Epistemic Levels in Argument: An Analysis of University Oceanography Students' Use of Evidence in Writing

ARTICLE *in* SCIENCE EDUCATION · MAY 2002

Impact Factor: 2.92 · DOI: 10.1002/sce.10024

Gregory J. Kelly
Pennsylvania State University

Epistemic Levels in Argument: An Analysis of University Oceanography Students' Use of Evidence in Writing

GREGORY J. KELLY, ALLISON TAKAO
Department of Education, University of California, Santa Barbara, CA 93106, USA

Received 30 April 2001; revised 3 September 2001; accepted 24 September 2001

ABSTRACT: The purpose of this paper is to examine university oceanography students' use of evidence in writing. Drawing from rhetorical studies of science writing and studies of argumentation in science education, a model for assessing students' arguments is proposed that considers the relative epistemic status of propositions comprising students' written texts. The study was conducted in an introductory university oceanography course in a large public university that utilized an interactive CD-ROM that provided geological data sets for student exploration of scientific questions. Student arguments were analyzed through a process of sorting propositions by epistemic level and identifying the explicit links within and across levels. These epistemic levels were defined by discipline-specific geological constructs from descriptions of data, to identification of features, to relational aspects of features, to theoretically formulated assertions. This form of argumentation analysis allowed for assessment of each student's writing on normative grounds and for comparisons across students' papers. Results show promise for the argumentation model as a methodological tool. The examination of epistemic status of knowledge claims provided ways of distinguishing the extent to which students adhered to the genre conventions specified by the task, i.e., providing evidentiary support for their argument concerning the theory of plate tectonics with real earth data. We draw on the findings to discuss ways argumentation theory can contribute to reform in science education. *Sci Ed* **86:**314–342, 2002; DOI 10.1002/sce.10024

INTRODUCTION

The purpose of this study is twofold. First, we develop a research methodology for analyzing students' written arguments. Second, we apply this method to assess university oceanography students' use of evidence in writing. Our approach to argumentation analysis

---

An earlier version of this paper entitled, "Epistemic levels in argument: An analysis of university oceanography students' use of evidence in writing," was presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA, April 28 – May 1, 2000.

Correspondence to: Gregory J. Kelly; e-mail: gkelly@education.ucsb.edu

Contract grant sponsor: National Academy of Education.
Contract grant sponsor: National Science Foundation.
Contract grant number: NSF Grant # 9952491.

draws on studies of scientific practices, particularly those associated with production of written knowledge (Bazerman, 1988). The model we propose for argumentation analysis considers the relative epistemic status of propositions comprising students' written texts. We draw conclusions based on these analyses to discuss ways of teaching writing and epistemology in science.

ARGUMENTATION IN SCIENCE AND EDUCATION

The fruitfulness of argumentation as a tool for understanding student reasoning, engagement in scientific practices, and development of conceptual and epistemic understandings has begun to get renewed interest in education (Duschl, Ellenbogen, & Erduran, 1999; Kelly, Druker, & Chen, 1998; Kuhn, 1992). Recently, Newton, Driver, and Osborne (1999) made a strong case for the use of argumentation in science pedagogy. However, in their study of secondary schools in the greater London area, they found that students were offered few opportunities to engage in substantive argument. Other applications of argumentation theory have similarly led to mixed results for uses of argument in both studies of classroom discourse and for analysis of written work. Analysis of student discourse in multiple interactional contexts (small group, whole class, student presentations) and across subject matters lends support to the potential of this approach to contribute to reforms proposed for teaching and learning science (National Research Council, 1996). Studies applying argumentation theory to classroom discourse have identified the roles of evidentiary authority in teacher–student discourse (Carlsen, 1997; Russell, 1983), the ways mathematical reasoning can be understood and unpacked (Forman et al., 1998), and the ways students reason about socioscientific issues (Patronis, Despina, & Spiliotopoulou, 1999). Argumentation theory has also been adapted for the analysis of small group conversations of students engaged in practical work. These studies of small group discourse evinced the importance of considering the interactional contexts in which arguments are formed, as social factors (Richmond & Striley, 1996) and assumed understandings (Kelly, Druker, & Chen, 1998) influence the formulation of student arguments.

A second application of argumentation theory has been the consideration of written argument. The production of written texts has played a central role in scientific communities. As studies of the experimental article as a cultural form suggest a crucial role of a relevant community in defining the conventionalized practices that shape the relationships of text and audience (Bazerman, 1988), there is a need to consider the writing of scientific argument as an embodied set of practices tied to particular goals and purposes. Furthermore, the cultural practices associated with the production of such texts have varied with changing mores in scientific communities (Atkinson, 1999; Bazerman, 1988). In educational settings, the writing of substantive argument has been identified as one of the scientific genres advocated by theorists from the writing to learn literature (Keys, 1999). Studies of classroom interaction examining the uses of scientific writing have focused on analysis of students' products and views about science, showing a range of applications of writing to learn and learning to write (Kelly & Chen, 1999; Kelly, Chen, & Prothero, submitted; Keys et al., 1999; Prain & Hand, 1999).

Previous research has applied an argumentation analysis derived from Toulmin's layout of arguments (Toulmin, 1958) to student classroom spoken discourse (Carlsen, 1997; Kelly, Druker, & Chen, 1998; Russell, 1983; for critique see Duschl, Ellenbogen, & Erduran, 1999). However, written discourse offers different challenges and possibilities for analysis (Prain & Hand, 1996). For example, one problem identified with Toulmin's layout of arguments is the ambiguity of the categorical system. Central components of an argument following Toulmin's model are relevant data, a claim asserted by the author, and warrants to be supported by theoretical backing. Yet in the context of actual argument, claims may serve as data in broader, more complex chains of reasoning. This is often the case in extended written arguments like those of the oceanography students presented in subsequent sections of this paper. Furthermore, Toulmin's argumentation approach does not consider the relative epistemic status (i.e., degree of abstractness of knowledge claims) of the speaker's assertions, nor the position of embedded claims in larger arguments. These issues are explored in this study through the examination of students' writing in the context of a university oceanography course.

EDUCATIONAL SETTING

The study was conducted in an introductory oceanography course at the University of California, Santa Barbara. Our choice to focus on this site for the study of science education was based on unique dimensions of oceanography as a discipline as well as on unique features of this particular course. As a discipline oceanography has rarely been examined by social scientists (although see Goodwin, 1995; Mukerji, 1989) and often receives less emphasis than other sciences in the official science curriculum. In addition, oceanography is an inherently multidisciplinary science, drawing from physics, geology, and chemistry, and in interaction with a number of life sciences (Mukerji, 1989).

This particular oceanography course was distinguished by the following features. First, the course was taught through the geological sciences department at the university in question and emphasizes the theory of plate tectonics. The focus on plate tectonics was educationally significant because of the theoretical development associated with these ideas (Duschl, 1990) and because geological sciences have been identified as underresearched in science education (Ault, 1998; Bezzi, 1999). Second, the student population was composed of primarily 1st-year university students, the vast majority of whom were not geological sciences majors. The course functioned as a terminal course for nonscientists and in addition was designed to attract students to the geological science academic major. Third, as identified in a previous ethnographic study (Kelly, Chen, & Prothero, submitted), writing was a key instructional component of this course. The course was designated a "writing intensive course" at this university and required that students write papers following the disciplinary genre conventions defined by the instructors.

Finally, the course included an interactive CD-ROM, "Our Dynamic Planet," created by the course professor (Prothero, 1995). This CD-ROM provided earth data to be used by students to solve problems associated with plate tectonics. These data representations were used by students in their writing of a technical scientific paper. For example, Figure 1 shows a topographical map of the Kurile Trench area, a profile elevation plot for a selected line crossing the trench, and volcano locations along the trench (represented as superimposed triangles). Other relevant cross sections such as earthquakes' depth profile plots could be generated by students at this and other geographical areas and incorporated in their writing. To achieve such representations, students select the cross-section endpoints through graphical "point and click" techniques and can choose the plot scales and vertical exaggeration if desired. Between any two locations, students could plot earthquakes and quake cross sections, seafloor elevation cross-sections, cenozoic volcano locations (on land); determine island ages; measure heat flow; and access movies and still graphics illustrating views or facts about particular locations. This large collection of raw earth data enables students to pose questions about and find support for the theory of plate tectonics, including the identification of plate boundary types through earthquake, volcano, elevation, and heat flow analyses and plate motion velocities through consideration of island ages and hot spots. With the same databases, more advanced studies could be conducted on variations in slab dip (angle of a subducting lithospheric slab from the horizontal plane), the

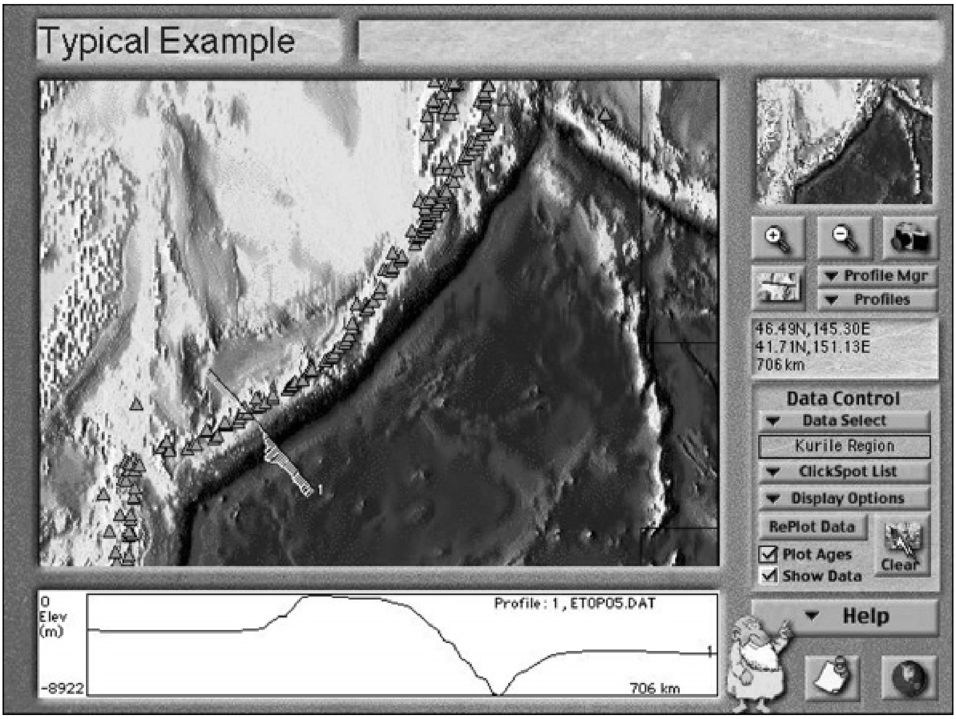

Figure 1. Typical data representation from the CD-ROM "Our Dynamic Planet" employed by students in their scientific writing. This figure shows a elevation profile is plotted along a line specified on the topographical map in the Kurile trench area. In addition, volcano locations are noted as triangles.

configurations of more complex plate boundaries, and comparisons of various plate boundaries at different geographical locations. Thus, the technology offers the possibility of engaging with geological knowledge through investigations, formulating arguments with real large scale data sets, and using specific, relevant data for given locations—tasks common to geological research (Ault, 1998). More information about the CD-ROM may be found at http://oceanography.geol.ucsb.edu/.

The instructional applications of the CD-ROM were focused primarily on student use of data representations to formulate a scientific argument. The students were instructed to write a technical paper, formulated in a scientific genre (Keys, 1999), in which they were to pose researchable questions, characterize several geographical areas with the use of relevant geological data, and reconcile their findings with plate tectonic theory. This served as their midterm for the course and was expected to be 1800 words in length (excluding figures). To assist students in the development of these midterm papers, the course instructor provided three opportunities to learn about scientific writing: discussions about scientific writing, a laboratory manual with instructions regarding the technical paper genre, and peer and instructor feedback on student writing.

The course lectures (given by the course professor) and laboratory sessions (with geology graduate students serving as instructors) included discussions regarding writing. The range of topics regarding scientific writing included how scientists select a problem, how evidence is used to support a theory or model, how observations are separated from interpretations, how these elements are formatted into a scientific paper, and how to generate and use feedback from other writers. Information regarding these topics was communicated through the course instruction focused more broadly on scientific practices that included issues relevant to writing such as uses of evidence, role of expertise, relevance of point of view, and limits to the authority of disciplinary inquiry (see Kelly, Chen, & Prothero, submitted).

The course laboratory manual offered students instructions and examples of the sort of writing expected of them. The laboratory manual presented the students with the expected format of a technical scientific paper including the different sections—listed as abstract, introduction, methods, observations, interpretations, conclusions, figures and captions, and references. Each of these sections included a short description followed by representative samples to illustrate the respective rhetorical tasks. For example, the section on "interpretation" began with "Here is where you take your individual observations and use your experience, insight, and knowledge to explain them. The physical mechanisms that are responsible for the observations should be discussed . . . " The explanation continued noting that "each interpretation must be backed up by one or more observations," and then reviewed issues of conflicts in data interpretation, discussed honesty in reporting results, and offered a number of samples of "interpretative statements" (e.g., "The North American plate is sliding against the Pacific plate at the San Andreas fault.").

The central writing focus of the course was on the use of large scale data sets and plate tectonic theory in the midterm assignment. This assignment was supported by other writing assignments, including a beach observation and a one page prepaper. The student prepapers were shorter versions of the midterm papers (analyzed in detail in subsequent sections), reviewed by their peers in laboratory recitation sections, and evaluated and graded by the graduate student teaching assistants. The students' midterm papers were graded by graduate student teaching assistants following evaluation criteria designed by the course professor. These evaluation criteria took the form of a grading rubric, as presented in Table 1.

DEVELOPMENT OF AN ARGUMENTATION ANALYSIS MODEL: EPISTEMIC LEVELS OF ARGUMENT

The development of the argumentation model presented in this paper was based on two previous studies by Kelly and his colleagues, as well as other scholarship in argumentation analysis and the rhetoric of science. Kelly, Druker, and Chen (1998) applied Toulmin's layout of arguments to the analysis of student dyadic spoken discourse. This study identified the potential uses of Toulmin's method, but surfaced methodological problems. The authors found that organizing student discourse into Toulmin's argument components required careful attention to the contextualized use of language. While the Toulmin model makes distinctions among statements of data, claim, warrant, and backing, among others (for reviews see Ramage & Bean, 1992; van Eemeren, Grootendoorst, & Kruiger, 1987), the scheme is restricted to relatively short argument structures and the argument components pose ambiguities. Statements of claims can serve as a new assertion to be proved, or can be in service of another claim, thus acting as a warrant. This point was not lost on Toulmin (1958) in his original *Uses of Argument*, a book centered on the field-dependence of substantive argument (not on developing heuristics for assessing argument). Thus, the layout of arguments applied in Kelly, Druker, and Chen (1998) need to be expanded to consider the longer chains of reasoning in the university students' writing. In addition, claims within broader arguments can be of different levels of generality, a problem not addressed in Toulmin's layout of argument.

In a study of students' written argument, Kelly and Chen (1999) modified Toulmin's layout of argument by drawing on the work of Latour (1987). Latour's analysis of scientific writing suggested that scientists typically try to move rhetorically from the particular contingencies of their actual experiments (i.e., very specific, grounded claims) to more

TABLE 1
Instructor's Grading Rubric (Columns 1 & 2) and Scores for Selected Students (Columns 3–6)[a]

Grading Rubric		Scores for Four Case Studies			
	Low-Med-High	Brad	Heather	Nick	Katie
Writing					
Clear (readable), focussed, and interesting.	0-1-2-3-4	2	4	2	4
Accurate punctuation and spelling.	0-1-2-3-4	3	4	1	4
Technical paper format is followed. Writings occur within correct section, e.g., interpretations and observations.	0-1-2-3-4	2	3	2	3
All paper sections are included and include the appropriate content.	0-1-2-3-4	1	3	1	3
Data/observations are adequately referenced and described for accuracy.	0-1-2-3-4	1	4	1	4
Inquiry					
A clear, solvable problem is posed, based on an accurate understanding of the theory of plate tectonics, and the inquiry assignment.	0-2-4-6-8	3	8	2	6
Observations are clearly supported by figures which show data and location of data when appropriate.	0-2-4-6-8	3	8	4	8
Clear distinction between observations and interpretations.	0-2-4-6-8	1	1	3	3
Available date are used effectively. Multiple kinds of data are used when available. Data are relevant to the investigation.	0-2-4-6-8	1	2	2	3
Conclusions are supported by the data.	0-2-4-6-8	1	2	1	3
The problem, model, and supporting data are clearly connected in the text and figures. The text clearly explains how the data support the interpretation/sketch. A clear distinction is made between portions of the model that are supported by data and/or background knowledge, and those that are unsupported by data and/or background knowledge.	0-2-4-6-8	2	6	2	7
Figures					
Figures are numbered, clear and easy to read, referred to in the text, and have informative captions.	0-1-2-3-4	3	4	3	4

Continued

TABLE 1
Instructor's Grading Rubric (Columns 1 & 2) and Scores for Selected Students (Columns 3–6)[a] (*Continued*)

Grading Rubric		Scores for Four Case Studies			
	Low-Med-High	Brad	Heather	Nick	Katie
Data/observations are adequately referenced in the text and illustrated in figures.	0-1-2-3-4	1	4	3	4
Interpretations are supported by data and clear sketches. Sketches incorporate data whenever applicable.	0-1-2-3-4	2	4	3	3
	Total	26/80	57/80	30/80	59/80

[a] Brad's paper was graded by teaching assistant one; Katie's by teaching assistant two. Heather and Nick were both graded by teaching assistant three.

generalized statements (i.e., theoretical claims). In this way they "stack" the facts, moving from low induction facts using the pictures, figures, and numbers to progressively higher induction, more abstract assertions (Knorr-Cetina, 1995). Kelly and Chen thus considered the epistemic status of students' claims in their writing and sorted these according to the model presented by Latour. This form of analysis allowed for the consideration of claims at multiple levels of theoretical generality and matched well with the categorical description of transactional use of language in informative writing developed by Britton et al. (1975). However, as suggested by Toulmin, and made apparent by Latour's studies of neuroendocrinologists, understanding substantive scientific arguments requires consideration of the specific constructs of the relevant disciplinary knowledge (Kelly, Chen, & Prothero, submitted; Kelly & Green, 1998; Strike, 1982; Toulmin, 1972). Therefore, in developing this model we needed to consider the disciplinary-specific knowledge in relation to the students' argument structures.

The development of our current model was thus based on previous work analyzing student argument structures and on consideration of disciplinary-specific knowledge in addition to the constraints imposed by the task. From previous work analyzing students' arguments, we learned the importance of considering the rhetorical needs of the academic task. For example, the spoken arguments provided by student dyads while working on electricity problems required a series of relatively short arguments with a limited number of claims (Kelly, Druker, & Chen, 1998). In this instance, the immediate audience was each of the students' laboratory partner. Thus, communication was sparse, and as analysts, we needed to fill in missing information. In the case of the university oceanography midterm papers, the rhetorical requirements of the academic task were considerably more complex. Constraints imposed by the task required that students make reference to large scale data sets, choose appropriate information from these data, and apply the information in an effort to connect to plate tectonic theory. The audience was other students and instructors. Furthermore, the information available in the provided data set (e.g., island ages, earthquake depth profiles) did not by itself speak to plate tectonics. Students needed to formulate a line of evidence based on multiple data representations, either of the same data sort (earthquake profiles along a mountain range) or of multiple types (seismic, volcanic, topographical). The work on formulating such lines of reasoning included using geological concepts, such as ridge, trench, convergent plate margin, and/or subduction. Thus, as a function of the task and the subject matter, students needed to reference specific data, name features, consider how features' geographical orientations were relevant, and use this information to characterize increasingly theoretical abstractions such as plates, boundaries, and zones. We applied these considerations as analysts to sample student arguments in a pilot study. This led to the development of the model presented in Figure 2.

Figure 2 represents our model of argumentation analysis. There are six epistemic levels, from the most specific, grounded claims (shown at bottom of the model) to progressively more general, theoretical claims (shown at top of the model). The epistemic levels are geological in nature from descriptions of data, to identification of topographical features, to relational aspects of geological structures, to use of geological theory or models with identified earth data, to general geological theory, and finally to claims so general as to be independent of the specifics of the argument in question (this last category was looked upon as unfavorable for the student authors by the course instructors). For analysis purposes all the propositions comprising a student's argument were sorted into these six epistemic levels. In addition to the epistemic levels, the model includes explicit links across the assertions made in the argument. This model allows for the use of multiple claims to support complicated arguments. The procedure of sorting the students' writing by proposition into different epistemic levels allows for the examination of the overall argument and the ways that specific references to data are tied to assertions of theoretical interest (e.g., identification of plate interfaces and relative motion). Thus, for each argument, we produce a semantic network showing links across epistemic levels. This allows for assessment of each argument on normative grounds and for comparisons across students' arguments.

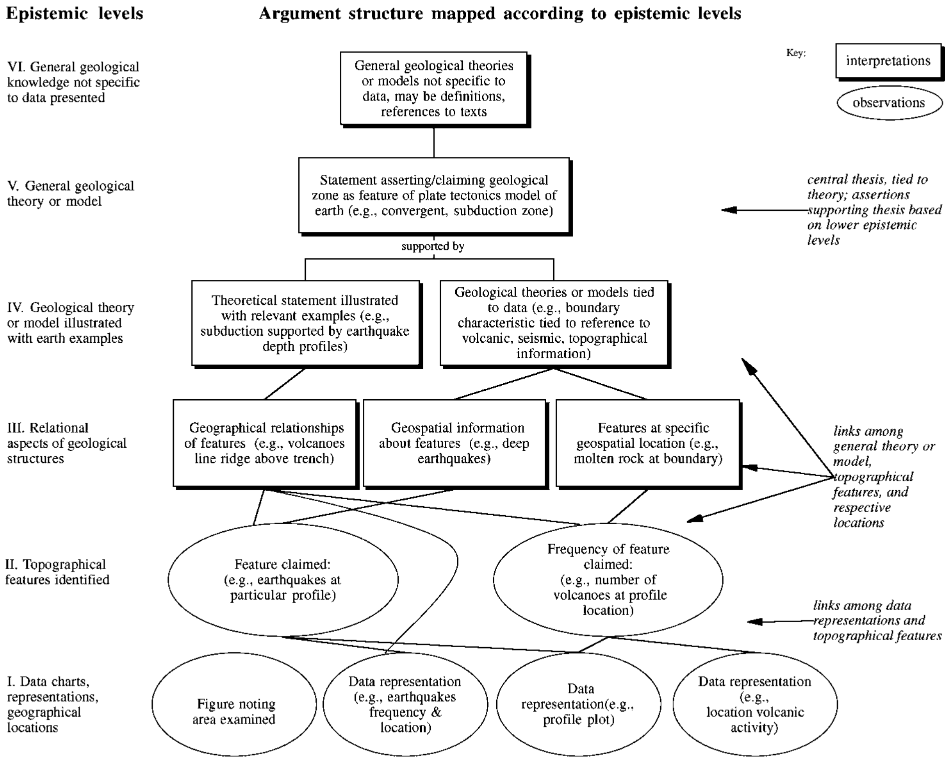

Figure 2. Model of epistemic levels for analysis of students' scientific arguments.

Definitions of the epistemic levels with representative examples are presented in Table 2. These definitions were used for the purposes of sorting the students' statements into the different epistemic levels. Following the "grounded metaphor," we placed the most specific, data-like statements on the bottom of the model (Figure 2) and began our numerical system from this level (Table 2). Several caveats are worth mentioning. First, one premise, based on theoretical considerations, was that student arguments could be arranged in a hierarchy from specific statements grounded in data (epistemic level I) to broader claims that are increasingly general and theoretical (epistemic levels IV and V). Second, the numerical system was designed for labeling and assisting in our rhetorical analysis; the numbers while

TABLE 2
Epistemic Levels for Analysis of Students' Scientific Papers: Definitions and Examples[a]

Category	Definition	Examples
Epistemic level VI	General propositions describing geological processes and referencing definitions, subject-matter experts, and textbooks. The knowledge represented may not necessarily refer to data that is specific to the area of study.	"An oceanic divergent margin means that the plates, which form the Earth, meet and disperse in opposite directions." (Brad, proposition #9)
Epistemic level V	Propositions in the form of geological theoretical claims or models specific to the area of study.	"Continental convergent margins result in earthquakes because the subducting plate fractures under the stress and releases energy due to its folding below the nonsubducting plate." (Nick, proposition #27)
Epistemic level IV	Propositions presenting geological theoretical claims or model illustrated with data specific to geographical area of study.	"The sea floor, which is the Pacific Plate is subducted beneath the more shallow sea floor and island chain of the Eurasian plate." (Heather, proposition #23)
Epistemic level III	Propositions describing relative geographical relations amongst geological structures specific to the geographical area of study.	"Shown in Figure 4 is the presence of over 60 volcanoes along the coast of the trench, reaching a distance inland approximately 230 km." (Katie, proposition #7)
Epistemic level II	Propositions identifying and describing topographical features of the geological structure specific to the geographical area of study.	"Up to 10.5 km marks the deepest recorded depth within the trench which makes it the second deepest known trench in the world." (Nick, proposition #3)
Epistemic level I	Propositions making explicit reference to data charts, representations, locations, and age of islands, or locating the geographical area of study.	"The first particular area observed was found on the eastern coast of Asia (Figure 1)." (Katie, proposition #1)

[a] Examples were taken from four case studies presented.

identifying an ordinal progression do not represent a quantitative measure of generality, nor should they be considered a measure of validity. We expected that a strong argument would use statements at a variety of epistemic levels.

PROCEDURES

Our analysis began by sampling from a large data set ($n = 200$) of student papers. The oceanography course included lectures by the course professor and laboratory sessions with one of three graduate student teaching assistants serving as the instructors. We took a stratified random sample: 8 student papers selected at random from each of the three teaching assistants' laboratory sessions. As identified in previous research (Kelly, Chen, & Prothero, submitted), the genre conventions defined in this class included dividing the paper into preset sections: abstract, introduction, observations, interpretations, conclusions, and figures. Upon review of the student writing we found that much of the inferential work was done in the sections labeled "observations" and "interpretations." Therefore, we centered our analysis on the arguments presented in these sections. Each student was instructed to include data both from one or more geographical areas based on their interest, and from one area common to all students—an area previously considered in homework and class discussions; this year the common area happened to be the west coast of South America. As a sampling procedure, we reviewed the first argument presented by each student, including all the text presented in the observations and interpretations sections as well as the relevant figures cited by the authors. This gave us a range of geographical areas chosen by the students.

The next step in our procedure was to type each of the arguments for the sampled geographical areas from the 24 student papers sampled into computer files. For simplicity's sake our unit of analysis was the grammatical sentence as punctuated by the student writers. For practical purposes, the student marked sentences served as the propositions sorted in the argumentation model. We labeled each sentence with a proposition number for future cross-reference and for identifying the sequential order of propositions as used by the respective authors. We then sorted each student sentence by epistemic level, placing each sentence in a summarized form on a semantic network in the form of Figure 2, retaining the proposition number. However, there were a few cases where students made two claims within a sentence (effectively two propositions); for these cases the sentence was noted at all appropriate epistemic levels. In addition, there were sentences that did not make claims relevant to the evidentiary basis of the argument, but rather provided metadiscursive comments about the argument (e.g., "The mid-Atlantic ridge is a very interesting part of our earth."). Such comments were not included in the model. Each statement was also labeled as derived from either the student-identified "observations" section (represented by proposition number within an oval) or "interpretations" section (represented by proposition number within a rectangle) (see Figure 2). The initial placement by epistemic level was completed by analyst 1 (Takao) and checked by analyst 2 (Kelly). Our decisions were based on the definitions presented in Table 2. All cases of disagreement between analysts were reviewed until a consensus was reached. Next, the explicit links among statements were added including reference using indexical phrases such as "this," or "its." These links were depicted as lines connecting propositions. Implicit links (e.g., sentences following a topic sentence with similar or illustrative content), while recognized as important for coherent writing, were not labeled because these posed prohibitive ambiguity. In subsequent studies we have considered in more detail uses of lexical cohesion in analysis of student arguments (Kelly & Bazerman, 2001; Takao & Kelly, 2001).

ANALYSES AND FINDINGS

After developing this model we conducted a series of exploratory analyses. First, we present some descriptive statistics to give an overview of the model's applications as well as considerations of reliability and validity. Second, based on the results of these quantitative measures, we present four case studies of students' specific arguments. These cases are illustrative of our general rhetorical analysis and provide information about the strengths and weaknesses of the model.

Applying the Model to Student Writing: Some Descriptive Statistics

Our first analysis was to consider the distribution of students' statements in their arguments across the six epistemic levels. We considered the aggregate numbers across all students' papers. Because two papers failed to make a distinction in their writing between observation and interpretation, they were omitted from the presentation of summary statistics—the overall argument structure for these two was not appreciably different from others of the same relative merit. Table 3 presents the distribution of the remaining 22 students' propositions across the epistemic levels. Each level shows how propositions fell into the students' classification of observation and interpretation. Some general trends in Table 3 are worth noting. First, the model was shown to be sensitive enough to distribute and sort propositions across the different epistemic levels. Level I (most descriptive) and level V (argument-specific theoretical) statements had the largest numbers (147 and 173 respectively). Level VI was used the least frequently. This level was so general as to be not specific to the arguments being made—propositions of this sort may serve as background information or to define a construct. Second, the distribution of observation and interpretation statements (defined by the student authors by section titles) varied across epistemic level. The more specific, grounded statements of epistemic levels I, II, and III were composed of more observation statements, while the more interpretative epistemic levels of IV and V included more interpretation statements than observation statements. This shows both the sensitivity of the argumentation model as well as students' understanding of the instructional task: creating arguments grounded in data, building toward theoretical assertions.

TABLE 3
Summary Statistics of Student Propositions Across Epistemic Levels (for 22 of 24 cases)

Epistemic Level	Statement Type	*n*	Percent	Mean	SD	Min	Max	Range
6	Observation	5	0.73	0.23	0.69	0	3	3
	Interpretation	35	5.14	1.59	2.61	0	11	11
5	Observation	25	3.67	1.14	1.98	0	8	8
	Interpretation	148	21.7	6.73	4.68	0	17	17
4	Observation	5	0.73	0.23	0.61	0	2	2
	Interpretation	91	13.4	4.14	3.85	0	16	16
3	Observation	88	12.9	4.00	2.49	0	8	8
	Interpretation	20	2.94	0.91	1.41	0	4	4
2	Observation	111	16.3	5.05	3.26	1	13	12
	Interpretation	6	0.88	0.27	0.63	0	2	2
1	Observation	126	18.5	5.73	1.98	1	8	7
	Interpretation	21	3.1	0.95	1.05	0	4	4

The next quantitative analysis we conducted was concerned with the reliability of the model across raters. To measure this we needed to define what the rhetorical model would predict as a successful scientific argument. Based on the argument structures depicted in Figure 2, we used the following three criteria to define strength of argument as predicted by our rhetorical model: (a) Our first criterion was integration of claims across levels. Our theoretical model suggested that distributing claims across levels would successively build the best case. We therefore rated those arguments stronger that distributed propositions across epistemic levels. (b) Our second criterion was the ratio of data statements to theory/model statements. Our theoretical model suggested that theoretically grounded arguments require sufficient use of data statements. We therefore rated arguments that had multiple data sources (epistemic level I) for theoretical claims (epistemic levels IV and V) stronger than those that made many theoretical assertions with little reference to data. We also believed that extremes in either direction were unlikely to produce strong arguments. (c) Our third criterion was the distribution of observation/interpretation statements across levels. Our theoretical model suggested that placing higher order epistemic level statements in the "interpretation" section of the technical paper denoted appropriate recognition of degree of inference by the student authors. We therefore rated those arguments stronger that had higher observation statement density at lower epistemic levels and higher interpretation statement density at higher epistemic levels. We used all three criteria with equal weight to rank each student argument. There are no numerical absolutes (e.g., an argument must have a certain number of level I data statements per theoretical claim; an argument must not use interpretation statement at level II, etc.). The 24 student arguments were rated best to worst following these criteria by two independent raters (with best scoring 24, worst 1). This was done by analyst 2 (Kelly) and a graduate student researcher (Chen) independent of the project, but familiar with argumentation analysis (as in Kelly, Chen, & Prothero, submitted; Kelly & Chen, 1999; Kelly, Druker, & Chen, 1998). Analyst 1 (Takao) was primarily responsible for sorting each argument into the epistemic levels, and was not included in this reliability study. The Spearman rank correlation between raters was .80, respectable given the complexity of the model.

Our next quantitative analysis concerned the relationship of the assessment of argument strength based on our rhetorical model with the grades given in the course by the geology graduate student instructors. Under the guidance of the course professor, the three graduate student instructors had assessed all the students' papers on the basis of three general categories: writing (including stylistic features of writing, format, labeling); inquiry (including identification of problem, statement of observations, distinction between observations and interpretations, conclusions supported by data); and figures (including norms for labeling and referencing and relationship to interpretations). See Table 2 for additional details. These categories were summed by each teaching assistant to give a total score for each student. For the 24 papers sampled in this study, each of the three components correlated highly with the total score. Pearson correlation coefficients with the total score were as follows: writing (0.94), inquiry (0.98), and figures (0.76). The numerical assignment of grades based on these criteria varied across instructors. Analysis of variance of instructors' total score for the 24 student papers in the stratified random sample showed significant differences (F-ratio $= 4.6$; $p \leq 0.022$) across instructors. There were no differences based on students' gender, for either the instructors' total score or the ranked score of quality based on the argumentation model.

We compared the instructors' score with the argument strength rank from the epistemic levels argumentation model in the following manner. On the basis of the numerical total point score, we created a ranked instructors' score for the students' papers (again using 24 for highest, 1 for lowest). The Spearman rank correlation of ranked instructors' score and ranked argument strength from the argumentation model was low ($r_{\text{ranks}} = 0.12$) and not statistically significant. In some respects this result is not surprising given the differences in assessment criteria, and sections of the papers examined. Nevertheless, these differences led us to reexamine the two assessment schemes (i.e., that of the instructors and that of the argumentation model based on epistemic levels). We constructed a scatterplot of ranked instructors' score by ranked argument strength from the rhetorical model to identify differences for particular cases (Figure 3). Clearly the low correlation and the scatterplot indicate that the instructors' rubric and the rhetorical argumentation model show differences in many cases for the assessment of argument strength. We use these differences to identify weaknesses in the argumentation model, but also in the instructors' rating. To consider how the two systems of assessment differed in detail, we present four cases. These cases were chosen strategically from the scatterplot shown in Figure 3. We chose one case (Brad) rated low by both systems, and one case (Heather) rated high by both. We then turned to discrepant cases. In one case (Katie), the instructor rated the student's writing high, while the raters using the argumentation model rated it low. In the other case (Nick), the instructor rated the student's work low, while the raters using the argumentation model rated the same work high in terms of quality. The data points for the four cases (Brad, Heather, Katie, and Nick) are labeled in Figure 3 for cross-reference purposes.

Case Studies: Unpacking Particular Arguments

In this section we present the argumentation structure for 4 of the 24 cases. We begin with the case of Brad, rated low by his graduate teaching assistant and by the epistemic levels model. The following represents the actual text presented by Brad in his paper as his observations and interpretations for the region of the mid-Atlantic ridge—the first area chosen by the student for presentation and thus sampled by us. We have added the numbering to his propositions.

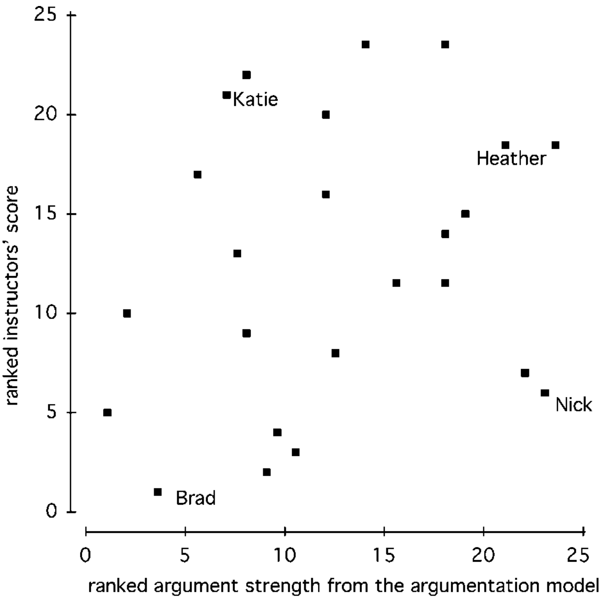

Figure 3. Scatterplot of ranked instructors' score vs. ranked argument strength from argumentation model.

Brad's observations and interpretations for the mid-Atlantic Ridge

> Observations: Mid-Atlantic Ridge
> (1) The Mid-Atlantic Ridge is a very interesting part of our Earth. (2) It is an underwater mountain range, also known as an oceanic divergent margin. (3) This ridge runs north to south down the center of the Atlantic from the North Pole to Antarctica. (4) Many different plates meet at the ridge including the North American, the Eurasian, the South American, and the African Plate. (5) The ridge extends at one point as deep as 5625 meters below sea level. (6) It stretches east to west from Europe and Africa to the east coast of the Americas, about 2547 kilometers. (7) This is evident in Figure 1. (8) Figure 1: Oceanic Divergent Margin/Mid-Atlantic Ridge between South America & Africa.
>
> Interpretations: Mid-Atlantic Ridge
> (9) An oceanic divergent margin means that the plates, which form the Earth, meet and disperse in opposite directions. (10) The resultant gap from these diverging plates is filled up with uprooted, low density magma. (11) This process leads to the series of volcanoes which form into a ridge in the gap left by the plates. (12) This process is known as seafloor spreading. (13) This is also illustrated in Figure 1. (14) The aging crust then sinks steadily down, while the mountains in the ridge slowly move outward while new ones fill in their place. (15) The mountains move in the direction of the plate. (16) This part of the process, combined with narrowness of the Atlantic and the shape of the continents, leads to the S shape formed by the ridge.
>
> Argument-specific figures cited:
> Figure 1: Oceanic Divergent Margin/Mid-Atlantic Ridge between South America & Africa. Shows geographical area and cross-section elevation plot.

Each of these propositions were sorted by epistemic level and mapped as shown in Figure 4, based on criteria presented in Table 2. The introductory comment in Brad's observations

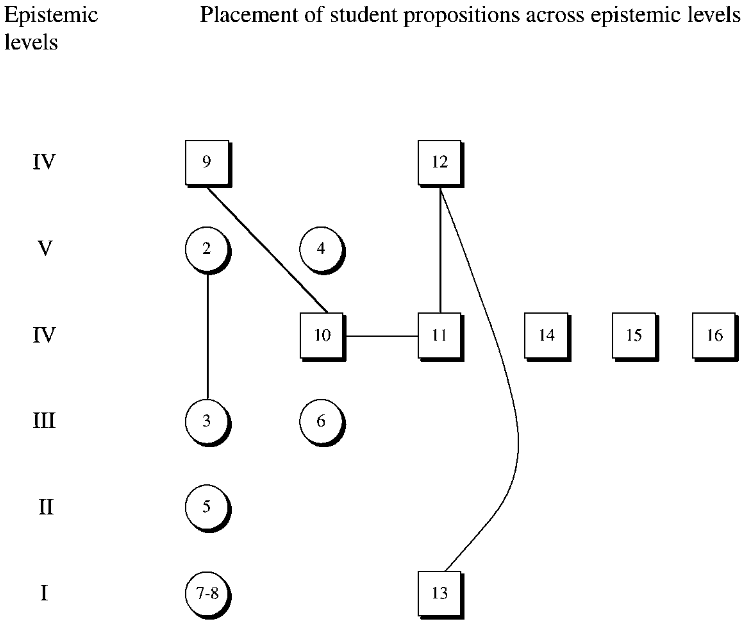

Figure 4. Argumentation structure by epistemic levels for Brad (rated low). Sentences specified by students as "observations" are depicted as circles, "interpretations" as squares.

was not classified, as comments of this sort did not contribute to the evidentiary basis of the argument. The second proposition referring to how the mid-Atlantic ridge is an underwater mountain range, known as an oceanic divergent margin, was classified at the fifth epistemic level, i.e., referring to general geological knowledge. Proposition 3 locates the mid-Atlantic ridge and thus was classified at the third epistemic level, i.e., relational aspects of geological structures. Geographical information is similarly offered in proposition 6. Proposition 5 identifies a topographical feature (ridge depth) offering some information (5,625 m below sea level). This proposition was classified at epistemic level II: topographical feature identified. Explicit reference to data representations (epistemic level I) does not enter the exposition until propositions 7 and 8. These propositions point to a cited figure which shows geographical area and a cross-section elevation plot.

Brad's interpretations fell into the three more theoretical epistemic levels. This is consistent with the general distribution across the 24 student papers (Table 2) and is consistent with the implied inference level of the assignment. Propositions 10, 11, 14, 15, and 16 use geological theory, citing specific examples from the geographical area. These were classified at epistemic level IV in Figure 4. Propositions 2 and 4 refer to geological theory, in a way more abstract and general than propositions 10, 11, 14, 15, and 16, and were thus classified at the fifth epistemic level. Proposition 13 cites a sketched figure, classified as epistemic level I. Finally, propositions 9 and 12 define divergent margins and name seafloor spreading respectively, and were classified as the most general sort of claim—claims referencing geological theory, but not specific to the argument being made.

Figure 4 is a schematic representation of Brad's argument. This argument was rated low by both raters using the epistemic levels rhetorical model. Using the criteria of the epistemic model (i.e., integration of claims across levels, ratio of data statements to theory/model statements, distribution of observation/interpretation statements across levels), Brad's argument shows considerable weakness. There are few statements referencing data representations, suggesting the argument may not have sufficient evidentiary support for the theoretical assertions. Furthermore, there is a poor integration of claims across levels; the argument looks top-heavy with many more theoretical statements than data statements. Substantively, Brad was trying to establish the reasons why the mid-Atlantic ridge has formed into an S-shape. He referenced very limited data to make this argument. He made use of an elevation profile plot across the mid-Atlantic ridge at one latitude (proposition 8). He added to this information about the ridge depth (proposition 5) and the geographical location of the ridge (propositions 3 and 6). These data points are not tied to the set of theoretical assertions and thus fail to make a strong case for the shape of the mid-Atlantic ridge.

Brad's argument was similarly rated low by his graduate student teaching assistant (see Table 1 and Figure 3). Although he was marked low across all areas of the grading rubric, four substantive areas were noted as particularly insufficient: distinction of observations and interpretations, effective use of available data, conclusions supported by data, and data adequately referenced in the text (see Table 1). In this case, we see that the epistemic model of argumentation offered a similar assessment to the graduate student teaching assistant. This lends some credibility to the ability of this model to pick out weaknesses in students' arguments.

We now turn to Heather's argument which was ranked high by both raters using the epistemic levels argumentation model (see Figure 3). Figure 5 shows the distribution of her propositions across the epistemic levels. The actual text of her argument is presented in the Appendix. Based on the criteria of the epistemic levels model, Heather constructed a rhetorically sound looking argument. She referenced data representations in seven propositions (1, 3, 7, 10, 16, 18, and 26). She made use of these propositions in the next two epistemic levels by identifying the topographical features of the data representations (propositions 4,

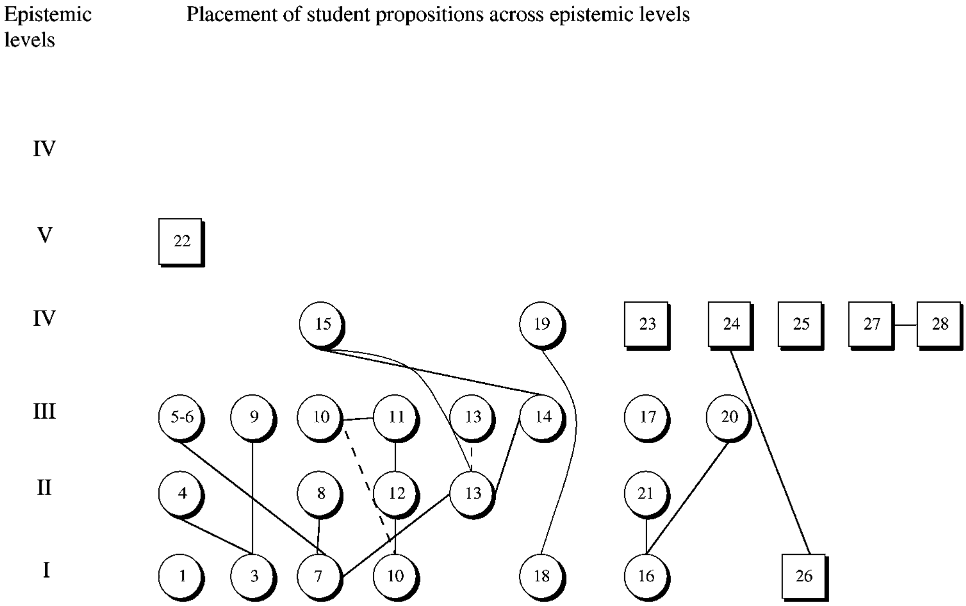

Figure 5. Argumentation structure by epistemic levels for Heather (rated high). Sentences specified by students as "observations" are depicted as circles, "interpretations" as squares.

8, 12, 13, and 21) and by demonstrating how these features were geographically orientated (propositions 5, 6, 9, 10, 11, 13, 14, 17, and 20). She has a number of theoretical assertions at levels IV and V (15, 19, 22, 23, 24, 25, 26, 27, and 28) and did not need to reference the very general theories usually found at level VI. She also made explicit ties across the levels. Arguments of this sort demonstrate how analysis by epistemic levels allows for understanding how the author "stacked the facts" (cf. Kelly & Chen, 1999; Latour, 1987). Substantively, Heather attempted to make the argument that the Kurile Island chain is a subduction zone at a convergent tectonic boundary. She referenced data referring to topographic (propositions 1, 3, 7, 10, and 12), seismic (propositions 16 and 18), and volcanic properties (propositions 13 and 14) of the specified area. She drew these data into the argument by using geological constructs (e.g., subduction, plate movement, lithosphere) to show how together they support her central assertion. Her argument is thus more extensive and extends more thoroughly across epistemic levels than that offered by Brad.

Heather's argument was similarly rated high by her graduate student teaching assistant (see Table 1 and Figure 3). She received high marks for her definition of problem, support for observations with figures, and ways she connected the geological model to the supporting data. However, she was rated low in three areas. She was interpreted as not making clear distinctions between observations and interpretations (scored 1 of 8 points, see Table 1). In addition, she was cited as not using data effectively nor sufficiently supporting her conclusions with data. These last two criteria were central to the academic task and thus the total score (assessed as high by the teaching assistant) seems surprising given these liabilities.

The next two arguments, written by Nick and Katie, represent cases where the review from the point of view of epistemic levels differed from the assessment by the graduate student teaching assistants. These cases pose validity questions for both the argumentation model, and perhaps the course grading schemes. Nick's argument is presented in Figure 6 (actual text in the Appendix). The map of Nick's argument shows many structural similarities to that of

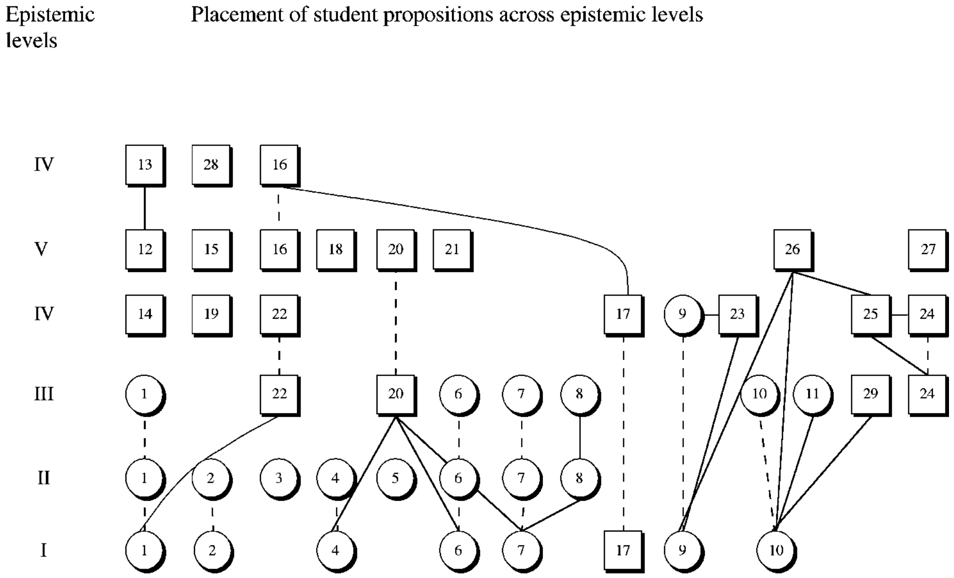

Figure 6. Argumentation structure by epistemic levels for Nick (mixed rating). Sentences specified by students as "observations" are depicted as circles, "interpretations" as squares.

Heather. Nick referenced seven data representations, formulated propositions across all six epistemic levels, and made many explicit ties across levels. His argument was rated higher than Heather's by both raters using the epistemic levels argumentation model. Substantively, Nick, like Heather, chose the Kurile Trench as the location for his analysis. In support of his argument, Nick referenced data referring to topographic (propositions 1, 2, 4, 6, and 7), seismic (proposition 10), volcanic (propositions 1 and 6), and heat flow (proposition 9) properties of the specified area. His argument made use of geological constructs (e.g., relative plate movement, subduction, energy), especially across two theoretical epistemic levels (IV and V): geological theory and geological theory illustrated with examples.

Interestingly, Nick's argument was assessed as considerably less quality by the same graduate student teaching assistant (30/80 for Nick vs. 57/80 for Heather)—the grader that had the lowest overall numerical scores of students' work. The largest differences in the assessment of the two papers based on the grading rubric were found in the following areas (see Table 1): Heather was seen to have a clearly defined problem; to have used observations clearly supported by her figures; and to have a problem, model, and supporting data clearly connected in the text and figures. In each of these areas she was rated higher than Nick by the same graduate teaching assistant.

The case of Katie also raises questions about the differences in the two assessments of student arguments. Figure 7 depicts Katie's argumentation structure by epistemic level. Her argument was rated 15th and 20th from the best by the raters using the rhetorical model of epistemic levels. From the point of view of the argumentation model, the argument has some drawbacks. First, the argument is sparse compared to that of Heather and Nick. (Note: In aggregate, the total number of propositions did not correlate with argumentation score, so there was not systematic bias in favor of longer papers.) There are few data references; of her six cited figures, one located the geographical area, two others were sketches, and three referenced data. The data representations referred to the trench location and depth (proposition 3), earthquake location and depth (proposition 4), and location and number of volcanoes (proposition 7) in the area. Second, there is only one reference to the relative

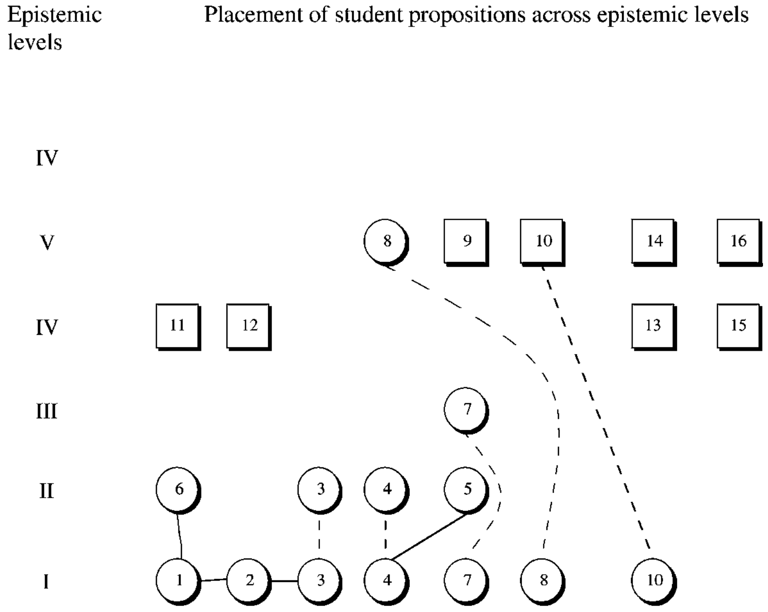

Figure 7. Argumentation structure by epistemic levels for Katie (mixed rating). Sentences specified by students as "observations" are depicted as circles, "interpretations" as squares.

location of the geological features (proposition 7). This is not necessarily a serious flaw—Katie did focus on a very specific area and examined elevation, earthquakes, and volcanic activity in this one area. Taken together, the evidence she brought to bear in her argument was more limited than that used by either Heather or Nick according to the argumentation model. Substantively, Katie attempted to argue that the eastern coast of Asia represents a plate boundary that can be considered a convergent margin. Her argument was constructed by considering the relative motion of two plates and the locations of earthquakes and volcanoes.

Of the four cases presented in detail in this paper, Katie's was graded highest (by her instructor who was not the same teaching assistant as for the other papers). She received high marks in a number of the "inquiry" criteria (see Table 1): posing a clear solvable problem; supporting observations with figures; making connections of problem, model, and supporting data in the text and figures. Thus, like the case of Nick, this argument represents a discrepant case. On one hand, the epistemic levels argumentation model was not able to make the distinctions of the instructors. One the other hand, the graduate student instructors differed significantly in their overall assessments and may have inconsistent ratings across instructors. This suggests possible weaknesses in the argumentation model, but also possible applications for offering an independent validity check on instructors' assessment of writing.

Much like the oceanography students, for these cases of ambiguity, we explored further evidence. The differences in assessment of student arguments by the graduate student instructors and our model led us to request a review by the course professor. We requested that he grade the papers following the same rubric originally designed by him and used by the instructors. He did this without knowledge of either of the other assessment scores. The results for the four cases across the three raters (teaching assistants, argumentation model, and course professor) is presented in Table 4. The addition of this third assessor did not make the picture clearer. The course professor, like the raters using the argumentation

TABLE 4
Comparison of Ratings of Four Student Papers Chosen for Case Studies

	Overall Teaching Assistant Rank Score[a] (and Point Total)		Average Rank Score[b] by Raters Using Epistemic Levels Model		Course Professor Rank[c]	
Student	4 Cases Among 24 Total	Among 4 Cases	4 Cases Among 24 Total	Among 4 Cases	Point Total	Among 4 Cases
Brad	5 (26)	Worst	1	Worst	26	Worst
Heather	19 (57)	Next-to-best	21	Next-to-best	60	Next-to-worst
Nick	6 (30)	Next-to-worst	23	Best	80	Best
Katie	21 (59)	Best	7	Next-to-worst	71	Next-to-best

[a] Rank score was calculated based on total point score assigning highest at 24, lowest at 1.

[b] Raters using the epistemic levels rhetorical model ranked the students' arguments, best to worst, assigning highest at 24, lowest at 1.

[c] Course professor ex post facto ranked four special cases after assigning total point score based on same rubric as course teaching assistants.

model, ranked Brad lowest and Nick best (in contradiction to the teaching assistant assessment). However, in contradiction to the assessment using the argumentation model, the course professor concurred with the teaching assistant about Katie's paper, ranking it second-to-best. Differences such as these in interrater agreement are common in assessing writing (Wolcott & Legg, 1998), but offer opportunities to identify varying interpretation of criteria and to pose new questions for exploration.

DISCUSSION

Our discussion of the application of the argumentation model is divided into two parts. We begin by discussing the strengths and weaknesses of our epistemic levels argumentation model: we consider what was learned from our empirical analysis. We then turn to broader issues of how theories of argument can contribute to thinking about reform in science education.

Epistemic Model of Argument: Strengths and Weaknesses

Our analysis of the 24 university students' arguments following our epistemic levels argumentation model demonstrated the potential of this approach for understanding students' use of scientific theory and data. Our model allows for schematically representing complex arguments composed of multiple claims of various degrees of generality. This is an improvement of the applications of the Toulmin method (e.g., Kelly, Druker, & Chen, 1998). Our model applies Latour's notion of the rhetorical changes of claims over time (Latour, 1987), but considers more carefully uses of claims at different epistemic levels within a written text. The empirical results indicate that the university students engaged in writing in ways adhering to specific disciplinary practices (Kelly, Chen, & Prothero, submitted; Keys, 1999). For example, they used inscriptions, formulated chains of reasoning, and considered evidence relevant to posed questions (Duschl, 1990; Roth & McGinn, 1998). Our analysis demonstrated a sensitivity to problems and strengths of student arguments (see in particular those of Brad and Heather, in Figures 4 and 5).

In this first empirical application, we identified weaknesses in our model by focusing on how subject-matter experts differed in their assessment of the students' arguments as compared to the predictions of our model. For example, our model assessed Nick's argument (Figure 6) as comparable to that of Heather (Figure 5). This was at odds with the grades assigned by the instructor. Interestingly, in his subsequent analysis following the same grading rubric, the course professor viewed Nick's argument as significantly better than Heather's—at odds with his teaching assistants, but more consistent with our argumentation model. As shown in Table 4, the professor's assessment of Katie's paper showed greater agreement with the course graduate student instructor than with our argumentation model. The argumentation model rated Katie's argument low because of the relatively sparse data referenced. In sum, our differences with the subject-matter experts points both to lack of consistency on their part, and more importantly for us as analysts, to a need to develop further the sensitivity of our argumentation model. In particular, the model needs to consider the ways that students use evidence given the scope of their argument. For example, Katie's focused argument may have led us to view it as disproportionally weak.

A second theoretical weakness was the lack of direct consideration of the inference logic of the authors. While the model identified the data and theoretical claims, the inferential connections were not sufficiently assessed. This may lead to characterizing arguments as evidentially supported when the key inference rules are not being followed. The relationship of propositions in an argument and inferences forming the logic of the argument pose considerable methodological problems. First, such assessment of inference requires expert subject-matter knowledge, and even at the level of scientific expertise, there may be serious disagreements. Second, coherent paragraphs may use assumed inferences based on judgments of reader experience (Bazerman, 1988; Halliday & Martin, 1993). These assumed tacit agreements on the part of the writer and readers may not give argumentation analysts readily available means to assess these inferences.

A third weakness concerns the inferences made by us as analysts. We examined the students' writing considering the propositions abstracted from the intentions of the writers. Our model can be further validated by examining how writers themselves considered their claims for a given writing task. Interviews with student writers may indicate that the extent to which the generality of student claims in their arguments were a function of the rhetorical task set by the assignment, rather than features of scientific writing generally. In an initial study considering students' perspectives, we used samples of other students' writing as a prompt to provide a context to discuss uses of evidence with student interviewees (Takao & Kelly, 2001). Results indicate that students were generally better at identifying stronger arguments than providing detailed reasons for their choice. Nevertheless, further research assessing writers' point of view concerning their arguments would begin to address this weakness.

Argumentation Theory, Science, and Knowledge

While current reform initiatives have identified the importance of developing communication and argumentation skills among students (e.g., National Research Council, 1996), few studies have considered the experimental article as a written genre and examined how students use evidence in this form. The students in this study participated in a university course that required the writing of a technical paper and provided the students access to geological data bases through the use of an interactive CD-ROM. The data modules provided students with earth data to solve many problems associated with plate tectonics.

Plate boundary types (using earthquake, volcano, elevation, and heat flow data) and plate motion could be determined (using island ages and hot spots information) with this technology (Kelly, Chen, & Prothero, submitted). This educational context then provided an excellent research site to examine how students can use scientific data as evidence in written forms. Furthermore, geological inquiry has unique features that are discipline specific (Ault, 1998). Examination of students' engagement with these unique ways of inquiring requires a methodology that is both attuned to the theoretical knowledge of the discipline and recognizes the epistemological constraints of such inquiry (Bezzi, 1999).

In considering which aspects of geological inquiry and primary sources merit inclusion in science instruction, Ault (1998) derived a set of discipline-specific criteria of excellence for geology education. These criteria are as follows: constraints on ambiguity; independent, converging lines of inquiry; proper taxonomies; extrapolating systems through time; and integration across temporal and spatial scales. Student generation of scientific argument, derived from real earth data and based in geological theory, offers one means for operationalizing such criteria in school situations. For example, Nick (Figure 6) presented an argument characterizing the Kurile Trench as a convergent continental margin. He developed independent, converging lines of inquiry drawing on information about topographical features (trench, basin), heat flow plots, earthquakes, and volcanoes. He imposed constraints on the ambiguity posed by the region's features by considering the effects of plate movement and the changes over time these movements would cause (e.g., subducted material heats up, plates move through convection, destruction of subducting plate forms back-arc basin). The use of these and other geological constructs shows knowledge of relevant taxonomies of the field. Therefore, while the criteria for excellence in geological inquiry cannot be comprehensively assessed through argumentation analysis (or any other one method we suspect), examination of student writing provides some insight into how students' understand these inquiry processes. By sorting Nick's and other students' arguments by epistemic level we were able to examine how they constructed evidence for their respective cases and engaged in some of the practices and knowledge suggested as central to geology by Ault (1998).

CONCLUSION

The epistemic levels argumentation model we proposed in this study offers a number of potentially fruitful avenues for assessment of student discourse, both spoken and written. The model helped identify the ways students align data in support of their central claims and allowed analysts to create a visible representation of the reasoning manifest in the overall argument. Applied to specific cases in a university oceanography course, the model enabled us to distinguish the extent to which students adhered to the genre conventions specified by the task, namely to develop an evidentiary supported argument concerning the theory of plate tectonics with real earth data.

Further research needs to consider four improvements of the model. First, a method of assessing the inferential logic of the students' claims needs to be more thoroughly developed. Textual analysis of the arguments with a thorough consideration of lexical cohesion (Halliday & Hasan, 1976; Hoey, 1983) may offer contributions in this direction (Kelly & Bazerman, 2001; Takao & Kelly, 2001). Second, interviews with students concerning their writing may provide insight into the intentions of their rhetorical moves; this was discussed above. Third, the issue of generalizability of the model needs to be investigated. The argumentation model presented was based on a general pattern in scientific texts that demonstrates the rhetorical processes of moving from specifics to theoretical statements (Latour, 1987). These processes have been identified across disciplines. For example, Myers' study of two biology texts considered how the extent of the claims made by the scientist authors was tempered through negotiations with reviewers and editors (Myers, 1997). Nevertheless, considerations of epistemic level of claim will undoubtedly have to be interpreted within the specific contexts of each sociohistorical situation (Atkinson, 1999; Bazerman, 1988), even if a pattern for science texts in general emerges. Furthermore, it is important to note the limitations of considering the rhetorical task facing the students in this study as a representation of geology writing. The specific task framed by the course activities may be more central than any generalized notion of writing in geology. In sum, it remains to be seen if the model is effective in other disciplines or with other rhetorical tasks within geology.

Finally, the possible pedagogical implications of the model need to be evaluated. In this first study, the epistemic levels argumentation model was not used for instruction. However, the model may have considerable utility as a writing and assessment heuristic. By making explicit the rhetorical processes of argument formulation, this model may help students evaluate the merits of others' arguments and assist them in the construction of their own arguments. Furthermore, the analysis of claims at different epistemic levels may offer a means for standardizing the assessment by the instructors. We hope to explore these directions in further work in this area.

APPENDIX: TEXTS OF STUDENT WRITING USED AS BASIS FOR RHETORICAL ANALYSIS

For the sampled texts from the four cases studies, Brad, Heather, Nick, Katie, numbers were assigned to each sentence for the sections analyzed in detail with the argumentation model. Figure captions relevant to analyzed text are presented; actual student figures while used in the analyses were omitted in this presentation for space considerations.

Case 1

Student pseudonym: Brad

Title: "Plate Tectonics"

Introduction . . .

Methods . . .

Observations: Mid-Atlantic Ridge
(1) The Mid-Atlantic Ridge is a very interesting part of our Earth. (2) It is an underwater mountain range, also known as an oceanic divergent margin. (3) This ridge runs north to south down the center of the Atlantic from the North Pole to Antarctica. (4) Many different plates meet at the ridge including the North American, the Eurasian, the South American, and the African Plate. (5) The ridge extends at one point as deep as 5625 meters below sea level. (6) It stretches east to west from Europe and Africa to the east coast of the Americas, about 2547 kilometers. (7) This is evident in Figure 1. (8) Figure 1: Oceanic Divergent Margin/Mid-Atlantic Ridge between South America & Africa.

Interpretations: Mid-Atlantic Ridge
(9) An oceanic divergent margin means that the plates, which form the Earth, meet and disperse in opposite directions. (10) The resultant gap from these diverging plates is filled up with uprooted, low density magma. (11) This process leads to the series of volcanoes which form into a ridge in the gap left by the plates. (12) This process is known as seafloor spreading. (13) This is also illustrated in Figure 1. (14) The aging crust then sinks steadily down, while the mountains in the ridge slowly move outward while new ones fill in their place. (15) The mountains move in the direction of the plate. (16) This part of the process,

combined with narrowness of the Atlantic and the shape of the continents, leads to the S shape formed by the ridge.

Observations: Subduction Zones . . .

Observations: Small Area Description . . .

Interpretations:

Conclusions . . .

Argument-Specific Figures Cited:

Figure 1: Oceanic Divergent Margin/Mid-Atlantic Ridge between South America & Africa. Shows geographical area and cross-section elevation plot.

Figures 2–3: Data for South American plate and Eastern Russian seaboard.

References

Case 2

Student pseudonym: Heather

Title: "Plate tectonics and the Pacific Plate: A discussion of the Kurile Island Chain, the East Pacific Rise and the San Andreas Fault"

Abstract . . .

Introduction . . .

Methods . . .

Observations: Kurile Small Area
(1) The Kurile Island Chain, part of the Eurasian Plate, is located Northeast of Japan and lies at the boundary zone between the Eurasian and Pacific plates (Figure 1). (2) To discern which type of tectonic boundary is represented in this area, I considered the topographic, volcanic and seismic properties of the area. (3) Figure 2 offers a closer view of the area to be considered. (4) The most notable topographic properties of this area are the trench, island chain and back-arc basin. (5) Notice the long, clearly defined trench along the East edge of the Eurasian plate, skirting the Kurile Island chain. (6) This is called the Kurile trench. (7) A more precise profile of the trench is plotted between 44.98N, 147.97E and 43.07N, 150.39E (Figure 3). (8) The depth of the trench at this profile reaches −8908 meters at the bottom (notice graph in Figure 3). (9) The island chain follows the shape of the plate boundary along the trench (Figure 2). (10) There is what is called a back-arc basin to the West of the island chain (detail, Figure 4). (11) This back-arc has the characteristic magmatic island arc and trench surrounding its shallow basin. (12) The profile in Figure 4, taken at 47.46N, 143.06E; 47.52N, 153.97E; indicate that at this location, the depth of the basin is a fairly constant −3360 m deep. (13) The Smithsonian Institution Global Volcanism Program and the corresponding dataset provides clear evidence of extensive volcanic activity on the rim of the Kurile Island chain, to the West of the trench, on the Eurasian plate as opposed to the Pacific (Figure 3). (14) The volcanic activity occurs approximately 96 km from the trench, in most cases. (15) Notice that the volcanic activity supports the notion that this island chain is indeed magmatic, and therefore characteristic of a back-arc basin. (16) The dataset from NEIC provides information on seismic activity in the region (Figure 5). (17) There is a notable tendency for extensive earthquake activity along the boundary zone along the island chain, occurring both in the trench area and to the west of it, including the islands and the subsiding nearer the back-arc basin. (18) For a more precise look at the earthquake activity in this small area, I updated the profile in

Figure 5 to include a detail of the earthquake activity along the profile. (19) This provided a remarkable example of a Benioff Zone, with its characteristic earthquake tending down and *towards* the plate, in this case tending toward the island chain. (20) Notice, in the graph in Figure 5, that the earthquake range in magnitude between 5 and 8 points on the Richter scale, and occur at shallower depths near the trench and increasingly deeper depths as they occur nearer the continental plate. (21) The graph suggests that the range in depth of these quakes runs from 1–147 km deep.

Observations: East Pacific Rise and San Andreas Fault Areas . . .

Interpretations: Kurile Island Chain Small Area
(22) This tectonic boundary seems to be a convergent boundary, and indeed a subduction zone. (23) The sea floor, which is the Pacific Plate is subducted beneath the more shallow sea floor and island chain of the Eurasian plate. (24) To come to this conclusion, I recognized that the characteristic deep trench with the corresponding earthquakes suggestive of a Benioff zone, point to the fact that the Pacific Plate sea floor is heavier than the continental Eurasian plate. (25) The force of converging with the Eurasian Plate literally drives the Pacific Plate downward where the crustal material re-enters the lithosphere and is, in effect, recycled. (26) See Figure 15, in which a cross-section of this zone is sketched. (27) The subduction zone is clearly marked and likewise, the back-arc basin which is another remarkable characteristic of this particular convergent boundary. (28) The volcanic activity helps to create this comparatively shallow basin, while it further evidences the subduction zone, convergent boundary.

Interpretations: East Pacific Rise and San Andreas Fault Areas . . .

Conclusions . . .

Figures:

Figure 1: World Map. Location of Kurile Island Chain and East Pacific Rise/San Andreas Fault Small Areas indicated.

Figure 2: Small Area Map of Kurile Island chain indicating topographic characteristics.

Figure 3: Kurile small area map with volcanic activity highlighted and a profile graph of the Kurile Trench at: 44.98N, 147.97E; 43.07N, 150.39E.

Figure 4: Kurile small area map including profile of the back-arc basin and corresponding graph at: 47.46N, 143.06E; 47.52N, 153.97E.

Figure 5: Kurile small area map with earthquake activity highlighted and graph of plotted profile, indicating Benioff Zone at: 4.98N, 147.97E; 43.07N, 150.39E.

Figures 6–14, 16: Data for East Pacific Rise and San Andreas Fault.

Figure 15: Sketch of cross-section of Kurile trench.

Case 3

Student pseudonym: Nick

Title: “Pacific Rim of Fire?”

Abstract . . .

Introduction . . .

Methods . . .

Observations: Kerile [Kurile] Trench
(1) Three profiles taken along the coastal region of the Khamchatka Peninsula display the topographic features of an oceanic trench and the thousands of volcanoes that exist

200–400 km inland of the trench (Figure 3). (2) The trench lies at 60 deg. N latitude and 160 deg. E longitude and extends for 2,200 km in length along this coast. (3) Up to 10.5 km marks the deepest recorded depth within the trench which makes it the second deepest known trench in the world. (4) One profile displays the gentle upward slope of the pacific Ocean Basin which becomes drastically altered by the sudden drop-off area of the trench (Figure 4). (5) Following the trench, a virtual linear rise occurs as the profile moves northwest and thus inland. (6) A second profile confirmed the presence of the trench 500 km to the south of the first profile, but added an accurate depiction of a 400 km long basin located behind the vertical rise of the volcanoes which dips 3000 m below sea-level (Figure 5). (7) A third profile affirms both the existence of the trench another 250 km to the south and the land features described by the first two profiles (Figure 6). (8) However, the vertical rise of volcanoes following the abyss of the trench maintains a steeper slope and the basin behind this rise of volcanoes becomes only 200 km long (Figure 6). (9) A heat flow plot taken along the middle topographical profile path indicates the upward and inland movement of heat flow coming away from the oceanic trench (Figure 7). (10) By plotting earthquakes' foci along the same path as the middle topographical profile of the Khamchatka coast, the plot shows earthquakes occur consistently along this trench (Figure 8). (11) Furthermore, the depths of the foci drop to below 100 km as the profile path proceeds farther inland from the trench (Figure 8).

Observations: Hawaiian islands and Emperor Seamounts . . .

Interpretations: Kerile [Kurile] Trench

(12) By relating the above observations to the theory of plate tectonics, an explanation for the lack of volcanic action on the middle of the California coastal area becomes inherent. (13) According to this theory, all three of the small areas which were observed lie amidst the Pacific Plate. (14) One of seven major lithospheric plates, this rigid slab of lithosphere or crust floats on a plastic-like layer of the mantle known as the asthenosphere. (15) The plates move relative to one another through a system known as convection where less dense material known as magma upwells from the mantle and denser pieces of the lithosphere downwell. (16) Areas such as the Kerile [Kurile] Trench along the Khamchatka Coast reveal the patterns of what geologists refer to as a continental convergent margin between two plates. (17) In this scenario, a plate containing oceanic crust collides with a plate made of continental crust which produces various effects (Figure 19). (18) The stereotypical description of a continental convergent margin possesses the same topographic effects, volcanism, and seismic activity observed along the Kerile [Kurile] Trench where the Pacific moves northwest into the North American Plates which moves southeast. (19) As the oceanic basin approaches the coast, the denser oceanic crust is forced, along with some oceanic sediments, under the less dense continental crust. (20) This process known as subduction, causes a trench which forms parallel to the margin of the two plates much like the Kerile [Kurile] Coast's topographical data suggests (Figure 4,5,6). (21) As the oceanic material subducts deeper into the Earth, it heats up do to the friction and hotter mantle material. (22) When the oceanic crust and sediments heat up, it transforms into magma and slowly rises through cracks of the crust to form volcanoes along the continental interior similar to those volcanoes observed inland from the Kerile [Kurile] Trench (Figure 3). (23) Figure 7 clearly depicts the upward and inland movement of heat flow in Khamchatka due to the upwelling of magma. (24) Behind the line of volcanoes, a back-arc basin forms due to rapid destruction of the subducting plate which sometimes surpasses the movement of the other plate. (25) Along the Kerile [Kurile] Coast this causes the nonsubducting plate, the North American Plate, to thin and stretch itself in order to compensate for the lack of reciprocal motion. (26) This process thus produces

a basin behind the volcanoes (Figure 6,7). (27) Continental convergent margins result in earthquakes because the subducting plate fractures under the stress and releases energy due to its folding below the nonsubducting plate. (28) Geologists typically describe the seismic activity of a convergence zones with earthquake foci below 100 m which become deeper the further inland on the nonsubducting plate they occur. (29) The data from the Kerile [Kurile] Coast suggests the earthquakes drop well below 100 m deep as the profile path moves further into the North American Plate (Figure 8).

Interpretations: Hawaiian islands and Emperor Seamounts . . .

Conclusions . . .

Figures

Figure 1. Mercator Projection map of locations of known volcanoes for three areas studied.

Figure 3: This map displays some topographic features of the Kerile [Kurile] coast or the western edge of the Pacific Plate. The area contains the Kerile [Kurile] Trench. The volcanoes' locations are displayed as well as the three profiles taken off the coast.

Figure 4: Profile off the Khamchatka coast which displays the trench and vertical rise of the continental interior.

Figure 5: Profile off of the Khamchatka coast which displays the trench and back-arc basin further south.

Figure 6: Profile off of the Khamchtka coast which displays the trench and back-arc basin even further to the south.

Figure 7: Heat flow map taken off of the Khamchtka coast along the profile path of Figure 5 which shows an increase in elevation of heat flow as the profile path moves away from the trench and into the continental interior.

Figure 8: Earthquake data plot taken along the profile path of Figure 5 which shows earthquakes increasing in depths below 100 km as the path moves inland.

Figure 19: A sketch of the continental convergent margin along the Kerile [Kurile] trench between the Pacific Plate and the North America Plate.

Figures 2, 9–18, 20–21. Refer to Hawaiian islands and Emperor Seamounts.

References

Case 4

Student pseudonym: Katie

Title: "Oceanography midterm: Determining the difference in plate boundaries"

Abstract . . .

Introduction . . .

Methods . . .

Observations: Convergent Zone

(1) The first particular area observed was found on the eastern coast of Asia (Figure 1). (2) This particular area was found to be located on the Pacific Plate. (3) In this area a trench was found (located at 41.25N longitude and 150.06E latitude) that extends a distance of 5190 km and reaches a depth of −8160 m (Figure 2). (4) Along with the trench I was able to observe the presence of many earthquakes, which reach a descending depth within 134 km of the surface (Figure 3). (5) The magnitude of these earthquakes range from a high

magnitude of 7.0 to a low magnitude of 5.0 (this also shown in Figure 3). (6) The presence of volcanoes was also apparent in this area. (7) Shown in Figure 4 is the presence of over 60 volcanoes along the coast of the trench, reaching a distance inland approximately 230 km. (8) I know that when a plate subducts below the continent it forces magma out of the earth's crust causing onshore volcanoes to develop (this is shown in Figure 5).

Interpretations: [Convergent Zone]

(9) With these findings I am able to presume that the plate boundary of this particular area is a convergent margin. (10) Some major factors in determining whether or not a plate is subducting goes back to the direction each plate is moving (Figure 6 shows the three plate boundaries and the directions each one moves in). (11) Because my findings prove to not have a ridge or large portions of volcanoes I can determine it's not a divergent zone. (12) The other case it could be is a transform margin, but due to the fact there proves to be no evidence of showing many shallow earthquakes then I know it is not a transform margin. (13) With the information I gathered from area I observed I am able to determine the Pacific plate is moving in a North-West direction, thus causing the Pacific plate to be pushed below the eastern coast of Asia and forcing the lithosphere to subduct below the continent shelf. (14) When the lithosphere subducts it then forces a trench that is scattered with many shallow and descending earthquakes; which are formed due to the bending of the plate (Ocean Sciences 97). (15) The volcanoes are formed due to the plate subducting below the continent forcing magma to rise to the surface from the earth's crust. (16) Over time these volcanoes will erupt and new ones will be formed.

Observation: Divergent zone [Mid-Atlantic Ridge] . . .

Interpretations [Divergent zone] . . .

Conclusions . . .

References

Figures

Figure 1: Location of convergent margin, eastern coast of Asia.
Figure 2: Trench depth, topology of the sea-floor observed in Figure 1.
Figure 3: Earthquake depth and magnitude.
Figure 4: Location of volcanoes.
Figure 5: Sketch of convergent zone, showing volcanoes and earthquakes.
Figure 6: Sketch of convergent subduction zone, divergent zone, and transform (fault).
Figures 7–10 Refer to divergent zone.

The data presented, the statements made, and the views expressed are solely the responsibility of the authors. We thank Catherine Chen for her help with the analysis and for her helpful comments on an earlier version of this manuscript. We thank Charles Bazerman for his methodological consultation and ongoing conversations regarding the rhetoric of science and its application to education. In addition, we thank Charles Bazerman, Julie Bianchini, Julie Esch, and William Prothero for their helpful comments on an earlier version of this manuscript.

REFERENCES

Atkinson, D. (1999). Scientific discourse in sociohistorical context: The philosophical transactions of the Royal Society of London 1675–1975. Mahwah, NJ: Erlbaum.

Ault, C. R. (1998). Criteria of excellence for geological inquiry: The necessity of ambiguity. Journal of Research in Science Teaching, 35, 189–212.

Bazerman, C. (1988). Shaping written knowledge: The genre and activity of the experimental article in science. Madison, WI: University of Wisconsin Press.

Bezzi, A. (1999). What is this thing called geoscience? Epistemological dimensions elicited with the repertory grid and their implications for scientific literacy. Science Education, 83, 675–700.

Britton, J., Burgess, T., Martin, N., McLeod, A., & Rosen, H. (1975). The development of writing abilities (11–18). London: Macmillan.

Carlsen, W. S. (1997). Never ask a question if you don't know the answer: The tension in teaching between modeling scientific argument and maintaining law and order. Journal of Classroom Interaction, 32, 14–23.

Duschl, R. A. (1990). Restructuring science education: The importance of theories and their development. New York: Teacher's College Press.

Duschl, R., Ellenbogen, K., & Erduran, S. (1999, April). Understanding dialogic argumentation among middle school science students. Paper presented at the annual meeting of the American Education Research Association.

Forman, E. A., Larreamendy-Joerns, J., Stein, M. K., & Brown, C. A. (1998). "You're going to want to find out which and prove it": Collective argumentation in a mathematics classroom. Learning and Instruction, 8, 527–548.

Goodwin, C. (1995). Seeing in depth. Social Studies of Science, 25, 237–274.

Halliday, M. A. K., & Hasan, R. (1976). Cohesion in English. Longman: London.

Halliday, M. A. K., & Martin, J. R. (1993). Writing science: Literacy and Discursive power. Pittsburgh: University of Pittsburgh Press.

Hoey, M. (1983).On the surface of discourse. London: George Allen & Unwin.

Kelly, G. J., & Bazerman, C. (2001, April). How Students Argue Scientific Claims: A Rhetorical-Semantic Analysis. Paper presented at the annual meeting of the American Education Research Association Seattle, WA.

Kelly, G. J., & Chen, C. (1999). The sound of music: Constructing science as sociocultural practices through oral and written discourse. Journal of Research in Science Teaching, 36, 883–915.

Kelly, G. J., Chen, C., & Prothero, W. (2000). The epistemological framing of a discipline: Writing science in university oceanography. Journal of Research in Science Teaching, 37, 691–718.

Kelly, G. J., Druker, S., & Chen, C. (1998). Students' reasoning about electricity: Combining performance assessments with argumentation analysis. International Journal of Science Education, 20(7), 849–871.

Kelly, G. J., & Green, J. (1998). The social nature of knowing: Toward a sociocultural perspective on conceptual change and knowledge construction. In B. Guzzetti, & C. Hynd (Eds.), Perspectives on conceptual change: Multiple ways to understand knowing and learning in a complex world (pp. 145–181). Mahwah, NJ: Erlbaum.

Keys, C. W. (1999). Revitalizing instruction in scientific genres: Connecting knowledge production with writing to learn in science. Science Education, 83, 115–130.

Keys, C. W., Hand, B., Prain, V., & Collins, S. (1999). Using the scientific writing heuristic as a tool for learning from laboratory investigations in secondary science. Journal of Research in Science Teaching, 36, 1065–1084.

Knorr-Cetina, K. (1995). Laboratory studies: The cultural approach to the study of science. In S. Jasanoff, G. E. Markle, J. C. Peterson, & T. Pinch (Eds.), Handbook of science and technology studies (pp. 140–166). Thousand Oaks: Sage.

Kuhn, D. (1992). Thinking as argument. Harvard Educational Review, 62(2), 155–178.

Latour, B. (1987). Science in action: How to follow scientists and engineers through society. Cambridge, MA: Harvard University Press.

Mukerji, C. (1989). A fragile power: Scientists and the state. Princeton, NJ: Princeton University Press.

Myers, G. (1997). Texts as knowledge claims: The social construction of two biology articles. In R. A. Harris (Ed.), Landmark essay on the rhetoric of science: Case studies (pp. 187–215). Mahwah, NJ: Erlbaum.

National Research Council (1996). National science education standards. Washington, DC: National Academy Press.

Newton, P., Driver, R., & Osborne, J. (1999). The place of argumentation in the pedagogy of school science. International Journal of Science Education, 21, 553–576.

Patronis, T., Potari, D., & Spiliotopoulou, V. (1999). Students' argumentation in decision-making on a socio-scientific issue: Implications for teaching. International Journal of Science Education, 21, 745–754.

Prain, V., & Hand, B. (1996). Writing for learning in secondary science: Rethinking practices. Teaching and Teacher Education, 12, 609–626.

Prain, V., & Hand, B. (1999). Students perceptions of writing for learning in secondary school science. Science Education, 83, 151–162.

Prothero, W. A. (1995). Taming the large oceanography class. Journal of Geological Education, 43, 497–506.

Ramage, J. D., & Bean, J. C. (1992). Writing arguments: A rhetoric with readings (2nd ed.). New York: MacMillan.

Richmond, G., & Striley, J. (1996). Making meaning in classrooms: Social processes in small-group discourse and scientific knowledge building. Journal of Research in Science Teaching, 33, 839–858.

Roth, W.-M., & McGinn, M. K. (1998). Inscriptions: Toward a theory of representing as social practice. Review of Educational Research, 68, 35–59.

Russell, T. (1983). Analyzing arguments in science classroom discourse: Can teachers' questions distort scientific authority? Journal of Research in Science Teaching, 20, 27–45.

Strike, K. (1982). Liberty and learning. Oxford, UK: Martin Robertson.

Takao, A. Y., & Kelly, G. J. (2001, March). Assessment of differences in university oceanography students' scientific writing. Paper presented at the annual meeting of the National Association for Research in Science Teaching, St. Louis, MI.

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

Toulmin, S. (1972). Human understanding. Vol. 1: The collective use and evolution of concepts. Princeton, NJ: Princeton University Press.

van Eemeren, F. H., Grootendoorst, R., & Kruiger, T. (1987). Handbook of argumentation theory: A critical survey of classical backgrounds and modern studies. Dordrecht, The Netherlands: Foris Publications.

Wolcott, W., & Legg, S. (1998). An overview of writing assessment: Theory, research, and practice. Urbana, IL: National Council of Teachers of English.