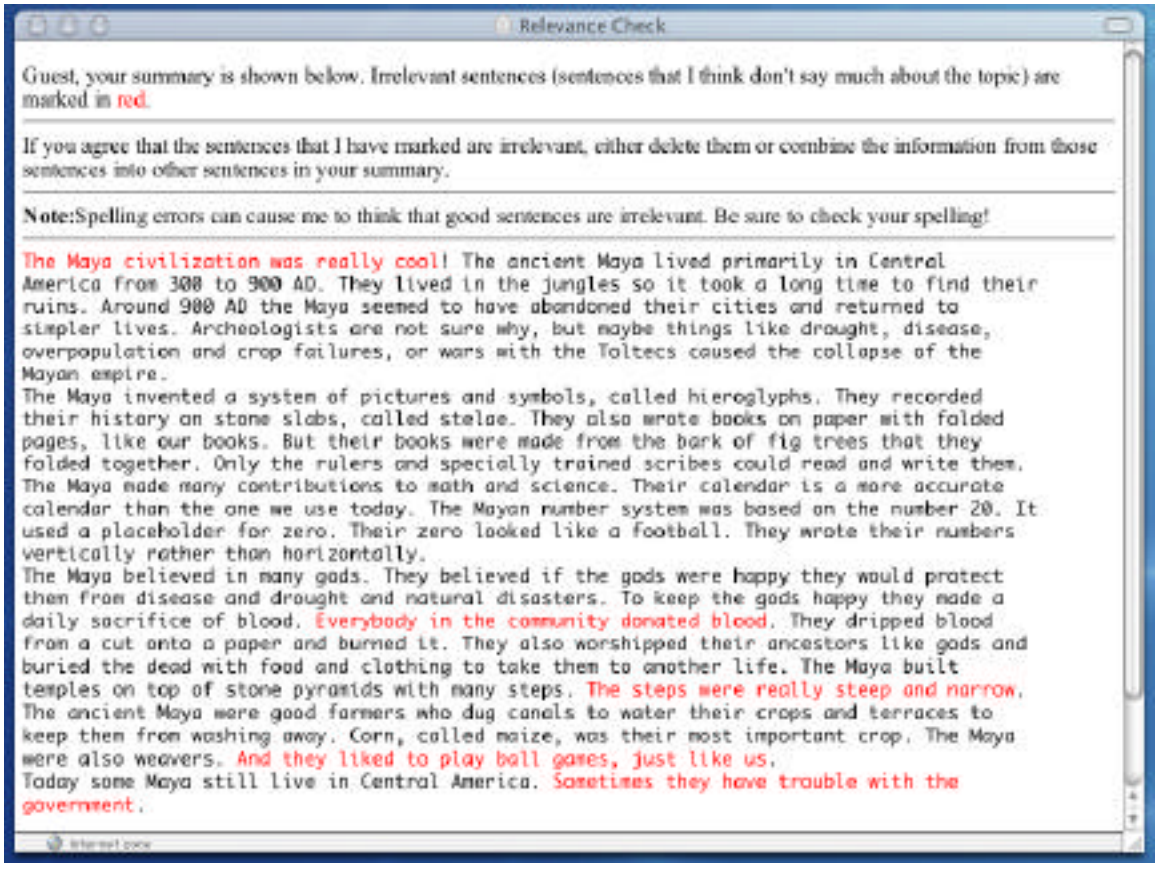

Your Summary
summary street
Your name: Guest
Text you are summarizing: Maya
Your summary should be 200 to 300 words long.
Congratulations, I could not find any misspelled words!
The Maya civilization was really cool! The ancient Maya lived primarily in Central America from 300 to 900 AD. They lived in the jungles so it took a long time to find their ruins. Around 900 AD the Maya abandoned their cities and returned to simpler lives. Archeologists are not sure why. Maybe they fought too many wars with the Toltecs.
The Maya invented a system of pictures and symbols, called hieroglyphs. They recorded their history on stone slabs, called stelae. They also wrote books on paper with folded pages, like our books. But their books were made from the bark of fig trees that they folded together. Only the rulers and specially trained scribes could read and write them.
The Maya made many contributions to math and science. Their calendar is a more accurate calendar than the one we use today. The Mayan number system was based on the number 20. It used a placeholder for zero. Their zero looked like a football. They wrote their numbers vertically rather than horizontally.
The Maya believed in many gods. They believed if the gods were happy they would protect them from disease and drought and natural disasters. To keep the gods happy they made a daily sacrifice of blood. Everybody in the community donated blood. They dripped blood from a cut onto a paper and burned it. They also worshipped their ancestors like gods and buried the dead with food and clothing to take them to another life. The Maya built temples on top of stone pyramids with many steps. The steps were really steep and narrow.
The ancient Maya were good farmers who dug canals to water their crops and terraces to keep them from washing away. Corn, called maize, was their most important crop. The Maya were . . . .
Save changes/check my spelling
Feedback on my summary
Format for Printing
Close

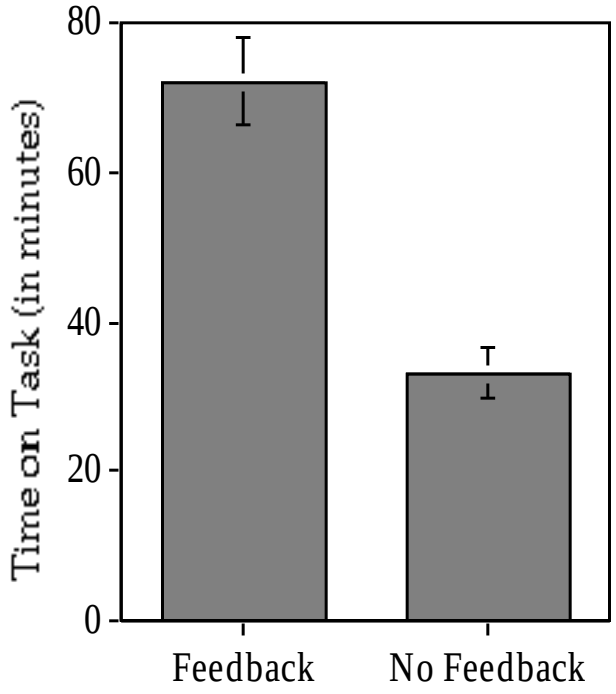

How Well Did You Cover Each Section?
summary street
History
Maya Writings, Math and Science
Religion, Architecture and Art
Maya People
Guest, these bars show how well your summary covered the sections of the text you read. If the bar passes the black line, then you've written enough information about that section. When your summary contains enough information about every section, you can advance to the next level and will receive more advice on how to improve your summary.
The blue dashed lines show how much information your previous summary contained, so you can see if you are improving or not.
Internet zone

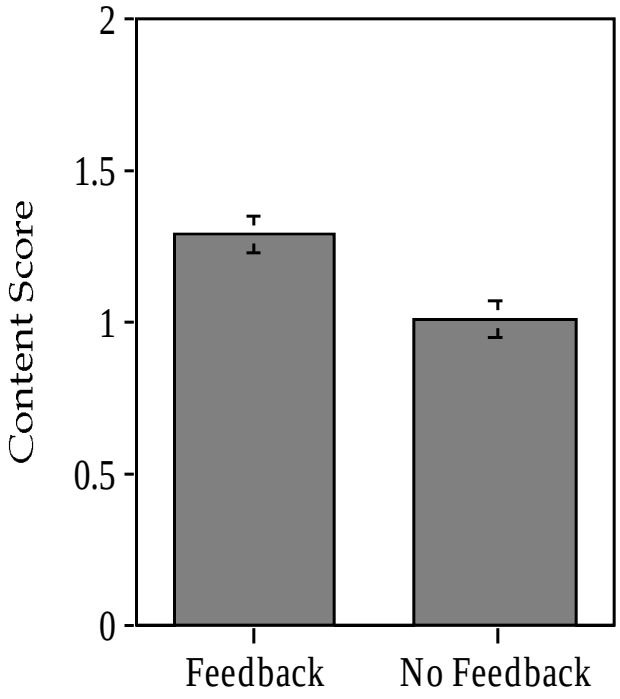

Your Summary
summary street
Your name: Guest
Text you are summarizing: Maya
Your summary should be 200 to 300 words long.
The Maya civilization was really cool! The ancient Maya lived primarily in Central America from 300 to 900 AD. they live in the jungles so it took a long time to find their ruins. Around 900 AD the Maya seemed to have abandoned their cities and returned to simpler lives. Archeologists are not sure why, but maybe things like drought, disease, overpopulation and crop failures, or wars with the Toltecs caused the collapse of the Mayan empire.
The Maya invented a system of pictures and symbols, called hieroglyphs. They recorded their history on stone slabs, called stelae. They also wrote books on paper with folded pages, like our books. But their books were made from the bark of fig trees that they folded together. Only the rulers and specially trained scribes could read and write them.
The Maya made many contributions to math and science. Their calendar is a more accurate calendar than the one we use today. The Mayan number system was based on the number 20. It used a placeholder for zero. Their zero looked like a football. They wrote their numbers vertically rather than horizontally.
The Maya believed in many gods. They believed if the gods were happy they would protect them from disease and drought and natural disasters. To keep the gods happy they made a daily sacrifice of blood. Everybody in the community donated blood. They dripped blood from a cut onto a paper and burned it. They also worshipped their ancestors like gods and buried the dead with food and clothing to take them to another life. The Maya built temples on top of stone pyramids with many steps. The steps were really steep and narrow.
The ancient Maya were good farmers who dug canals to water their crops and terraces to keep
Save changes/check my spelling
Feedback on my summary
Format for Printing
Close

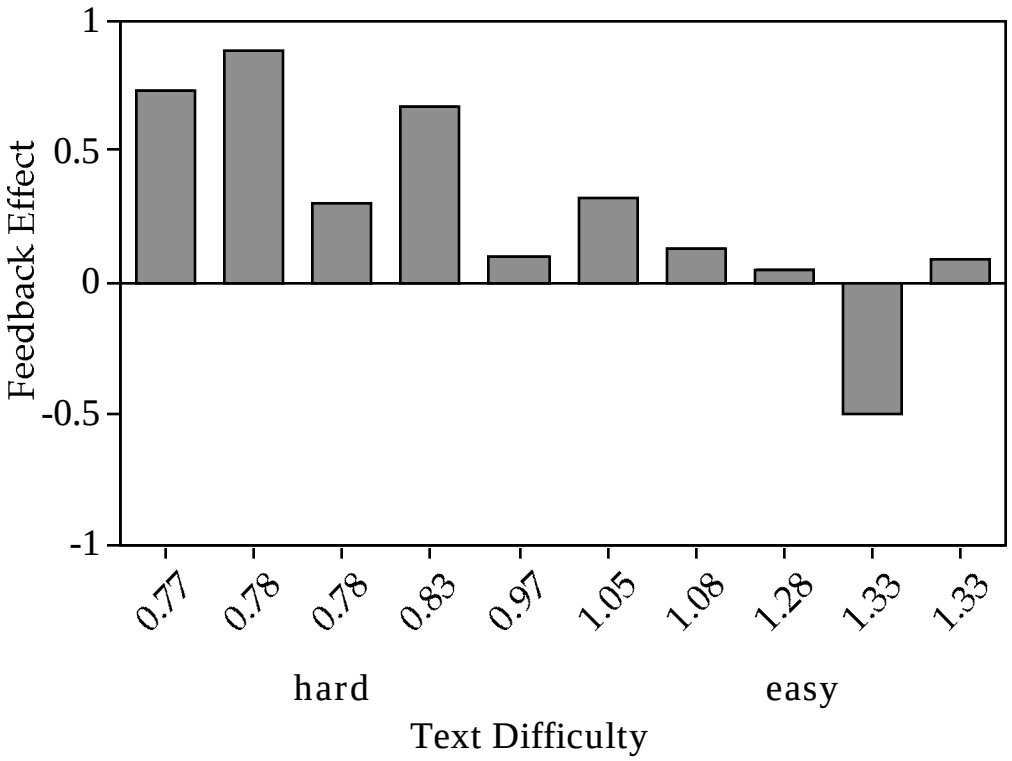

How Well Did You Cover Each Section?
summary street
History
Maya Writings, Math and Science
Religion, Architecture and Art
Maya People
Guest, these bars show how well your summary covered the sections of the text you read. If the bar passes the black line, then you've written enough information about that section. When your summary contains enough information about every section, you can advance to the next level and will receive more advice on how to improve your summary.
The blue dashed lines show how much information your previous summary contained, so you can see if you are improving or not.

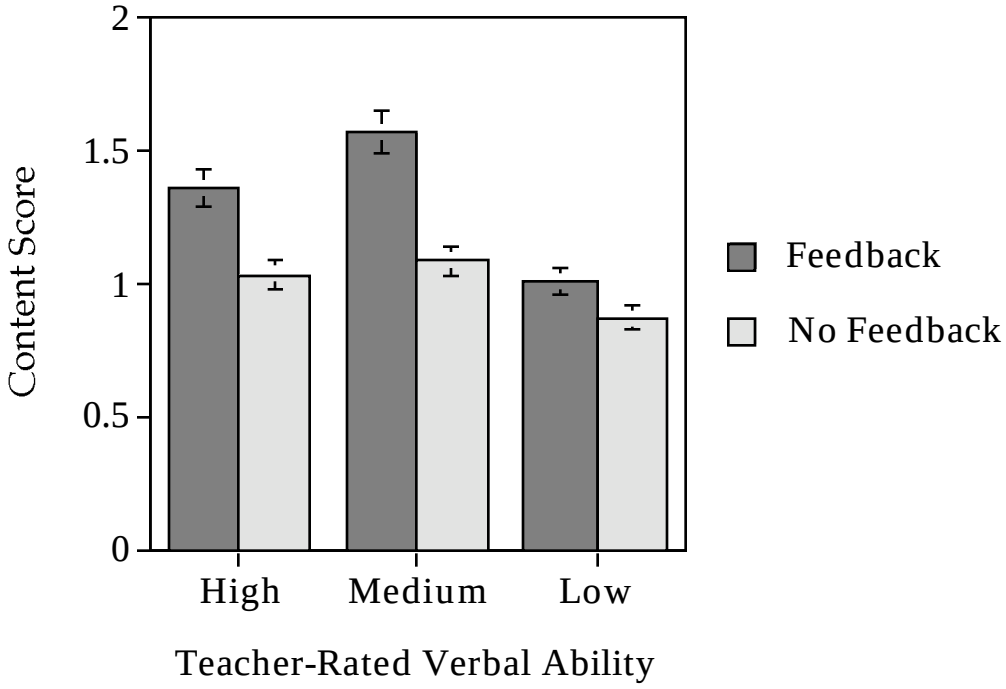

Redundancy Check
Guest, your summary is shown below. Redundant sentences (sentences that I thought have too much overlap with other sentences) are marked in red, green, blue, purple, or orange. I have also numbered the redundant sentences to make it easier to see the sentences that I thought overlapped with one another.
If you agree that the sentences that I have marked are redundant, try to combine the information from those sentences into a single sentence, or get rid of one of them if you feel the information is contained in the other sentence.
The Maya civilization was really cool! The ancient Maya lived primarily in Central America from 300 to 900 AD. They lived in the jungles so it took a long time to find their ruins. Around 900 AD the Maya seemed to have abandoned their cities and returned to simpler lives. Archeologists are not sure why, but maybe things like drought, disease, overpopulation and crop failures, or wars with the Toltecs caused the collapse of the Mayan empire.
The Maya invented a system of pictures and symbols, called hieroglyphs. They recorded their history on stone slabs, called stelae. #1 They also wrote books on paper with folded pages, like our books. #1 But their books were made from the bark of fig trees that they folded together. Only the rulers and specially trained scribes could read and write them.
The Maya made many contributions to math and science. Their calendar is a more accurate calendar than the one we use today. The Mayan number system was based on the number 20. #2 It used a placeholder for zero. #2 Their zero looked like a football. They wrote their numbers vertically rather than horizontally.
The Maya believed in many gods. They believed if the gods were happy they would protect them from disease and drought and natural disasters. To keep the gods happy they made a daily sacrifice of blood. Everybody in the community donated blood. They dripped blood from a cut onto a paper and burned it. They also worshipped their ancestors like gods and buried the dead with food and clothing to take them to another life. The Maya built temples on top of stone pyramids with many steps. The steps were really steep and narrow.
The ancient Maya were good farmers who dug canals to water their crops and terraces to keep them from washing away. Corn, called maize, was their most important crop. The Maya were also weavers. And they liked to play ball games, just like us.
Today some Maya still live in Central America. Sometimes they have trouble with the government.
Internet zone


Relevance Check
Guest, your summary is shown below. Irrelevant sentences (sentences that I think don't say much about the topic) are marked in red.
If you agree that the sentences that I have marked are irrelevant, either delete them or combine the information from those sentences into other sentences in your summary.
Note:Spelling errors can cause me to think that good sentences are irrelevant. Be sure to check your spelling!
The Maya civilization was really cool! The ancient Maya lived primarily in Central America from 300 to 900 AD. They lived in the jungles so it took a long time to find their ruins. Around 900 AD the Maya seemed to have abandoned their cities and returned to simpler lives. Archeologists are not sure why, but maybe things like drought, disease, overpopulation and crop failures, or wars with the Toltecs caused the collapse of the Mayan empire.
The Maya invented a system of pictures and symbols, called hieroglyphs. They recorded their history on stone slabs, called stelae. They also wrote books on paper with folded pages, like our books. But their books were made from the bark of fig trees that they folded together. Only the rulers and specially trained scribes could read and write them.
The Maya made many contributions to math and science. Their calendar is a more accurate calendar than the one we use today. The Mayan number system was based on the number 20. It used a placeholder for zero. Their zero looked like a football. They wrote their numbers vertically rather than horizontally.
The Maya believed in many gods. They believed if the gods were happy they would protect them from disease and drought and natural disasters. To keep the gods happy they made a daily sacrifice of blood. Everybody in the community donated blood. They dripped blood from a cut onto a paper and burned it. They also worshipped their ancestors like gods and buried the dead with food and clothing to take them to another life. The Maya built temples on top of stone pyramids with many steps. The steps were really steep and narrow.
The ancient Maya were good farmers who dug canals to water their crops and terraces to keep them from washing away. Corn, called maize, was their most important crop. The Maya were also weavers. And they liked to play ball games, just like us.
Today some Maya still live in Central America. Sometimes they have trouble with the government.


80
60
40
20
0
Time on Task (in minutes)
Feedback
No Feedback


2
1.5
1
0.5
0
Content Score
Feedback
No Feedback


1
0.5
0
-0.5
-1
Feedback Effect
0.77
0.78
0.78
0.83
0.97
1.05
1.08
1.28
1.33
1.33
hard
easy
Text Difficulty


2
1.5
1
0.5
0
Content Score
High
Medium
Low
Teacher-Rated Verbal Ability
Feedback
No Feedback