Taylor & Francis
Taylor & Francis Group

DISFLUENCIES IN COMPREHENDING ARGUMENTATIVE TEXTS

MEREDITH LARSON, M. ANNE BRITT, and AARON A. LARSON

Northern Illinois University, Dekalb, Illinois, USA

In two experiments, we examine university students' ability to comprehend authentic argumentative texts and factors that influence their application of this skill. Participants read several relatively lengthy arguments and identified the main claim and reasons. Experiment 1 shows that participants are not skilled at identifying key elements from an argumentative text (only 30% accuracy). The performance of participants of all skill levels was aided by explicit markers to signal the critical elements. Claim and reason identification errors showed interesting patterns such as identifying either an uncontroversial statement of the theme or a counter-argument as the claim. Based on the results from Experiment 1, a short argument tutorial was created and tested in Experiment 2. The tutorial was helpful when readers were reading simply to comprehend but not when they were reading to rebut. The results point to the need for further instruction in argument comprehension skills and the need for additional research into the processes involved.

Students read and understand arguments in a variety of classroom situations and must be prepare to continue this practice in the workplace and in real-life decision-making situations. In high-school and college classrooms, students are required to read arguments to complete class assignments (e.g., research papers, debates, and literary criticism), to meet National Standards in disciplines (National Center for History in the Schools, 1996), and to take standardized tests such as the NEAP science test and persuasive writing test. Beyond the classroom, argument comprehension is required by many college and post-graduate entrance exams such as the SAT, ACT, LSAT, GRE and MCAT. In addition to

Address correspondence to M. Anne Britt, 363 Psychology-Math building, Northern Illinois University, Dekalb, IL 60115. E-mail: britt@niu.edu

their scholastic importance, argument comprehension skills are essential for learning across the lifespan, for many careers and for participating in our democracy. Despite the importance of reading arguments and the countless opportunities we have to practice it, many high-school graduates have difficulty forming and understanding written arguments. For instance, the most recent national assessment found that only 12% of 12th graders were proficient in argument skills (National Assessment of Educational Progress, 1996) and only 13% of 12th graders' essays were judged "skillful" or "excellent" (National Assessment of Educational Progress, 1998). Given the importance of argument comprehension, it is essential that all students become proficient text-based reasoners and an important step toward this goal is to assess college students' ability to read and comprehend complex argumentative text.

Comprehending arguments involves many higher-order reading and reasoning processes; the most basic of which is representing essential argument information. In much the same way as students use narrative schema to comprehend stories (Goldman, Graesser, & van den Broek, 1999; Mandler & Johnson, 1977; Stein & Glenn, 1979; Trabasso & van den Broek, 1985), it is reasonable to assume that students use argument schemas to comprehend arguments. The essence of any argument is a claim (e.g., Cell phone use while driving should be banned) supported by one or more reasons (Toulmin, 1958). A claim forms an argument once it is supported by a reason (e.g., Distracted drivers are dangerous drivers). Reasons can, themselves, be supported (e.g., the statistics from the NHTSA director). The claim holds the top position of an argument schema because all other elements of the argument are presented to either support or oppose this main proposition. As with other schema, awareness of an argument schema can allow readers to accurately represent the author's intended meaning. It can also aid in organizing subsequent text and guide coherence inferences.

Two studies have looked directly at high-school (Chambliss, 1995) and grade school students' (Chambliss & Murphy, 2002) use of argument schemas to guide argument comprehension. They presented participants instructional-type argument texts similar to the prototypic structure of argument texts taught in secondary composition courses. In particular, the texts were highly structured with strong topic sentences and introductory and concluding

paragraphs summarizing the structure. These researchers found that advanced high-school readers and a majority of the younger readers were able to use their knowledge of an argument structure to correctly organize the basic element in the texts, especially as the external organizers were increased. These argument texts were specially written to conform to the structure taught in class and may have artificially increased performance. It is unclear how well young readers or adults would be able to comprehend authentic arguments of the type that they would encounter outside the classroom.

Much more is known, however, about children's *production* of arguments. It has been shown that even young readers can use their knowledge of argument structure to produce arguments. Around the age of seven, children appear to become capable of rudimentary argument production (Golder & Coirier, 1994; Miller, 1986, 1987; Stein & Miller, 1993). For instance, they can state an explicit position and generally support their claim with a reason (Coquin-Viennot & Coirier, 1992; Golder & Coirier, 1994). By the end of their twelfth year, children begin to support their claims with more than a single reason, show an increased ability to detect invalid inferences, and begin to mention counter-arguments (Golder & Coirier, 1994). For many children, argument production ability continues to improve gradually with age, though several aspects of argumentation (e.g., counter-arguments and qualifiers) are still not mastered by the end of high school (Golder & Coirier, 1994; Knudson, 1992; 1994). Given that students can create simple arguments, they must have at least a rudimentary argument schema that could be used to comprehend written argumentative discourse.

Special lexical and syntactic elements in a narrative text can serve as processing cues for constructing a mental representation of the text (Graesser, Millis, & Zwaan, 1996). Blakemore (1987, 2000) points out that some words and constructions, such as "that is" "but" and "because," do not contribute to the truth-conditions of a proposition but that they do provide processing or conceptual information. In addition to such connectives, discourse markers include headings, numbering, summaries, and typeface changes. These devices are used by writers to emphasize the structure of texts and to signal importance. Prior research has shown that readers' representation of the structure of narrative and expository texts is aided by signals that clarify the structure (Beck, McKeown,

Sinatra, & Loxterman, 1991; Kozmnsky, 1977; Lorch & Lorch, 1995, 1996; Lorch, Lorch, & Inman, 1993; Meyer & Rice, 1989; Millis & Just, 1994). This is especially true for challenging texts (Spyridakis & Standal, 1987) and for low-knowledge participants (Goldman, Saul, & Cote, 1995; Lorch & Lorch, 1996).

In a similar manner, there are discourse cues that can signal argument elements (e.g., claims and reasons) and their relationships. The types of discourse markers that can signal argument elements include connectives (e.g., *because, therefore, although*), headings (e.g., "evidence for"), and organizational markers (e.g., *first, second, finally*). These discourse markers can direct the reader to appropriately connect the various argument elements into a coherent structure. In fact, Britt and Larson (2003) found that statements marked by modals (e.g., *should*) and uncertainty markers (e.g., *probably*) signaled controversial assertions requiring support. Although attending to such features may be largely automatic for some students, they require inferences that may present less-able students with a significant challenge. Such students may require instruction and practice to become accustomed to attending to these cues.

The purpose of the current experiments is to explore and improve students' comprehension of written argumentative texts. In Experiment 1, we assessed college students' ability to identify claims and supporting reasons and the extent to which this process is aided by explicit marking of the argument's rhetorical structure. This experiment was designed to provide a baseline for students' argument comprehension ability, assess students' sensitivity to argument structure as signaled by discourse markers, and provide a corpus of errors to guide the development of an argument tutorial. Then, based on an analysis of the type of errors from Experiment 1, we created a tutorial to aid students in argument comprehension. Experiment 2 tested the effectiveness of this tutorial and instructions to evaluate the argument during reading.

Experiment 1

In Experiment 1, participants read several authentic argumentative texts with the goal of identifying the main claim and reasons. Performance on this task indicates the degree of skill that students possess in reading natural arguments and the extent to which they

notice crucial elements of argumentative discourse. In addition, half of the participants read modified versions of the texts that included explicit discourse marking to signal the construction of an argument schema. If students are not sensitive to such discourse markings, then one would not expect to see improved identification scores when they are present. Lack of improvement also suggests that students could benefit from instructions that explain and illustrate the use of discourse markers. Improved performance would suggest that students are aware of the organizational structure of arguments but have difficulty identifying major elements without explicit markings.

There is reason to believe, however, that there may be expertise differences. McNamara, Kintsch, Songer, and Kintsch (1996) found that low-knowledge readers comprehended a science text better when it had explicit connectors signaling the high-level text structure. In contrast, high-knowledge readers performed better with a less explicit text. That there was an effect of knowledge leads to a concern that good reasoners may be hindered by having explicit markings. The explicit markings may lead good reasoners to more passive reading because there is no need to make inferences to construct an argument schema. Therefore, it may be the case that explicit markings are helpful only for poor and middle reasoners and that they may in fact lead to worse comprehension for good reasoners. For this reason, we used a subsection of the Law School Admissions Test (LSAT) to classify participants on verbal-reasoning ability.

Method

PARTICIPANTS

Seventy-six native English-speaking students from an introductory-level Psychology class at Northern Illinois University participated for partial course credit.

MATERIALS

The materials were excepts from 7 argumentative essays[1]. Six of the arguments were excerpts from naturally-occurring

---

[1]Two additional arguments were in the packet for piloting for a later experiment. The task was identical.

documents taken from textbooks on argumentation. The other argument was written by an experimenter. The essays were shortened without eliminating distracting information such as titles, rhetorical questions, and counter-arguments. Three arguments were policy arguments (The proposed Delaware expressway should not be built; It is not in our interests to pass death penalty laws; Northern Illinois University should impose a more rigorous two-year math requirement for all students), two arguments were value arguments (TV has a negative influence on children and family life; The U.S.'s actions in the Panamanian Revolution were proper and guiltless), and two arguments were factual arguments (Hypnosis is sneaking ever so quietly into the medical mainstream; The law requires the insanity defense). Each argument had one main claim and between 2 and 10 reasons ($M = 5.14$, $S = 2.79$). The arguments varied in length from 103 to 732 words with an average of 426.21 words. The texts were complex with an average Flesch-Kincaid grade level of 11.21 ($S = 0.90$).

DESIGN

Participants were randomly assigned to one of two marker-explicitness conditions: Implicit or Explicit. In the Implicit version, participants read the original excerpts. In the Explicit version, a variety of discourse markers were added to emphasize aspects of the rhetorical structure. Primary markers signaled specific argument elements for the claim (e.g., "we propose that," "I submit that," "for all of these reasons," "we conclude that," "the evidence seems to prove conclusively that"), reasons (e.g., "most indicative of X is the fact that," "several reasons to support," "another reason is," "a stronger reason," "X is needed because"), and backing (e.g., "consider the fact that," "Evidence for . . . . can be found in a study"). In addition, rhetorical connectives (e.g., *however, therefore, although*) and organizational markers (e.g., *first, second, finally*) were included. The Explicit and Implicit Versions of one text is presented in the Appendix to show the type of marking added.

The second factor was Reasoning Skill as measured by a selection of 18 questions from past versions of the Law School Admissions Test (LSAT) that require verbal reasoning ability (e.g., judge implications, assumptions or flaws in a passage or to determine what would strength or weaken an author's arguments). The LSAT is an achievement test of reading and verbal reasoning skills that

is used to screen applicants for admittance to law school which requires skilled analysis of written arguments. The LSAT test has been shown to correlate with academic successes in the first year of law school in which verbal reasoning skills are paramount (Anthony, Harris, & Pashley, 1999) and consistently obtains re-test reliability coeffients between .90 to .95 (Law School Admission Council, 2001).

PROCEDURE

Participants were given instructions followed by a practice argument. The instructions told participants that they would have to write down the author's main claim and underline (and number) any reasons mentioned by the author to support that claim. They were also told that the author may make more than one claim and that they were to focus on the most important one. Then, participants completed the practice argument and checked their answers with those provided by the experimenter. If they had no questions, they received the packet of the arguments and were to perform the same task (i.e., state claim and underline and number the reasons) for each essay. After completing the packet, participants completed the verbal reasoning test (i.e., LSAT).

During a debriefing session in the class following the experimental session, participants were given two questions to rate their interest and skill on a 7-point Likert scale with 1 being "strongly agree" and 7 being "strongly disagree." The first question was "I enjoy reading arguments" and the second was "I am skilled at understanding arguments."

Results

SCORING

The main claims and reasons were scored for accuracy. Claims were scored as correct if they captured the main proposition of the claim. Paraphrases were correct even when the focus was changed from a policy (e.g., The proposed Delaware expressway should not be built) to a value (e.g., the highway is bad). Reasons were scored according to the accuracy of the statement underlined in the text. Most participants included numbers in front of each underlined reason as directed. Occasionally a participant did not include numbers for reasons. In these cases, a single unbroken line stroke

TABLE 1. Descriptive Statistics for the Correct Identification of Argument Elements (of the 43 Possible) by Marker Explicitness and Reasoning Skill.

Reasoning skill	Implicit			Explicit		
	Mean	*SD*	Proportion (%)	Mean	*SD*	Proportion (%)
Less skilled	8.62	3.10	20	11.67	5.82	27
Moderately skilled	12.07	3.99	28	14.72	6.20	34
Highly skilled	14.00	3.43	33	17.56	8.57	41

was considered a single reason. If part of the segment underlined included an actual reason, then it was scored as correct. Inter-rater reliability was 81%.

ARGUMENT IDENTIFICATION ACCURACY

Argument identification scores were obtained by combining the number of correctly identified claims (of a possible 7) and reasons (of a possible 36). Overall, accuracy tended to be low. Across all conditions, participants identified only an average of 12.92 ($S = 5.86$) argument elements of the 43 possible. Table 1 presents the argument identification scores for each Reasoning Skill group (Less Skilled, Moderately Skilled, and Highly Skilled) and each Marker Explicitness condition (Explicit vs. Implicit marking). A 2 (Marker Explicitness) × 3 (Reasoning Skill) Between-subjects ANOVA revealed a significant main effect of Reasoning Skill, $F(2, 70) = 5.94$, $p < .05$. A Bonferroni post-hoc analysis, at 0.05 level of significance, showed that high-skilled reasoners ($M = 15.78$) correctly identified more elements than only the less-skilled reasoners ($M = 10.08$). There was also a significant main effect of Marker Explicitness, $F(1, 70) = 5.81$, $p < .05$. Those in the explicit condition found more of the reasons and claims ($M = 14.44$) than those in the implicit condition ($M = 11.32$). There was no interaction between Reasoning Skill and Marker Explicitness, $F(2, 70) = 0.031$, $p = 0.90$.

CLAIM IDENTIFICATION ERRORS

In addition to analyzing group differences, interesting patterns emerge from a detailed analysis of the type of errors students make. The *content* of the claim errors were categorized into 6 argument error types. The content of the errors reveals that

most (37%) were non-controversial main thesis statements such as general statements or questions on the controversy without taking a stance (e.g., "The death penalty from a positive, negative and economical point of view," "The validity of the death penalty," "Should the death penalty exist or not?"). The next two most common errors were mis-identifying supported reasons (16%) and counter-arguments (16%) as the main claim. Supported reasons are actually sub-claims that are elaborated with backing. For example, one of the reasons provided to support not having the death penalty is that it is unfair. This claim was then supported by evidence that it is both racially and economically biased. While this was supported, it is not the primary claim presented by the author. While the mis-identification of supported reasons for the main claim is a problem, it is not as serious as the mis-identification of counter-arguments for the main claim. These are the opposite of the author's intended message. Finally, participants rarely mistook unsupported reasons (10%), backing (8%) or alternative claims (additional claims on the same side and hierarchical level of the main claim) (8%) for the main claim. The final 9% of the responses were unsystematic.

The structure of the claim identification errors also provide interesting data. We examined the structure of the arguments that students identified when they got the main claim wrong (49% of the possible claims were incorrectly identified). On a case-by-case basis, we looked at the number of reasons that could support the claim they identified. If we accept the claim identified by a participant, then only 33% of the reasons actually provided support for their identified claim. A small proportion of the reasons (9%) were merely restatements of their claim and as such could not be considered support. All of the other reasons selected (59%) could not have actually supported their proposed claim. For example, one student thought the claim to the death penalty argument was that "The death penalty is an ineffective punishment." The student then identified two of the reasons to be: "Death penalty is not fair" and "It discriminates by race," neither of which directly support the student's statement of the main claim.

To examine whether marker explicitness and reasoning skill influenced the frequency of disconnected reasons, a 2 (Marker Explicitness) × 3 (Reasoning Skill) between-subjects ANOVA was performed on the number of non-supporting reasons for incorrectly identified claims. The means and standard deviations for

TABLE 2. Average Number of Non-Supporting Reasons for Mis-Identified Claims by Marker Explicitness and Reasoning Skill.

Reasoning skill	Implicit	Explicit
Less skilled	11.08 (1.63)	9.17 (1.70)
Moderately skilled	7.93 (1.52)	6.67 (1.70)
Highly skilled	6.67 (1.96)	1.33 (1.96)

Note. Standard deviations in parentheses.

each condition are shown in Table 2. This analysis revealed a significant main effect of Reasoning Skill, $F(2, 70) = 5.69$, $p < .05$. A Bonferroni post-hoc analysis, at 0.05 level of significance, showed that less-skilled reasoners identified a greater number of non-supporting reasons ($M = 10.16$) than high-skilled reasoners ($M = 4.00$). There was also a significant main effect of Marker Explicitness, $F(1, 70) = 4.15$, $p < .05$. The explicit condition led to fewer non-supporting reasons ($M = 6.21$) than the implicit condition ($M = 8.73$). There was no interaction between Reasoning Skill and Marker Explicitness, $F(2, 70) = .041$, $p = 0.96$. Explicit marking aided all skill groups equally in not presenting reasons that fail to support their claims.

RATED INTEREST AND SKILL

Participants rated their own level of enjoyment for reading arguments and skill at understanding arguments on a 7-point scale. Participants moderately enjoyed reading arguments ($M = 3.85$) and believed they were only moderately skilled at reading arguments ($M = 3.76$). Surprisingly, these ratings were not statistically correlated with their LSAT scores on actual argument comprehension skill. The correlation of LSAT score with enjoyment ratings was -0.14 and with perceived skill was -0.11.

Discussion

Participants correctly identified only 30% of the 43 main argument elements. The baseline performance, in the implicit condition, was significantly lower at 27%. Note that this was a search task not a memory task. Participants had the text in front of them while identifying the key argument elements. Therefore, any difficulty could

not be attributable to memory limitations. This low performance shows that participants require support from either the author, such as discourse markings, or from argument structure training. Participants were able to make use of explicit marking, leading to an average of 34% of correct identification. All skills groups were aided by explicit marking signaling argument structure. Thus, it was not the case that participants performed at a ceiling level, and it was not the case that high reasoners performed so well that explicit markers interfered with comprehension. Therefore, if the author's intention is to make the argument understandable, explicit signals help with little cost.

These results point to the need for remediation. One problem with any attempt at instruction may be hampered by students' lack of awareness of their skill level as shown by the low correlation between perceived argument skill and LSAT score. If participants already perceive themselves to be skilled argument comprehenders, they may not gain the full benefit of a tutorial. Nevertheless, we have used the results from Experiment 1's error analysis to create a tutorial to provide direct instruction in argument comprehension.

CREATING AN ARGUMENT TUTORIAL

The identification of the claim is especially important and an analysis of the distribution of errors provides some direction for possible instruction. First, mistaking the theme for the main claim suggests that students may need a definition of a claim as a controversial stance not just a question or theme as taught in high-school composition courses. Second, mistaking a supported sub-reason could be aided by prompting students to verify that this claim does in fact support their identified higher-level claim. Finally, participants' selection of a counter or opposing claim is partially explained by an order effect. Most of the mis-identified counter-arguments occurred prior to a statement of the main claim. Interestingly, most of the alternative claims occurred after the main claim. Thus, perhaps some of these errors were a result of students' prematurely jumping to a conclusion before reading the entire argument. This will have to be examined more closely in follow-up experiments. However, it may be that students could be instructed to read the entire article before deciding on the main claim. In authentic arguments, authors frequently do not state their main claim in the first paragraph. In fact, if students are using a heuristic

to determine the main point of the first paragraph, they would have been wrong in all 5 of the arguments that were longer than a single paragraph. Thus, instruction could include a statement to read the complete argument before deciding upon the main claim.

Similarly, an analysis of the relationship between the provided reason and the stated claim when the claim was incorrectly identified also provide interesting avenues for instruction. These structural errors suggest that students failed to check whether the stated reasons supported the stated claim. Some of these errors may be eliminated if students reviewed the claim-support structure of their response before submitting it. This may even lead students to re-evaluate incorrect main claims. Many participants actually correctly identified reasons but mis-identified the claim. Perhaps they are able to use cues to identify important items, but fail to verify that these identified items formed a coherent argument.

Experiment 2

The poor performance of all skill groups in comprehending authentic argumentative essays shows that students do not adequately process and represent the main structure of such texts. Given the educational goals of producing students skilled in comprehending argumentative texts, how might such deficiencies be ameliorated? Perkins (1995) and Kuhn (1995) both concluded that these failings may be a result of lack of practice, exposure, and explicit instruction in the skill of argumentation. Our results from Experiment 1 suggest that students may benefit from further instruction, and our analysis of their errors provides guidance for the type of information that should be included in the argument tutorial.

In Experiment 2, we constructed a short tutorial that defined key argument terms, challenged common misconceptions about arguments, and explained a series of steps to comprehend written arguments. Participants were provided an example and an opportunity to practice the presented steps. The other half of the participants received no instruction but received an opportunity to practice the task for the same amount of time. Then participants were given 4 of the texts from Experiment 1 and were asked to identify the main claims and reasons.

The effect of reading goal was also examined as a factor to improve readers' argument comprehension. Half of the participants

were asked to read for comprehension while the other half were asked to read to rebut the argument. According to the Evaluation hypothesis, argument comprehension may be improved when people read with the goal of evaluating the author's argument. If so, then those participants asked to rebut the argument should identify argument elements better. In contrast, according to the Comprehend First hypothesis, argument evaluation should occur in two separate phases. First one should fully understand what the author is arguing and only later should they worry about evaluating the message. This may be especially true for low-knowledge participants and participants still learning to understand arguments. If this is the case, then readers in the comprehension condition should out perform participants in the rebuttal condition. Of course, reading instruction may interact with the presence or absence of a tutorial. It may be the case that any effect of reading goal (e.g., the Evaluation hypothesis or the Comprehend First hypothesis) is only effective when participants are taught how to read arguments. In the absence of such instruction, there may be no effect of reading goal.

Method

PARTICIPANTS

Seventy native-english speakers enrolled an introductory-level Psychology course at Northern Illinois University participated for course credit.

MATERIALS

Four of the argument essays from Experiment 1 were used in this experiment (death penalty, Panama canal, hypnosis, and insanity defense). Because longer excerpts were used for each argument, ranging from 523 to 1233 words with an average of 915.30 words, participants read only 4 texts. The difficulty level of the texts was similar to Experiment 1 with a Flesch-Kincaid grade level of 11.28 ($S = 0.91$).

The Argument Tutorial began with a brief definition of key argument elements (i.e., claim, reason, qualifiers, counter-claim, and rebuttal) and a humorous example argument (whether one should ask for directions when lost). To help students be more amenable to this instruction, we presented and refuted four

common misconceptions: that arguments are fights, that they are
not useful, that they are not comprehensible by everyone, and
that they are clearly written. The first three of these misconcep-
tions were identified during an informal debriefing from a prior
study in answer to why they avoid or don't like reading arguments.
Finally, the tutorial elaboratively presented a series of steps to com-
prehend and evaluate arguments. These included: Identify the
main claim and state it clearly; Identify all important reasons and
state them clearly; Evaluate whether each reason is true; Evalu-
ate how strongly each reason supports the claim; and Identify the
weaknesses in the argument. Throughout the tutorial, participants
were asked to practice each skill.

DESIGN

The design was a 2 Training (Argument Tutorial or No-
feedback Practice) × 2 Reading Goal (Comprehension vs. Re-
buttal) between-subjects design. For the Training factor, half of
the participants received an short concise argument tutorial. The
other half of participants received No-feedback Practice. In this
condition, participants were given a practice argument (927 words,
11.4 grade level). First they wrote a summary and then stated the
main claim and listed all the reasons the author provided for sup-
port. The tasks in both of these conditions took approximately the
same amount of time (10 minutes).

There were two conditions for the Reading Goal factor. Partic-
ipants in the Comprehension condition were instructed to read for
comprehension, and after each argument, they had to create one
additional reason that supports this claim but was not mentioned
in the text. Participants in the Rebuttal condition were asked to
critically evaluate the argument while reading. After each argu-
ment, they were asked to write down one reason to rebut one of
the author's arguments.

PROCEDURE

Participants were randomly assigned to one of four condi-
tions when they arrived for the experiment. After completing the
Argument Tutorial or No-feedback Practice depending on the con-
dition assigned, participants were given a packet of the arguments
that included instructions. The instructions told participants that
they would have to write down the author's main claim and list

any reasons mentioned to support that claim. They were also told that the author may make more than one claim and that they were to focus on the most important one. After each argument, the Comprehension participants were asked to list an unmentioned reason to support the author's main claim while Rebuttal participants were asked to list an unmentioned rebuttal to the author's main argument. After completing the packet, participants rated the strength of each of the author's reasons. For each argument, the author's main claim was presented at the top of sheet of paper, and the reasons were listed below. For each reason, participants rated on a 10-point scale the degree of strength, from very weak to very strong and justify their ratings.

Results

Argument identification scores were obtained by combining the number of correctly identified claims (of a possible 4) and reasons (out of a possible 24). Accuracy was generally low (43%) but improved from Experiment 1. Participants correctly identified an average of 12.13 ($S = 6.81$) of the 28 possible argument elements. A 2 (Training: Argument Tutorial or No-feedback Practice) × 2 (Reading Goal: Comprehension vs. Rebuttal) Between-subjects ANOVA was conducted. There was a significant Training × Reading goal interaction, $F(1, 66) = 3.94$, $p < .05$. The Argument Tutorial for the Comprehension group identified more argument elements ($M = 16.06$, $S = 7.89$) then the Argument Tutorial for the Rebuttal condition ($M = 10.39$, $S = 6.21$) or either No Feedback group ($M = 10.82$, $S = 6.91$ for Comprehension, and $M = 11.39$, $S = 5.03$ for Rebuttal condition). Neither main effect of Training, $F(1, 66) = 1.82$, $p > .05$, nor Reading Goal, $F(1, 66) = 2.64$, $p > .05$, was significant.

Discussion

Providing participants a short tutorial aided argument identification but only when participants focused on the single goal of comprehending the argument. Thus, the tutorial was effective, but directions to evaluate while reading were not helpful and in fact removed the positive effect of the tutorial. Contrary to the Evaluation Hypothesis, a rebuttal goal did not lead to better argument

analysis. These results support the Comprehend-first Hypothesis, stating that comprehending should be done first, but this is qualified by the need for a tutorial prior to reading. Evaluating during comprehension may require more resources than our participants had available, at least for non-expert argument comprehenders reading an unfamiliar topic.

General Discussion

The results of these experiments show that our sample of college students were not proficient comprehenders of natural, written arguments. They identified only about one-third of the main claims and reasons, they selected reasons that could not possibly support their stated claim, and they often identified a stated counter-claim as the main claim (16% of errors). This disappointing performance must be addressed through instruction. Argument comprehension and evaluation are required in the college classroom (e.g., research papers and literary criticism), in many professions (e.g., law, medicine, and teaching), and to participate as a citizen in our democracy. These experiments explore the efficacy of two methods of enhancing argument comprehension. Experiment 1 showed that participants were aided by support from the author in the form of explicit discourse markers; Experiment 2 showed that participants were aided by a brief tutorial that explained the process of argument comprehension as long as they were not given the dual task of reading to rebut.

The current experiments also provide general suggestions for remedying the situation. First, although students do better when reading well-structured texts, instruction must go beyond the neatly structured texts to more complex argument presentations. The claim-identification errors point to the limits of teaching students to comprehend and compose arguments using only highly-structured arguments that are not typical of the structure of authentic argumentative texts. Many arguments in real-life do not conform to the prescribed structure, and this may actually lead students to mis-understanding. In fact, the errors from Experiment 1 showed that students often mis-perceived the author's point as the opposite of the actual message. This may be a result of students making a mistaken assumption: that authors state their main point in the first sentence or first paragraph. In our unsystematic

sampling of arguments, we noticed that authors commonly present and dismiss a counter-argument before presenting the central argument. It may be that students are unaware of this practice, and they may not understand why opposing points are mentioned in the same text (Means & Voss, 1996; Perkins, 1985). These problems can be addressed by including more complex arguments as part of instruction.

Second, students' performance did not demonstrate care in re-constructing coherent arguments. They identified claims and reasons without verifying that the reasons actually supported their selected claim. While in some cases this may have been a result of a lack of domain knowledge, it may indicate a problem in skilled argument comprehension. For example, some students may require practice in understanding the connection between reasons and claims. Toulmin (1958) refers to general rules, laws, or principles that permit the conclusion given the reason as warrants. Perhaps practice aimed at helping students notice and evaluate warrants may lead to better comprehension of complete arguments.

It appears that college students are still novice argument comprehenders. A relatively short (10 minutes) tutorial was able to improve their performance. Thus, unlike narrative comprehension, this area seems ripe for looking at the acquisition of a skill in young adults. There are many open questions: What are the component skills of argument comprehension? What is the nature of argument representations? Do certain characteristics of the reader influence acquisition? Existing theories of text processing provide valuable predictions on these questions and argument comprehension should serve as an interesting testbed for these theories.

References

Anthony, L. C., Harris, V. F., & Pashley, P. J. (1999). *Predictive Validity of the LSAT: A National Summary of the 1995–1996 Correlation Studies 3.* Law School Admission Council Research Report Series Technical Report, 97–01.

Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly, 26*(3), 251–276.

Blakemore, D. (1987). Linguistic constraints on pragmatic interpretation: A reassessment of linguistic semantics. *Behavioral and Brain Sciences, 10*(4), 712–713.

Blakemore, D. (2000). Indicators and procedures: Nevertheless and but. *Journal of Linguistics, 36*, 463–486.

Britt, M. A., & Larson, A. A. (2003). Constructing representations of arguments. *Journal of Memory & Language, 48*(4), 794–810.

Chambliss, M. J. (1995). Text cues and strategies successful readers use to construct the gist of lengthy written arguments. *Reading Research Quarterly, 30*(4), 778–807.

Chambliss, M. J., & Murphy, P. K. (2002). Fourth and fifth graders representing the argument structure in written texts. *Discourse Processes, 34*(1), 91–115.

Coquin-Viennot, D., & Coirier, P. (1992). Structures of argumentative discourse: Effects of type of referential space. *European Journal of Psychology of Education, 7*(3), 219–229.

Golder, C., & Coirier, P. (1994). Argumentative text writing: Developmental trends. *Discourse Processes, 18*, 187–210.

Goldman, S. R., Saul, E., & Cote, N. (1995). Paragraphing, reader, and task effects on discourse comprehension. *Discourse Processes, 20*(3), 273–305.

Goldman, S. R., Graesser, A. C., & van den Broek, P. (1999). *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso.* Princeton, NJ: Erlbaum.

Graesser, A. C, Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*, 163–189.

Knudson, R. E. (1992). The development of written argumentation: An analysis and comparison of argumentative writing at four grade levels. *Child Study Journal, 22*, 167–184.

Knudson, R. E. (1994). An analysis of persuasive discourse: Learning how to take a stand. *Discourse Processes, 18*, 211–30.

Kozmnsky, E. (1977). Altering comprehension: The effect of biasing titles on text comprehension. *Memory & Cognition, 5*(4), 482–490.

Kuhn, D. (1991). *The skills of argument.* New York: Cambridge University Press.

Law School Admission Council. (2001). *10 actual, official LSAT preptests.* Newton, PA: Law School Admission Council, Inc.

Lorch, R. F., & Lorch, E. P. (1995). Effects of organizational signals on text-processing strategies. *Journal of Educational Psychology, 87*(4), 537–544.

Lorch, R. F., & Lorch, E. P. (1996). Effects of organizational signals on free recall of expository text. *Journal of Educational Psychology, 88(*1), 38–48.

Lorch, R. F., Lorch, E. P., Inman, W. E. (1993). Effects of signaling topic structure on text recall. *Journal of Educational Psychology, 85*(2), 281–290.

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9*, 111–151.

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

Means, M. L., & Voss, J. F. (1995). Star Wars: A developmental study of expert and novice knowledge structures. *Journal of Memory & Language, 24*(6), 746–757.

Meyer, B. J. F., & Rice, G. E. (1989). Prose processing in adulthood: The text, the reader, and the task. In L. W. Poon & D. C. Rubin (Eds.), *Everyday cognition in adulthood and late life* (pp. 157–194). New York, NY: Cambridge Academic Press.

Miller, M. (1986). Argumentation and cognition. In M. Hickmann (Ed.), *Social and functional approaches to language and thought* (pp. 225–249). San Diego, CA: Academic Press.

Miller, M. (1987). Culture and collective argumentation. *Argumentation, 1*, 127–154.

Millis, K. K, & Just, M. A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory & Language, 33*(1), 128–147.

National Assessment of Educational Progress. (1996). *NAEP 1994 U.S. history report card: Findings from the National Assessment of Educational Progress.* Princeton, NJ: Educational Testing Service.

National Assessment of Educational Progress. (1998). *NAEP 1998 writing report card: Findings from the National Assessment of Educational Progress.* Princeton, NJ: Educational Testing Service.

National Center for History in the Schools. (1996). *National standards for United States history: Exploring the American experience.* Los Angeles, CA: University of California, Los Angeles.

Perkins, D. N. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology, 77*, 562–571.

Spyridakis, J. H., & Standal, T. C. (1987). Signals in expository prose: Effects on reading comprehension. *Reading Research Quarterly, 22*(3), 285–298.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing* (pp. 53–120). Hillsdale, NJ: Erlbaum.

Stein, N. L., & Miller, C. A. (1993). The development of memory and reasoning skill in argumentative contexts: Evaluating, explaining, and generating evidence. In R. Glaser (Ed.), *Advances in instructional psychology 4* (pp. 285–335). Hillsdale, NJ: Erlbaum.

Toulmin, S. E. (1958). *The uses of argument.* Cambridge, MA: Cambridge University Press.

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

Appendix

Two versions of an excerpt from an argument text used for Experiment 1.

Note: The differences between the two versions are brackets with the Implicit version first, followed by the Explicit version. A double dash indicates that the implicit version does not have additional text.

Hypnosis: Trick or Treatment?

There was nothing remarkable about Victor Rausch's gallbladder operation. Nothing at all except that he underwent the surgery without so much as swallowing an aspirin. Rausch, then a young dentist from Waterloo, Ontario, wanted to see if he could skip the anesthetic and rely on hypnosis to keep himself relaxed and free of pain while his gallbladder was removed. As the surgeon sliced into his abdomen, Rausch entered a trance-like state, and throughout the 75-minute operation Rausch maintained steady blood pressure and pulse rate; he even talked and joked with the surgical team. And implausible as it seems, he swears he felt no pain—only a little tugging. After the surgery was over he stood up, walked down the hall, and rode the elevator to his hospital room.

Sounds like a medical parlor trick? Yes, surgery without anesthesia is a bit of a mind-bender. But the truth is, even in its less startling applications, hypnosis still evokes the image of its sideshow past. Just murmur, "You are getting sleepy, very sleepy," and people envision one of those 1950s mad-doc movies in which creepy old men hypnotize lovely young women to do all sorts of unspeakable things.

Nevertheless, [– / it has become increasing clear that] as researchers learn more about the mind–body connection, hypnosis is sneaking ever so quietly into the medical mainstream. [– / There are several reasons to support conclusion]. Doctors and therapists often use hypnosis to help people quit smoking, lose weight, and overcome phobias—some of the more typical uses of the method. [And / Furthermore,] health maintenance organizations and major insurers are generally willing to pay. [Patients / Additionally patients] are being taught self-hypnosis to ward off asthma attacks and epileptic seizures; and hemophiliacs are using it to stop their own bleeding. [And last summer after reviewing the medical literature, the / Strong support also came after a review of the medical literature in which the] National Institutes of Health concluded that the technique is effective for easing several kinds of discomfort, including headaches and pain associated with cancer.

[It's easy to imagine the / The increased use of hypnosis by the medical profession can be partially explained by looking at its] advantages. Once you become proficient at hypnotizing yourself, you can do it anywhere at anytime. There are no side effects. And it doesn't cost a dime. Such control is a powerful tonic for many patients, even when hypnosis is used as an adjunct to conventional remedies.