

## Article

# A Teaching Strategy with a Focus on Argumentation to Improve Undergraduate Students' Ability to Read Research Articles

Edwin B. Van Lacum,\* Miriam A. Ossevoort,<sup>†</sup> and Martin J. Goedhart<sup>†</sup>

\*College of Science, University of Amsterdam, 1098 XH Amsterdam, The Netherlands; <sup>†</sup>Department of Education, Faculty of Mathematics and Natural Sciences, University of Groningen, 9747 AG Groningen, The Netherlands

Submitted June 18, 2013; Revised March 14, 2014; Accepted March 15, 2014  
Monitoring Editor: Hannah Sevian

The aim of this study is to evaluate a teaching strategy designed to teach first-year undergraduate life sciences students at a research university how to learn to read authentic research articles. Our approach—based on the work done in the field of genre analysis and argumentation theory—means that we teach students to read research articles by teaching them which rhetorical moves occur in research articles and how they can identify these. Because research articles are persuasive by their very nature, we focused on the rhetorical moves that play an important role in authors' arguments. We designed a teaching strategy using cognitive apprenticeship as the pedagogical approach. It was implemented in a first-year compulsory course in the life sciences undergraduate program. Comparison of the results of a pretest with those of the posttest showed that students' ability to identify these moves had improved. Moreover, students themselves had also perceived that their ability to read and understand a research article had increased. The students' evaluations demonstrated that they appreciated the pedagogical approach used and experienced the assignments as useful. On the basis of our results, we concluded that students had taken a first step toward becoming expert readers.

## INTRODUCTION

The research article, the most common type of primary literature published in specialized scientific journals, is an important medium of communication within the scientific community. We define primary literature as reports of original observations, theories, or opinions, written for peers in the scientific community. Owing to electronic publishing, scientists today have instant access to an almost limitless amount of research articles (Björk *et al.*, 2009). Interestingly, this increased accessibility has changed scientists' reading habits.

Research has shown that scientists read more articles than they did before, although the total time spent on reading has only increased a little (Tenopir *et al.*, 2009). Therefore, the ability to read research articles in an efficient way is now more than ever an essential skill for scientists. This is why we argue that science students at research universities should be introduced to research articles at an early stage in their academic training, so they have enough time to develop this specific skill.

A study by Coil and colleagues (2010) highlights the relevance for students of reading research articles. They showed that members of the faculty think that it is important for undergraduate science students to acquire skills such as the ability to interpret data, write reports, and critically analyze research articles. However, these faculty members are also of the opinion that it is very time-consuming to teach these skills. Because they feel a pressing need to cover content in their introductory courses, they are unable to pay much attention to students' acquisition of these skills.

In this article, we will report on the evaluation of a teaching strategy for reading research articles. This strategy was integrated into a first-year compulsory course in the life sciences bachelor's degree program of a research university.

DOI: 10.1187/cbe.13-06-0110

Address correspondence to: Miriam A. Ossevoort (m.a.ossevoort@rug.nl).

© 2014 E. B. Van Lacum *et al.* CBE—Life Sciences Education © 2014 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution-Noncommercial-Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.

### Reading Research Articles

The ability to read research articles is a form of scientific literacy (Norris and Phillips, 2003). According to these authors, reading and writing are not merely interchangeable tools for the storage and transmission of information in science. Without reading and writing, the social practices that enable practicing science would not be possible. The use of texts allows scientists to record and present data, undergo peer review, and critically re-examine previously published ideas, and so forth. According to Norris and Phillips (2003), the reading and writing of scientific texts needs more emphasis in science education. Reading in science education, they argue, involves learning how to comprehend, interpret, analyze, and criticize texts—activities that are central to science. In this article, we will follow this broad definition by Norris and Phillips (2003). This means that reading will include not only the passive absorption of information but also active and complex interpretation processes such as analyzing and criticizing.

Nowadays science reading can be described as an interactive process in which the reader shifts between text-based information, concurrent experiences (such as discussions), and prior knowledge (Holliday *et al.*, 1994). Science reading is seen as an interactive constructive model, which is compatible with the constructivist model of science learning. This means that science reading is no longer viewed as an individual enterprise; instead, it should include opportunities for discussions among students, as well as between students and teachers.

Teaching undergraduates how to read research articles requires some special considerations for educators. Research articles are not easy to read, partly because of their language (Fang, 2005) and partly because students are used to textbooks and experience difficulty in coping with the persuasive aspects of research papers (Gillen, 2006). Therefore, it is important to develop specific teaching strategies that introduce students in higher education to research articles.

Several studies describe the reading behaviors of scientists and those of graduate and high school students. For instance, Bazerman (1988), as well as Berkenkotter and Huckin (1995), established the fact that experts read selectively (skipping parts of the text; reading only the parts they deem new and important) and nonsequentially (not reading the sections of an article in order). Charney (1993), like Bazerman (1988), demonstrated that scientists read articles “as is convenient for their own purposes (they read parts selectively and out of order); they weigh the plausibility of claims and evidence; they struggle to understand unfamiliar technical terms; they cheer and get mad” (p. 228). In contrast to scientists, graduate or high school students read sequentially and nonselectively and tend to do no more than understand the text and integrate it with their prior knowledge (Charney, 1993; Brill *et al.*, 2004).

Several studies are available on the use of primary literature in colleges and universities (e.g., Janick-Buckner, 1997; Levine, 2001; Kuldell, 2003; Peck, 2004; Robertson, 2012). In almost every study, students read research articles via individual guided reading (e.g., answering guiding questions about certain aspects of the article) followed by group discussions. These good practices, as well as study guides for reading research articles in a critical way (e.g., Yudkin, 2006), provide us with useful information about how to introduce

primary literature to students. However, these studies have a number of limitations. For instance, the effectiveness of these courses is often determined by evaluations, in which students assess the course and/or their own abilities, and not by measuring objective learning outcomes. Furthermore, these studies often lack descriptions of a theoretical framework, and it is not clear which teaching models were used to design the courses.

### Teaching Strategy for Reading Research Articles

The teaching strategy for reading research articles we developed consists of a heuristic based on the work done in the fields of genre analysis and argumentation theory. We used cognitive apprenticeship as the pedagogical approach.

In our case, the genre that we are working with is the research article. A genre consists of “a class of communicative events, the members of which share some set of communicative purposes. These purposes are recognized by the expert members of the parent discourse community and thereby constitute the rationale for the genre” (Swales, 1990, p. 58). Several genre analysis studies have determined the frequency of rhetorical moves in the different sections of a research article (Swales, 1990; Thompson, 1993; Dudley-Evans, 1994; Nwogu, 1997; Williams, 1999; Peacock, 2002; Kanoksilapatham, 2005). A rhetorical move refers to “a section of a text that performs a specific communicative function. Each move not only has its own purpose but also contributes to the overall communicative purposes of the genre” (Connor *et al.*, 2007, p. 23). The arrangement of rhetorical moves in a text is called the *rhetorical structure*. As an example, the moves present in the Discussion section of a research article are, for instance: 1) information move, 2) highlighting overall research outcome, 3) explaining specific research outcome, 4) referring to the previous literature, 5) claim, 6) limitation, and 7) recommendations (Swales, 1990; Thompson, 1993; Dudley-Evans, 1994; Nwogu, 1997; Williams, 1999; Peacock, 2002; Kanoksilapatham, 2005). Using a focus on rhetorical structure for teaching students how to read genre-specific texts has been proposed by several authors (Hill *et al.*, 1982; Blanton, 1990; Swales, 1990). As Swales (1990) wrote, “There may be pedagogical value in sensitizing students to rhetorical effects, and to the rhetorical structures that tend to recur in genre-specific texts” (p. 213).

As mentioned above, research articles are persuasive in nature (Suppe, 1998). Authors use data to convince readers that the conclusions presented are correct. In addition, they use references to other studies to consolidate their claims (Latour, 1987). The combination of data and references to support a conclusion in a research article can be called an argument. As stated by Du Boulay (2012), an argument refers to “an author’s claims (including their degree of strength), his or her theoretical orientation, the quality of the evidence produced or demonstrated and how this is linked to theory” (p.148). Over the past two decades, argumentation has attracted increasing attention in the field of science education in both pre-university and university education (Erduran and Jiménez-Aleixandre, 2008; Andrews, 2010; Kuhn, 2010). Several argumentation models, such as, for example, the Toulmin (1958) model, are used in an educational setting for teaching or analyzing students’ use of evidence in discourse or in

written products (e.g., Kelly and Takao, 2002; Sampson and Clark, 2008). Being able to read and understand argumentation is important for appreciating “the power and limitations of scientific knowledge claims” (Evagorou and Dillon, 2011, p. 191). As a result, reading research articles is closely related to understanding an author’s argument. Studies have demonstrated that students have difficulty in identifying an author’s argument in research articles (Kolokant *et al.*, 2006; Van Lacum *et al.*, 2012). This is why our heuristic focuses especially on those rhetorical moves that play an important role in the author’s argument. This way, students can become familiar with the persuasive aspects of research articles: the persuasiveness element, as explained above, is possibly one of the reasons why novice readers find it difficult to read research articles.

On the basis of the aforementioned studies, we developed the scientific argumentation model (SAM), a heuristic consisting of a set of seven moves (unpublished data) that play an important role in an author’s argument as given in research articles. We provided the students with a description of each move. In addition, we provided the students with clear, transferable criteria (organizational and lexical features) for identifying those rhetorical moves. Several examples of each move, taken from authentic research articles, were included. The moves are named and described as follows:

1. Motive: Statement indicating why the research was done (e.g., a gap in knowledge, contradictory results). The motive leads to the objective.
2. Objective: Statement about what the authors want to know. The objective may be formulated as a research question, a research aim, or a hypothesis that needs to be tested.
3. Main conclusion: Statement about the main outcome of the research. The main conclusion is closely connected to the objective. It answers the research question, it says whether the research aim was achieved, or it states whether the hypothesis was supported by evidence. The main conclusion will lead to an implication.
4. Implication: Statements indicating the consequences of the research. This can be a recommendation, a statement about the applicability of the results (in the scientific community or society), or a suggestion for future research.
5. Support: The statements the authors use to justify their main conclusion. These statements can be based on their own data (or their interpretation) or can be statements from the literature (references). Supports may be presented in so-called support chains. For example: table → interpretation of the table’s data in the Results section (statement of finding) → further interpretation of the table’s data in the Discussion section (preliminary conclusion).
6. Counterargument: Statements that weaken or discredit the main conclusion. For example, possible methodological flaws, anomalous data, results that contradict previous studies, or alternative explanations. Counterarguments are sometimes presented as limitations.
7. Refutation: Statements that weaken or refute a counterargument.

As stated above, in our teaching strategy we teach students how to read research articles by teaching them where those rhetorical moves occur in the different sections of research articles and how they can identify these moves. We used cognitive apprenticeship as the pedagogical approach. According to Collins and colleagues (1991), cognitive apprenticeship involves three characteristics: 1) the processes of the task should be identified and made visible to students; 2) abstract tasks should be situated in authentic contexts (so students will understand the relevance of the task); and 3) the diversity of situations should be varied and common aspects should be articulated (so students may transfer what they learn). The first characteristic is put into practice by letting students acquire an integrated set of skills by observing an expert who performs a task (thinking is made visible via *modeling*) and guides the newcomers when they practice this task (*coaching*). Students are given support that helps them to carry out the task (*scaffolding*). Eventually, support is gradually removed until students are able to accomplish a task on their own (*fading*). During the group sessions, students articulate their knowledge, reasoning, or problem-solving processes (*articulation*), and compare their own processes with those of other students, an expert, or—ultimately—an internal cognitive model of expertise (*reflection*).

To create an authentic context (second characteristic), we used research articles that had not been edited, translated, or adapted. Confronting students with the complexity of investigations described in research articles, enables them to learn not only science content but also about the scientific method (e.g., Epstein, 1970; Hoskins *et al.*, 2011). In addition, by learning to read and understand scientific language, students will slowly become part of the “community of practice” of science (Lave and Wenger, 1991). Ultimately, mastering the language of science will enable students to communicate and function in this community and to identify themselves as scientists.

To stimulate articulation and reflection (third characteristic), we encouraged students to discuss with peers and more experienced readers what they have read. This is in accordance with the aforementioned studies about reading courses in higher education, which suggest that discussing primary literature may be a useful method for increasing the understanding of the text. To achieve this, we used cross-year, small-group tutoring (tutoring by students from other years). It has to be noted that peer tutoring has some disadvantages. For example, the student tutor’s mastery of content is probably less than that of a professional instructor. However, Topping (1996) reviewed a number of studies on the effectiveness of cross-year, small-group tutoring and concluded that “much of the research is not of the highest quality, but good quality studies . . . do clearly demonstrate improved academic achievement” (p. 327). To create the diversity of situations (third characteristic), we let students read a number of different research articles to put the transfer of their reading skills into practice.

## AIM

The aim of this study is to evaluate a teaching strategy for life sciences students at a research university, in which they practiced reading authentic research articles by focusing on the rhetorical moves that play an important role in the authors’

argumentation. The teaching strategy, implemented in a first-year undergraduate course, follows cognitive apprenticeship as the pedagogical approach. For this study, our research questions were:

1. What is the actual progress made by undergraduate life sciences students in terms of their ability to identify rhetorical moves in research articles after following our teaching strategy?
2. What are the students' own perceptions as to their ability to read a research article, both before and after our teaching strategy?
3. How does students' reading behavior change during the teaching strategy?
4. How do the students evaluate our teaching strategy?

## DESCRIPTION OF THE TEACHING STRATEGY AND ASSIGNMENTS

### Context and Course Organization

Courses at the undergraduate level of the life sciences curricula at the University of Groningen, The Netherlands, integrate the teaching of knowledge and skills. In this study, we will evaluate our teaching strategy as implemented in a course called Biomedical Research. This compulsory course was part of the last quarter of the first-year undergraduate program. The educational aims of the course were that students would 1) understand the physiology and pharmacology of the cardiovascular system; 2) know the possibilities and limitations of *in vitro* animal research and develop research skills during lab assignments; and 3) be able to read scientific texts and communicate both orally and in writing.

The course (which lasted 11 wk) consisted of lectures presenting four main topics (related to the first educational aim), lab work (related to the second educational aim), and tutorial group meetings (related to the first and third educational aims). The main topics were the autonomous nervous system (week 1), the heart (weeks 2–4), the cardiovascular system (weeks 5–8), and healthy aging (weeks 9–10). The practical work concerned the function of the heart and blood vessels. A schematic representation of the outline of the Biomedical Research course is shown in Figure 1.

At the end of the course, students took a multiple-choice exam (knowledge test about the aforementioned topics) and

an oral examination, during which they individually presented their summary of a research article (Figure 1). The assignments and lectures were in Dutch. The textbooks and research articles were in English.

Our teaching strategy for reading research articles was implemented in eight weekly tutorial group meetings (Figure 1). During the preceding lectures, students were made familiar with the concepts discussed in the research articles.

The total number of first-year undergraduate life sciences students who took the course was 125, randomly divided over 14 tutorial groups. The students were ~18–20 yr old and their native language was Dutch. The tutors were between 20 and 23 yr old, were studying life sciences ( $n = 8$ ) or medicine ( $n = 6$ ), and were third-year bachelor's or master's students. The tutors had ample experience with reading research articles. The tutors had applied for positions as teaching assistants.

### Teaching Strategy

**Tutorial Group Meetings.** Each week, at the end of the tutorial group meeting, students received a new research article, an assignment, and instructions from the tutor. After the first meeting, students received information sheets, in which we had listed the seven different moves, together with definitions of all moves and examples taken from authentic research articles (i.e., scaffolding). In a 2-h meeting with all the tutors, we explained our teaching design and pedagogical approach, including how to conduct the meeting and how to give feedback. We also emphasized that they should demonstrate how they read a research article and identify the moves (i.e., modeling). The assignment for the students consisted of reading the research article and answering questions on paper. Students did the assignment as homework. During the tutorial group meeting that followed, students discussed the article and their answers to the assignment. For the discussion, we provided tutors with discussion prompts concerning the methodology, the meaning of certain technical terms, the interpretation of results, and the connections with other articles. In addition to the feedback given during the tutorial group meeting, students received individual written feedback on their answers from the tutor as soon as possible, so the students could implement the suggestions in time for the next homework assignment. The tutorial meetings were held in small meeting rooms at the university. Meetings usually lasted 2–2.5 h and were mandatory.

**Homework Assignments.** During the teaching sessions, students received six research articles and six homework assignments. The central part of these assignments was the identification of the seven moves of our heuristic: motive, objective, main conclusion, implication, supports, counterarguments, and refutations. We did not want to overload students at the beginning of the teaching sessions by having them identify all the moves at once. That is why we followed a cumulative approach. Once we did introduce a move, the students had to identify this move in all subsequent assignments (Figure 1). We hoped that by repeating the identification process, students would rely less on the information sheets (our scaffolding method) as the course progressed. In this way, fading could occur. For tutorial group meeting 2 (T2), students focused on the general structure of research articles. For T3,

Lectures											
Practicals											
PRE-TEST	Reading research articles, identification of:										POST-TEST
	-	-	O,M	O,M,MC,I	all moves			practice for oral exam		FINAL EXAM	
	T1	T2	T3	T4	T5 /	T6 /	T7	T8			
Wk	1	2	3	5	6	7 /	8 /	9	10	11	

**Figure 1.** A schematic representation of the outline of the course Biomedical Research, including our teaching strategy. The final exam consisted of a knowledge test and an oral examination. Abbreviations of the rhetorical moves: O = Objective; M = Motive; MC = Main Conclusion; I = Implication.



students identified the motive and objective. For T4, students identified the motive, objective, main conclusion, and implication. For T5, T6, and T7, they identified all seven moves (Figure 1). During T6 and T7, students received instruction about the linkage between all the moves, which would serve to demonstrate the argumentation structure of a research article.

**Additional Assignments.** For the homework assignments, students not only identified rhetorical moves but each week also had to answer additional questions that were partly based on assignments published by Yudkin (2006) and Gillen (2007). For T2, the students answered questions about the general structure of the research articles. We asked them to formulate the function of the different sections (Abstract, Introduction, Method, Results, and Discussion) and to summarize these sections. The students formulated criteria for a good title and determined whether the article of the week's title followed these criteria. Furthermore, they answered questions about the article's references, the time between acceptance and publication, and the funding for the research. For T3, the students chose five important concepts mentioned in the Introduction and explained their meaning. We asked students to describe the field of research and why the authors referred to previous research in their Introduction. They then summarized this previous research. For T4, the students were asked whether the main conclusion and implication were related to each other, and how certain the authors were about their main conclusion. For T5, the students identified the most essential figure or table, and justified their choice. Students evaluated the quality of the refutations. Finally, the students devised a new counterargument (not mentioned by the authors) and a refutation. For T6, the students described the article's experimental and control groups and summarized four experiments. For T7, no additional questions were asked. During the tutorial group meetings in weeks 7, 8, and 9, the students practiced summarizing a research article for the oral examination.

## RESEARCH DESIGN

We measured the effectiveness of our teaching strategy in a single field experiment in which the variables were isolated and controlled using a so-called pre-experimental design with one group for pretest and posttest (Cohen *et al.*, 2008). In educational research it is often—as in this case—not possible to conduct true experiments using control groups. The final assessment of the course in which we implemented our teaching strategy partly consisted of students' ability to summarize a research article (oral examination). If we had used a control group, one in which students were not expected to read a research article, we would quite obviously be disadvantaging them. Furthermore, the lectures did not allow us to include a control group. Therefore, we studied the students as a single group, measuring their ability to identify the moves in a research article using a pretest and posttest design. We assigned the differences between these scores as being an effect of our teaching strategy. It is in any case unlikely that other extraneous variables influenced this outcome, since during the experiment students were taking only this course and no others.

## Pretest and Posttest

To measure the effectiveness of our teaching strategy, we had students take a pretest and posttest as homework. The pretest article and assignment were given at an introductory meeting a week before the teaching sessions started. The posttest article and assignment were given at the end of T7. Students handed in their answers to the pretest and posttest assignments by email before the first tutorial group meeting and before T8, respectively. Both assignments consisted of the following questions:

1. What was/were the researcher's/researchers' motive(s) for conducting this research?
2. What was/were the researcher's/researchers' research question(s) or objective(s)?
3. What is/are the conclusion(s) drawn by the researcher(s) from the results?
4. Give the author's/authors' support(s) for this/these conclusion(s).
5. What is/are the main conclusion(s) drawn by the researcher(s) from the results?
6. What are, according to the researcher(s), the implications of the research?
7. Which factors does/do the author(s) mention that weaken the results or conclusion(s)?

When we formulated the questions, we deliberately did not use the terms *counterargument* and *refutation* but, in general, asked which factors the authors mentioned that weakened their conclusions. At the time of the pretest, students had not received instructions about the meaning of these terms, and we wanted to make sure that students would not become confused.

We implemented the pretest and posttest via a parallel test. The tutorial groups were divided into two groups, A ( $n = 72$ ) and B ( $n = 53$ ). For logistical reasons, the group sizes differed. Because some students did not hand in answers to all their assignments, we only analyzed data for 108 students (group A: 66 students; group B: 42 students). Female ( $n = 60$ ) and male ( $n = 48$ ) students were evenly distributed between both groups. We determined that there was no significant difference between the two groups regarding their academic performance (range of total score = 0–90) in the courses of the first semester preceding this course (group A [ $n = 66$ ]: mean total score: 63.6 [SD = 9.3]; group B [ $n = 42$ ]: mean total score: 61.8 [SD = 12.2] [independent Student's  $t$  test:  $p = 0.414$ ,  $t(106) = 0.819$ ]). At the beginning of the teaching strategy, the students (minus 12) filled out the questionnaire about their reading experiences regarding primary literature. Two of the students had not read any of the research articles at all. Seventy-five students had read one to six articles, 16 students had read seven to 12 articles, and three students had read more than 12 articles. Thus, the students were all novice readers of research articles. The distribution between both groups was similar (unpublished data).

During the pretest, group A received article 1 by Bas *et al.* (2007), and group B received article 2 by Ozen *et al.* (2008). During the posttest, group A received article 2, and group B article 1. By switching the articles, we were able to eliminate any possibility that the improvement measured was due to a

posttest article that was easier to read and understand than the pretest article.

Articles 1 and 2 described the effects of a fish oil (n-3 essential fatty acids [EFA]) diet on cerebral injury in rats and were similar in style and content. Both studies used a control group (with rats on a normal diet) and an experimental group (with rats on a diet enriched with fish oil). In the rats of both groups cerebral injury was produced. Then, the apoptotic neurons (data A) were counted and the levels of several biomarkers (data B–E; B = malonaldehyde [MDA]; C = superoxide dismutase [SOD]; D = nitric oxide [NO]; E = catalase [CAT]) were measured to determine the amount of damage to the rats' brains. In contrast to the articles used in the teaching sessions, the concepts mentioned in articles 1 and 2 were closely related to the topics discussed in the lectures; they were not, however, explicitly discussed by the course lecturers. The main body of article 1 contained ~4000 words, one figure, and two tables. Article 2 contained ~3300 words, two figures, and one table. Readability of the articles was measured using the Flesch Reading Ease Score (Flesch, 1948). The Flesch Reading Ease Score for articles 1 and 2 were 50 and 46, respectively. This means that the articles were "fairly difficult" and "difficult" to read. Although the articles fall into different scoring categories, we will assume that the articles are more or less equivalent in readability, as the differences between the two scores is very small.

### ***Students' Perceived Ability to Identify the Moves of SAM and Their Reading Behavior***

After taking the pretest and posttest, the students completed an online questionnaire, in which they indicated on a 5-point rating scale (from strongly disagree [1] to strongly agree [5]) how much they agreed with seven different statements (see Table 3 later in the *Findings* section). The statements dealt with students' ability to read a research article, understand the experimental procedures, and identify certain rhetorical moves in research articles. In addition, we asked the students whether they had used a dictionary for translation purposes or a textbook for looking up certain technical terms or concepts when reading the article and completing the assignment. Furthermore, to determine their reading behavior, the students had to indicate on a 4-point rating scale how well they read the different parts (Abstract, Introduction, Method, Results, Discussion section, and the figures and tables) of the research article: *not*, *casually*, *good*, or *very good*. Their answers were used to assess to what extent they had read selectively. The students also had to indicate whether they had read the article sequentially or nonsequentially.

### ***Course Evaluation***

The course was evaluated by the students by means of a standard course evaluation form used for all courses and with an additional evaluation form that focused on our teaching strategy. For this study, we used the students' answers on six items (5-point Likert scale) and two open-ended questions. The six items are listed in Table 5 in *Findings*. The two open-ended questions were: 1) Which research articles should not be used in the teaching strategy, and why? 2) Which parts of the course should absolutely be retained?

### ***Data Analysis***

***Pretest and Posttest.*** For the analysis of the pretest and posttest, we devised a scoring model based on our own analysis of the rhetorical moves in the two articles. For each move, we determined a number of elements that should be present in a student's answer. For example, the objective in article 1 was stated as: "The aim of this study was [to investigate] the [antioxidant] and [neuroprotective effects] of [fish n-3 EFA] on [cerebral ischemia(I)/reperfusion(R) injury] Sprague Dawley [rats'] [hippocampal formation]." For each element placed between brackets, the student was awarded one point. Thus, the students could earn 7 points for the objective in article 1. The maximum numbers of points that could be scored for the motive, objective, main conclusion, implication, and counterargument of article 1 were, respectively, 6, 7, 7, 5, and 4.

For article 2, the maximum numbers of points that could be scored for the motive, objective, main conclusion, implication, and counterargument were, respectively, 4, 7, 7, 5, and 4. We then calculated the scores for each answer. For the analysis of the answers to questions 1, 2, 5, 6, and 7 of the pretest and posttest assignments, the first author blindly rated the students' answers (i.e., he did not know which answers came from the pretest and which answers came from the posttest). To check the reliability of our method, the second author also blindly rated 120 randomly chosen items. Krippendorff's alpha (Hayes and Krippendorff, 2007) was 0.98, which indicates that there was a high interrater agreement. We used the first author's ratings to calculate the score per group per move.

To calculate the students pretest to posttest improvement in identifying the different moves, we combined all the pretest and posttest data of all the students (groups A and B) and ran a paired-test statistical analysis (Wilcoxon signed-rank test: alpha = 0.05). For the motive, we first calculated the percentage of the total score for each article, because the maximum number of points that could be scored between the motives in articles 1 and 2 differed.

Looking at the data for the whole student population, we observed a trend whereby three clusters of scores became apparent. Therefore, for a clearer representation of the data we used the classifications of "incorrect," "semicorrect," and "correct." A student's score was classified as correct if the maximum number of points was scored. It was classified as incorrect when less than half of the maximum number of points was scored. We are of the opinion that an answer with a score of less than half of the maximum number of points clearly does not reflect a text fragment representing the particular move. Answers that did not fall into the aforementioned categories, those scoring between half and the full number of points, were classified as semicorrect. For determining the significance of the differences between the pretest and posttest scores of each group, we used the chi-square test (alpha = 0.05).

For the analysis of the supports, we used the students' answers on questions 3 and 4 of the pretest and posttest assignments. First, we selected those students who pointed out the correct text fragments as being a conclusion in question 3. Then, using the answer to question 4, we scored how many correct supports were mentioned by each of these students. The maximum number of supports for article 1 and article 2 are 5 and 4, respectively. The score 0 represents those students who did not note any of the above-mentioned supports

or had written down: “all data support the conclusion.” We also identified which supports (data A–E, see above for a description) the students indicated.

We are aware that, when determining students’ improvement in identifying the moves, different articles were used in generating the pretest and posttest data. Careful selection of the articles meant that the influence of this factor was minimized. Both articles are more or less equivalent in readability and content knowledge (see above). Moreover, by switching the articles, we were able to eliminate the possibility that the improvement measured was due to a posttest article that was easier to read and understand than the pretest article was.

**Students’ Perceived Ability to Identify the Moves of SAM and Their Reading Behavior.** Regarding the statements from the questionnaire that dealt with students’ perceived ability to read a research article, understand the experimental procedures, and identify certain rhetorical moves in research articles, significant differences between the beginning and end of the teaching sessions were determined via a Wilcoxon signed-ranked test. This test was also used to analyze the students’ answers to our question about how well they read the different sections of the research articles. The percentage of the whole student population reading nonsequentially or sequentially during the pretest and posttest was calculated.

**Course Evaluation.** The mean scores of agreement for each item were calculated for the whole student population. The answers to the two open-ended questions were categorized, and the percentages of students who mentioned a certain category were calculated.

## FINDINGS

### Students’ Ability to Identify the Moves of SAM

We determined students’ ability to identify the five moves of SAM (motive, objective, main conclusion, implication, and counterargument) by attributing points to their answers, based on our scoring model. First, we analyzed the improved

identification of the five moves of SAM for all the students (groups A and B) by comparing the pretest and posttest data using a paired-test statistical analysis (Wilcoxon signed-rank test:  $\alpha = 0.05$ ). For all the moves, except the identification of the counterargument ( $p = 0.587$ ), the students’ ability to identify the moves significantly improved ( $p < 0.001$ ). We did not determine the identification of the refutation, because we asked the students to list the authors’ statements that weakened the results and/or conclusion. We deliberately did not use the terms *counterargument* and *refutation* because the students at the time of the pretest had not received any instruction as to the meaning of these terms.

To determine whether our premise—that the two articles were more or less equivalent in content knowledge and readability—was correct, we separately analyzed the pretest and posttest scores for the students of groups A and B (Table 1). With the exception of the identification of the objective by group A, the implication by group B, and the counterargument by groups A and B, the students of groups A and B showed significant improvement in their identification of the moves of SAM in the posttest as compared with the pretest (Table 1). Overall, in the posttest, the percentage of students who correctly identified the motive, objective, and the implication was higher than the percentage of students who correctly identified the main conclusion and the counterargument (Table 1).

We distinguished five and four essential supports, respectively, for the main conclusion of articles 1 and 2. We found no difference in the amount of supports the students indicated in the pretest versus the posttest for articles 1 and 2 (Table 2). In article 1, three supports were identified by more students in the posttest than in the pretest (data B: 14 [pre] to 23% [post]; data C: 24 [pre] to 32% [post]; data E: 17 [pre] to 32% [post]). The two other supports (data A and D) were identified by slightly fewer students. Regarding article 2, we saw an increase for two of the supports (data B: 33 [pre] to 41% [post]; data C: 22 [pre] to 41% [post]), and a small decrease for two other supports (data A and E). Data D was not determined in article 2. It has to be noted that only nine students indicated the main conclusion as being one of the conclusions for article 2 during the pretest, so the sample size is very small.

**Table 1.** Students’ ability to identify moves of SAM in the pretest and posttest (in percentage)

	Group <sup>a</sup>	Pretest			Posttest			Chi-square test <sup>b</sup>
		Incorrect	Semicorrect	Correct	Incorrect	Semicorrect	Correct	
Motive	A	44	21	35	42	3	55	$p = 0.003$
	B	71	0	29	17	7	76	$p < 0.001$
Objective	A	6	32	62	1	26	73	$p = 0.250$
	B	5	59	36	2	19	79	$p < 0.001$
Main conclusion	A	79	15	6	63	14	23	$p = 0.024$
	B	83	7	10	60	7	33	$p = 0.027$
Implication	A	73	9	18	20	13	67	$p < 0.001$
	B	53	21	26	43	12	45	$p = 0.159$
Counterargument	A	59	8	33	82	0	18	$p = -0.006$
	B	88	0	12	69	2	29	$p = 0.088$

<sup>a</sup>Group A ( $n = 66$ ) read article 1 in the pretest and article 2 in the posttest. Group B ( $n = 42$ ) read article 2 in the pretest and article 1 in the posttest.

<sup>b</sup>The chi-square test ( $\alpha = 0.05$ ) was used for statistics.

**Table 2.** Percentage of students mentioning a certain number of supports for the correct main conclusions of article 1 (pretest,  $n = 29$ ; posttest,  $n = 31$ ) and of article 2 (pretest,  $n = 9$ ; posttest,  $n = 29$ )

Number of supports <sup>a</sup>	Article 1		Article 2	
	Pretest	Posttest	Pretest	Posttest
0	48	52	45	45
1	17	13	0	10
2	7	10	33	4
3	14	0	0	10
4	7	3	22	31
5	7	22	-	-

<sup>a</sup>Article 1 contains five supports. Article 2 contains four supports.

### Students' Perceived Ability to Identify the Moves of SAM and Their Reading Behavior

After the aforementioned pretest and posttest, the students assessed their abilities regarding their reading and understanding of the research articles. Comparing the answers to the pretest questionnaire and posttest questionnaire, for five of the statements we saw a significant increase in the degree to which the students agreed with them (Table 3). These statements involved the ability to read a research article in a structured way and the identification of certain rhetorical moves. The students' perceived ability to understand the materials and methods, in addition to the experimental design used in the article, had not significantly increased.

These results are corroborated by the decreased use of textbooks and dictionaries when the students read the articles and did the assignments. Our data show that 76% of the students used a textbook during the pretest versus 40% during the posttest. Furthermore, 61% of the students used a dictionary during the pretest versus 39% during the posttest.

The students' reading behavior also changed after they had followed our teaching strategy. After completion of the pretest and posttest, 87% of the students ( $n = 96$ ) said they had read the article during the pretest in a sequential way versus 68% ( $n = 103$ ) during the posttest.

**Table 4.** Frequencies of students' answers (in percentages) when asked how well they read the different sections of the pretest ( $n = 96$ ) and posttest ( $n = 103$ ) articles<sup>a</sup>

	Pretest		Posttest	
	Not/casually	Good/very good	Not/casually	Good/very good
Abstract	23	77	43	57
Introduction	27	73	20	80
Method	36	64	74	26
Results	10	90	18	82
Discussion	12	88	8	92
Figures and tables	40	60	35	65

<sup>a</sup>For this table we grouped the students who answered "not" or "casually" and the students who answered "good" or "very good."

We also asked the students to indicate on a 4-point rating scale how well they had read the different sections of the research articles during the pretest and posttest (Table 4). When we compared the posttest with the pretest, the students said they had paid significantly less attention to the Abstract ( $z = -2.9, p = 0.003$ ; Wilcoxon signed-rank test, two-tailed), Methods ( $z = -5.3, p < 0.001$ ), and Results ( $z = -3.7, p < 0.001$ ) sections (Table 4). The attention paid to the Introduction section, Discussion section, and the figures and tables remained the same.

### Evaluation of the Teaching Strategy by the Students

Students evaluated the teaching strategy with a questionnaire filled in at the end of the course. The results show that the students generally evaluated the different parts of the teaching strategy positively (Table 5). These results are in accordance with the fact that 78% of the students answered "practicing reading research articles during the tutorial group meetings" to the question as to what parts absolutely should stay in the Biomedical Research course. In their comments, the students stated, for instance [translated into English]: "The tutorial

**Table 3.** Students' perceived ability (in percentage) to read a research article, understand the experimental procedure, and identify a certain rhetorical move (pretest,  $n = 95$ ; posttest,  $n = 102$ ; 1 = strongly disagree, 2 = disagree, 3 = disagree/agree, 4 = agree, 5 = strongly agree)

I am able to ...		1	2	3	4	5	Mean (SD)	Wilcoxon signed-rank test
read a research article in a structured way.	Pre	0	18	24	54	4	3.4 (0.8)	$z = -4.3$ $p < 0.001$
	Post	0	4	10	78	8	3.9 (0.6)	
identify the research question.	Pre	0	5	17	71	7	3.8 (0.6)	$z = -5.9$ $p < 0.001$
	Post	1	3	2	52	42	4.3 (0.7)	
understand the choice of materials and methods used.	Pre	1	20	41	37	1	3.2 (0.8)	$z = -1.5$ $p = 0.135$
	Post	1	14	38	44	3	3.3 (0.8)	
understand the experimental design.	Pre	0	4	14	79	3	3.8 (0.6)	$z = -1.2$ $p = 0.229$
	Post	0	7	21	66	6	3.7 (0.7)	
identify the results.	Pre	0	5	8	77	10	3.9 (0.6)	$z = -3.2$ $p = 0.002$
	Post	0	4	3	67	26	4.2 (0.7)	
identify the conclusion.	Pre	0	4	17	71	8	3.8 (0.6)	$z = -3.2$ $p = 0.001$
	Post	0	4	7	66	23	4.1 (0.7)	
identify the supports used to justify the conclusion.	Pre	0	10	36	51	3	3.5 (0.7)	$z = -2.7$ $p = 0.008$
	Post	0	7	20	67	6	3.7 (0.7)	



**Table 5.** Student evaluations ( $n = 104$ ) of the teaching strategy for reading research articles<sup>a</sup>

Item	Mean score (SD)
Assistance of the tutor	4.3 (0.8)
Content of the lectures and research article parallels	3.8 (0.7)
Quality of the homework assignments	3.4 (0.7)
Order of the homework assignments	3.7 (0.6)
Quality of the additional assignments	3.3 (0.8)
Preparation for the oral examination during the tutorial group meetings	4.2 (0.8)

<sup>a</sup>Mean scores on a Likert 5-point scale; range: 1 = very bad, 2 = bad, 3 = neutral, 4 = good, 5 = very good.

group meetings were instructive and enjoyable”; “I liked getting feedback on my assignments”; “This is very useful for me!”; and “Now I’m able to read a research article much faster.” About half of the students ( $n = 39$ ) stated that all the research articles were suitable.

## DISCUSSION

The main purpose of our teaching strategy, which was evaluated in this article, was to increase students’ ability to identify components of the authors’ argument in an authentic empirical research article. Our teaching strategy was based on the SAM—a heuristic consisting of a set of seven moves—and we followed cognitive apprenticeship as the pedagogical approach. In comparing the pretest and posttest, we demonstrated that the students showed improvements in their abilities to identify the motive, objective, main conclusion, and implication. Regarding the identification of supports and counterarguments, there was no significant difference between the pretest and posttest. The students’ self-assessment showed an increase in their perceived ability to identify all the rhetorical moves involved. The students’ perceived ability to understand the materials and methods, as well as the experimental design used, did not significantly increase. In our teaching strategy, we did not pay much attention to these aspects, so this might explain this finding.

Furthermore, at the end of the teaching strategy, the students reported a change in their reading behavior: fewer students read in a sequential manner, and more students read selectively. Their evaluation of the teaching strategy demonstrated that the students appreciated the pedagogical approach used and experienced the assignments as very useful. Given our results, it seems likely that our students have become better readers. We therefore conclude that the focus on the research articles’ rhetorical structure may be a powerful tool for introducing undergraduate students to primary literature.

The identification of the objective by the students in group A did not significantly improve, but this may be due to their high score during the pretest. The objective in article 1 (pretest) starts with “The aim of this study was to investigate . . .,” whereas the objective in article 2 (posttest) starts with “In this study, we investigated. . .” The objectives in both articles contained the rhetorical cue “investigate.” However, the

use of the word “aim” might have been a trigger for novice readers during the pretest (before instruction had begun) to recognize the objective.

Our study also showed that a majority of the students still had difficulty in identifying the main conclusion. This was quite puzzling, because in a previous study we found that almost all of the students had found the most important conclusion in a research article (Van Lacum *et al.*, 2012). In the previous study, we suggested that the students might have relied on lexical features like reporting verbs (e.g., suggests, found, show) and transition words/phrases (e.g., overall, so, in summary) to identify the conclusions. The main conclusions in articles 1 and 2 do contain these kinds of lexical features. There are sentences before and after the main conclusions that also contain the lexical features of conclusions, but by looking carefully at the content of these sentences it should be possible to tell that they are not the main conclusion. This could explain why a majority of the students were unable to identify the correct sentences as the conclusions. These findings stress the importance of paying attention to—in addition to the rhetorical and lexical features—the content feature of the moves in our instructional approach.

We observed that the students had difficulty finding counterarguments. This accords with Kuhn (1991), who found that students had difficulty recognizing the critical status of the counterclaims they encountered. One possible explanation is that counterarguments are scattered throughout the Discussion section (unlike the main conclusion, which is in most cases placed at the beginning or end of the Discussion section) and often do not contain distinctive lexical features. Another possible explanation is that the students are used to reading textbooks. Textbooks tend to present knowledge claims without explaining how these claims came to be (Goldman and Bisanz, 2002), so these texts seldom contain counterarguments. As a result, students tend to have limited experience with the nature of counterarguments.

We observed that the total number of supports the students indicated had not increased at the end of the teaching sessions. We also observed that three of the five (in the case of article 1) and two of the four (for article 2) essential supports were identified by more students at the end of the teaching sessions. However, no increase was observed for the other supports. In addition, the support that was most frequently identified was identified by only half of the students. Therefore, it seems that, despite our teaching strategy, the students still had problems finding all the evidence used to justify a conclusion. We can only speculate about possible reasons for this result. One reason may be that supports, like counterarguments, often lack distinctive features. Another reason might be that students do not realize which supports are important as evidence for the main conclusion and do not realize that more supports are usually needed as evidence in order to justify a claim.

As stated in the *Introduction*, other studies have, in accordance with our results, shown an increase in students’ confidence in reading and analyzing research articles via a Likert-style survey after taking a primary literature reading course (e.g., Kozeracki *et al.*, 2006; Hoskins *et al.*, 2011). In our study, the students stated via the self-assessment that they were able to identify the main conclusion, although only 23–33% of the students were actually indeed able to identify the main

conclusion during the posttest. This means that our study suggests that self-assessments are limited in what they can tell about students' improvement.

Our study indicates that the students adopted different reading behaviors during the teaching sessions. There was an increase in the number of students who stated that they had read their article nonsequentially. Compared with the pretest, the students said they read the Abstract, Method, and Results sections less intensively, which indicates that they had become more selective in their reading. These two findings suggest that students' reading became more goal directed. Maybe the students seemed to pay less attention to these sections of the articles because they thought they could not find answers to the assignment questions in these sections. For the Methods section this seems likely, since we asked no questions about the Methods section. But for the Results section it is puzzling, since the data (evidence to support the claim) are presented in this section. Maybe the students thought that they did not have to read this section, since the evidence is also mentioned in the Discussion section. In future studies, to elucidate the students' rationale as to whether they are reading more nonsequentially and selectively, task-based interviews observing students' reading behavior will be carried out. In addition, in our teaching design, we want to stress the importance of reading the Results section more.

It could be argued that the students performed better on the posttest because their knowledge of the concepts discussed in the articles had increased. This is unlikely, however, since the subject of the pretest and posttest articles (the effects of a fish oil diet on cerebral injury) was only loosely related to the subject of the course (the cardiovascular system). Furthermore, the effects of priming were probably minimal, because there were 6 wk between the pretest and posttest.

Our students were nonnative speakers of English. We have no reason to suspect that this had a noticeable influence on our results, because Dutch students are generally well versed in the English language. All the students' textbooks are in English, so they have ample experience in reading English science texts. Furthermore, research suggests that students' language skills play a much less important role than conceptual knowledge with respect to the comprehension of scientific texts (Chen and Donin, 1997). We did find that ~40% of the students still used a dictionary during the posttest. In future studies using task-based interviews, we would like to investigate which English terms or rhetorical cues the students did not understand, so we can pay more attention to these during the tutorial group meetings or lectures.

Norris and Phillips (2003) emphasize that reading is not "a simple concatenation of word meanings, is not characterized by a linear progression or accumulation of meaning as the text is traversed from beginning to end, and is not just the mere location of information" (p. 229). Instead, they characterize reading as a process that (among other things) "requires the active construction of new meanings, contextualization, and the inferring of authorial intentions" (p. 229). That is why teaching students how to identify the rhetorical moves in a research article is just a first step. Ultimately, among other things, readers should be able to connect prior knowledge to new information in the text, monitor their comprehension, and draw inferences during and after reading (Pearson *et al.*, 1992). However, by focusing on identifying rhetorical moves, we hopefully have ensured that the students will be able to

apply their skills to other disciplines. Because research articles "tend to be rhetorically standardized with regard to paragraph organization, choice of vocabulary and grammatical means of expression" (Knorr-Cetina, 1981, p. 95), we expect that, if students are able to recognize the rhetorical moves in biomedical research articles, they will also be able to recognize the rhetorical moves in research articles from related disciplines. In this way, they will develop what Bhatia (2004) calls "generic competence."

To our knowledge, our research is one of the first studies in which the identification of rhetorical moves in authentic empirical research articles by novice readers has been examined. Earlier studies have used rhetorical moves to teach students how to write genre-specific texts (e.g., Marshall, 1991; Henry and Roseberry, 1998) but not how to read these texts.

This paper has demonstrated how ideas from the field of genre analysis and argumentation theory may be used to improve undergraduates' reading strategies. Our teaching strategy with its focus on rhetorical structure may be an effective method for introducing novice readers to primary literature. A key issue in the future will be to improve students' ability to identify supports, counterarguments, and refutations. In future studies using task-based interviews, students' reading behavior; their use of lexical, rhetorical, and content features; and their ability to identify certain moves can be closely monitored. The outcome can be used to improve the instructional approach. A next step after the ability to identify the seven moves is to develop students' awareness of the links between these rhetorical moves by drawing a diagram illustrating these seven moves. This activity might help students to become aware of whether data are being correctly used to support the conclusion or whether supports are missing, thereby enabling students to criticize the authors' argument.

## ACKNOWLEDGMENTS

We thank J. H. Buikema, PhD, and Prof. R.H. Henning for giving us the opportunity to gather data during the Biomedical Research course of the first-year undergraduate program in life sciences at the University of Groningen, The Netherlands.

## REFERENCES

- Andrews R (2010). *Argumentation in Higher Education: Improving Practice through Theory and Research*, New York: Routledge.
- Bas O *et al.* (2007). The protective effect of fish n-3 fatty acids on cerebral ischemia in rat hippocampus. *Neurochem Int* 50, 548–554.
- Bazerman C (1988). *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*, Madison: University of Wisconsin Press.
- Berkenkotter C, Huckin T (1995). *Genre Knowledge in Disciplinary Communication: Cognition/Culture/Power*, Hillsdale, NJ: Lawrence Erlbaum.
- Bhatia VK (2004). *Worlds of Written Discourse: A Genre-Based View*, London: Continuum.
- Björk BC, Roos A, Lauri M (2009). Scientific journal publishing—yearly volume and open access availability. *Inf Res* 14, 391.
- Blanton WE (1990). The role of purpose in reading instruction. *Read Teach* 43, 486.

- Brill G, Falk H, Yarden A (2004). The learning process of two high-school biology students when reading primary literature. *Int J Sci Educ* 26, 497–512.
- Charney D (1993). A study in rhetorical reading: how evolutionists read “The spandrels of San Marco.” In: *Understanding Scientific Prose*, ed. J Selzer, Madison: University of Wisconsin Press, 203–231.
- Chen Q, Donin J (1997). Discourse processing of first and second language biology texts: effects of language proficiency and domain-specific knowledge. *Mod Lang J* 81, 209–227.
- Cohen L, Manion L, Morrison K (2008). *Research Methods in Education*, London: Routledge.
- Coil D, Wenderoth MP, Cunningham M, Dirks C (2010). Teaching the process of science: faculty perceptions and an effective methodology. *CBE Life Sci Educ* 9, 524–535.
- Collins A, Brown JS, Holum A (1991). Cognitive apprenticeship: making thinking visible. *Am Educ* 15, 6–11, 38–46.
- Connor U, Upton TA, Kanoksilapatham B (2007). Introduction to move analysis. In: *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*, ed. D Biber, U Connor, and TA Upton, Amsterdam: John Benjamins, 23–41.
- Du Boulay D (2012). Argument in reading: what does it involve and how can students become better critical readers? *Teach High Educ* 4, 147–162.
- Dudley-Evans T (1994). Genre analysis: an approach to text analysis for ESP. In: *Advances in Written Text Analysis*, ed. M Coulthard, London: Routledge, 219–228.
- Epstein HP (1970). *A Strategy for Education*, Oxford, UK: Oxford University Press.
- Erduran S, Jiménez-Aleixandre MP (eds.) (2008). *Argumentation in Science Education. Perspectives from Classroom-based Research*, Dordrecht, The Netherlands: Springer.
- Evagorou M, Dillon J (2011). Argumentation in the Teaching of Science. In: *The Professional Knowledge Base of Science Teaching*, ed. D Corrigan *et al.*, Dordrecht, The Netherlands: Springer, 189–201.
- Fang Z (2005). Scientific literacy: a systemic functional linguistics perspective. *Sci Educ* 89, 335–347.
- Flesch R (1948). A new readability yardstick. *J Appl Psych* 32, 221–233.
- Gillen CM (2006). Criticism and interpretation: teaching the persuasive aspects of research articles. *Cell Biol Educ* 5, 34–38.
- Gillen CM (2007). *Reading Primary Literature: A Practical Guide to Evaluating Research Articles in Biology*, San Francisco, CA: Pearson/Benjamin Cummings.
- Goldman SR, Bisanz GL (2002). Toward functional analysis of scientific genres: implications for understanding and learning processes. In: *The Psychology of Science Text Comprehension*, ed. J Otero, JA León, and AC Graesser, Mahwah, NJ: Lawrence Erlbaum, 19–50.
- Hayes AF, Krippendorff K (2007). Answering the call for a standard reliability measure for coding data. *Commun Methods Measures* 1, 77–89.
- Henry A, Roseberry RL (1998). An evaluation of a genre-based approach to the teaching of EAP/ESP writing. *TESOL Q* 32, 147–156.
- Hill SS, Soppelsa BF, West GK (1982). Teaching ESL students to read and write experimental-research papers. *TESOL Q* 16, 333–347.
- Holliday WG, Yore LD, Alvermann DE (1994). The reading-science learning-writing connection: breakthroughs, barriers, and promises. *J Res Sci Teach* 31, 877–893.
- Hoskins SG, Lopatto D, Stevens LM (2011). The C.R.E.A.T.E. approach to primary literature shifts undergraduates’ self-assessed ability to read and analyze journal articles, attitudes about science, and epistemological beliefs. *CBE Life Sci Educ* 10, 368–378.
- Janick-Buckner D (1997). Getting undergraduates to critically read and discuss primary literature. *J Coll Sci Teach* 27, 29–32.
- Kanoksilapatham B (2005). Rhetorical structure of biochemistry research articles. *Engl Specif Purp* 24, 269–292.
- Kelly G, Takao A (2002). Epistemic levels in argument: an analysis of university oceanography students’ use of evidence in writing. *Sci Educ* 86, 314–342.
- Knorr-Cetina K (1981). *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*, Oxford, UK: Pergamon.
- Kolokant YBD, Gatchell DW, Hirsch PL, Linsenmeier RA (2006). A cognitive-apprenticeship-inspired instructional approach for teaching scientific writing and reading. *J Coll Sci Teach* 36, 20–25.
- Kozeracki CA, Carey MF, Colicelli J, Levis-Fitzgerald M (2006). An intensive primary-literature-based teaching program directly benefits undergraduate science majors and facilitates their transition to doctoral programs. *Cell Biol Educ* 5, 340–347.
- Kuhn D (1991). *The Skills of Argument*, Cambridge, UK: Cambridge University Press.
- Kuhn D (2010). Teaching and learning science as argument. *Sci Educ* 94, 810–824.
- Kuldell N (2003). Read like a scientist to write like a scientist: using authentic literature in the classroom. *J Coll Sci Teach* 33, 32–35.
- Latour B (1987). *Science in Action*, Milton Keynes, UK: Open University Press.
- Lave J, Wenger E (1991). *Situated Learning: Legitimate Peripheral Participation*, Cambridge, UK: Cambridge University Press.
- Levine E (2001). Reading your way to scientific literacy. *J Coll Sci Teach* 31, 122–125.
- Marshall S (1991). A genre-based approach to the teaching of report-writing. *Engl Specif Purp* 10, 3–13.
- Norris SP, Phillips LM (2003). How literacy in its fundamental sense is central to scientific literacy. *Sci Educ* 87, 224–240.
- Nwogu KN (1997). The medical research paper: structure and functions. *Engl Specif Purp* 16, 119–138.
- Ozen OA *et al.* (2008). The protective effect of fish n-3 fatty acids on cerebral ischemia in rat prefrontal cortex. *Neurol Sci* 29, 147–152.
- Peacock M (2002). Communicative moves in the discussion section of research articles. *System* 30, 479–497.
- Pearson PD, Roehler LR, Dole JA, Duffy GG (1992). Developing expertise in reading comprehension. In: *What Research Has to Say about Reading Instruction*, ed. SJ Samuels and A Farstrup, 2nd ed., Newark, DE: International Reading Association, 145–199.
- Peck WH (2004). Teaching metastability in petrology using a guided reading from the primary literature. *J Geosci Educ* 52, 284–288.
- Robertson K (2012). A journal club workshop that teacher undergraduates a systematic method for reading, interpreting, and presenting primary literature. *J Coll Sci Teach* 41, 25–31.
- Sampson VD, Clark DB (2008). Assessment of the ways students generate arguments in science education: current perspectives and recommendations for future directions. *Sci Educ* 92, 447–472.
- Suppe F (1998). The structure of a scientific paper. *Philos Sci* 65, 381–405.

Swales JM (1990). *Genre Analysis: English in Academic and Research Settings*, Cambridge, UK: Cambridge University Press.

Tenopir C, King DW, Edwards S, Wu L (2009). Electronic journals and changes in scholarly article seeking and reading patterns. *Aslib Proc* 61, 5–32.

Thompson DK (1993). Arguing for experimental “facts” in science: a study of research article Results sections in biochemistry. *Writ Commun* 10, 106–128.

Topping KJ (1996). The effectiveness of peer tutoring in further and higher education: a typology and review of the literature. *High Educ* 32, 321.

Toulmin SE (1958). *The Uses of Argument*, Cambridge, UK: Cambridge University Press.

Van Lacum E, Ossevoort M, Buikema H, Goedhart M (2012). First experiences with reading primary literature by undergraduate life science students. *Int J Sci Educ* 34, 1795–1821.

Williams IA (1999). Results sections of medical research articles: analysis of rhetorical categories for pedagogical purposes. *Engl Specif Purp* 18, 347–366.

Yudkin B (2006). *Critical Reading: Making Sense of Papers in Life Science and Medicine*, London: Routledge.