Reading primary literature

van Lacum, Edwin Bernard

READING PRIMARY LITERATURE

Van Lacum, E.B.

Reading Primary Literature: Introducing Undergraduate Life Science
Students to the Rhetorical Structure of Research Articles

Omslagontwerp: Monique Koers (www.themocompany.nl)
Afbeelding op omslag gebruikt onder licentie van Shutterstock.com
Druk: Wöhrmann Print Service, Zutphen

ISBN: 978-90-367-6249-6
NUR: 846

www.edwinvanlacum.nl

rijksuniversiteit
groningen

Reading Primary Literature

Introducing Undergraduate Life Science Students to the Rhetorical Structure of Research Articles

PROEFSCHRIFT

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
vrijdag 28 juni 2013
om 16:15 uur

door

Edwin Bernard van Lacum
geboren op 19 augustus 1981
te Groningen

Promotor:	Prof. dr. M.J. Goedhart
Copromotor:	Dr. M.A. Ossevoort
Beoordelingscommissie:	Prof. dr. K.T. Boersma Prof. dr. C.M. de Glopper Prof. dr. R.H. Henning

Voor mijn ouders

Contents

Chapter 1
Introduction

1.1 General Aim

The general aim of this thesis is to develop and evaluate an effective teaching strategy for introducing first-year undergraduate students to research articles. Research articles are the canonical form for communicating original scientific results (Gross, Harmon, & Reidy, 2002). In science, the research article is the most important type of what is known as primary literature: reports of original observations, theories, or opinions, written for peers in the scientific community. In science, other types of primary literature besides research articles are opinion pieces and conference proceedings.[1]

Learning to read research articles is particularly important for students who will become scientists. The relevance of this ability has become more urgent in recent years, because scientists today have easy access to a rapidly growing amount of journal articles due to the internet and electronic databases.[2] Consequently, this has influenced their reading habits. Surveys indicate that the average number of annual article readings per scientist has increased: from 188 in 1993 to 280 in 2005 (Tenopir, King, Edwards,

[1] In contrast, secondary literature offers descriptions of primary literature. An important type of secondary literature is the review article. In such articles, authors present a summarized description of the primary literature that is available about a certain topic and provide comments on contradictions and controversies (Ziman, 1969; Noguchi, 2006). Other types of secondary literature are book chapters, textbooks, and journalistic reports.

[2] In the 20th century, the professionalization of science led to an enormous growth in the number of articles that are published each year in scientific journals (Lariviere, Archambault, & Gingras, 2008). Between the end of the Second World War and the end of the 1970's, the number of journal articles published each year grew exponentially. After that, the number of publications kept growing, although not exponentially. Björk, Roos, and Lauri (2009) estimated that 1,346,000 journal articles were published in 2006.

& Wu, 2009). However, the total time spent on reading primary literature has only increased marginally. This suggests that scientists are reading more selectively. Therefore, it is more important than ever for (future) scientists to develop skills that will enable them to read research papers in an efficient way.

Of course, not all students will pursue a career in science. However, the ability to read research articles may also be useful in other careers (e.g. teachers using research articles to stay informed of the latest developments in their discipline, government workers basing their policies on research, science journalists using research articles as a news source, and so forth). It has been suggested that reading research articles stimulates active learning (Gillen, Vaughan, & Lye, 2004), stimulates critical thinking (Gillen et al., 2004; Hoskins, Stevens, & Nehm, 2007), improves writing skills (Kuldell, 2003), and increases the understanding of how science works in practice (Baram-Tsabari & Yarden, 2005).

The ability to read research articles is a form of scientific literacy. Norris and Phillips (2003) state that scientific literacy is understood in two distinct ways. In the fundamental sense, scientific literacy means the ability to read and write scientific texts. In the derived sense, scientific literacy has a broader meaning: it relates to knowledgeability, learning, and education. Definitions of scientific literacy often emphasize the derived sense, as they encompass skills such as understanding science and its applications, knowledge of the risks and benefits of science, appreciation and comfort with science, and so forth. These definitions rarely refer to features of scientific literacy that fall under the fundamental sense. Norris and Phillips argue that this fundamental sense of scientific literacy is too often neglected in these definitions. According to them, reading and writing are not merely exchangeable tools for the conservation and communication of scientific information. Without reading and writing, the social practices that make science would not be possible. The use of texts allows scientists to record and (re)present data, preserve scientific knowledge, peer review, critical reexamine previously published ideas, et cetera. It is worth quoting Norris and Phillips' lengthy description of scientific literacy:

> Scientific literacy comprises both the concepts, skills, understandings, and values generalizable to all reading, and knowledge of the substantive content of science. (…) Interpretation of science text involves, to be sure, knowledge of substantive scientific content. However, substantive scientific knowledge is not enough for understanding. Generalizable concepts, skills, understandings, and values are also needed for reading and other intellectual tasks (…). Determining how scien-

> tific texts are to be taken involves determining, for example, when something is an inference, a hypothesis, a conclusion, or an assumption; when something is an asserted truth, an expressed doubt, or a proffered conjecture; when something is evidence for a claim, a justification for an action, or a stated fact to be explained. If these general meanings are missed, then the reader not only has read poorly, the reader has failed to grasp the scientific meaning beyond the surface content level and failed to grasp science. (p. 235)

According to Norris and Phillips, scientific literacy in the fundamental sense should be emphasized more 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. Thus, a failure to read science is also a failure to understand science. In this thesis, we will follow this broad definition by Norris and Phillips. This means that *reading* does not only include the passive absorption of information, but also active and complex interpretation processes such as analyzing and criticizing.

In this chapter, we will start with presenting an outline on the role and relevance of generic skills in higher education (Sections 1.2 and 1.3). Then, we will discuss the history of the research article and its characteristics (Sections 1.4 and 1.5). The use of research articles in higher education and the involved pedagogy will be discussed in Section 1.6, followed by our views on reading as an activity in the scientific community (Section 1.7). This chapter ends with an outline of the thesis (Section 1.8) and the description of the course in which we conducted our research (Section 1.9).

1.2 The Relevance of Generic Skills

Since the Second World War, Europe has become increasingly integrated. The main idea behind this process of integration (which has resulted in the creation of what is now called the European Union) is that it will assure peace, stability, and prosperity on the continent. In the 1990s it was felt that European integration was too much focused on economical aspects (e.g. the common market and the single currency). A more complete Europe had to be established; a Europe that would be able to strengthen its intellectual, cultural, social, scientific, and technological dimensions. This Europe of Knowledge would ensure social and human growth, consolidate and enrich European citizenship, give citizens competences to face the new millennium, and raise awareness of belonging to a shared social and cultural space (European Ministers of Higher Education, 1999). To achieve this, the Education Ministers from 29 European countries signed in 1999 the Bologna Declaration. It was the formal beginning of the Bologna Process, which aims "to strengthen the competitiveness and attractiveness of European higher education and to foster student mobility and employability" (Benelux Bologna Secretariat, 2009, p. 3). As a result, European higher education has been transformed in significant ways. Most importantly, the education systems of the different countries were made more compatible with each other. For example, a system of easily readable and comparable degrees was adopted. These reforms led in 2010 to the launch of the European Higher Education Area (EHEA).

In 2013, the Bologna Process is still in progress. During a recent policy meeting in Bucharest, Ministers from 47 European countries set the following goals regarding the further development of the EHEA: providing quality higher education for all, enhancing graduates' employability, and strengthening mobility as a means for better learning (European Ministers of Higher Education, 2012). To enhance employability, higher education should focus more on "generic skills and competencies like analytical skills, communication skills, ethical awareness, the ability to assess risks in a longer time perspective as well as the capacity to reason at a level of abstraction and to learn further" (Benelux Bologna Secretariat, 2009, p. 9).

Clanchy and Ballard (1995) present a systemic analysis of generic skills in higher education. They describe generic skills as "generic 'attributes' which graduates of an (…) university take from their institution into the workplace and the community" (p. 157). They further state that generic skills are higher order intellectual skills and should not be confused with lower order skills that comprise specific and often technical abilities (e.g. being able to work with a word processor). As Clanchy and Ballard state, generic skills can only be developed within specific contexts of knowledge. It is not possible to develop critical thinking skills *in absentia* of some subject matter. Furthermore, these contexts determine the form of those skills. For example, the skills researchers need to communicate to their peers depend on their discipline: researchers in the empirical sciences use research articles, while researchers in the humanities may use essays. So, science students need to be taught different genres than humanities students.

Clanchy and Ballard group generic skills (and attitudes) into three fields of activity: (1) analytical thinking and reasoning, (2) research, and (3) communication. The first field of activity, analytical thinking and reasoning, involves strategies such as questioning, hypothesizing, testing against evidence, and so forth. The objects of analytical thinking may be concepts, laws, models, and theories. The second field of activity is research (including methods of inquiry and management of information): library searching, experimental work, fieldwork, computer modeling, and so forth. Reading research articles also belongs to this field of activity. When doing research, it is important to approach knowledge in a critical way and to acknowledge the contributions of other researchers. The third field of activity is communication (oral and written), involving the ability (1) to choose an appropriate medium when communicating products of thinking and research and (2) to use the conventions of this medium. This includes skills such as using sources and data in a correct way and adopting a suitable tone of voice.

As the Bologna Process intends, generic skills should be given more attention in higher education. However, this is not always an easy task. An American study by Coil, Wenderoth, Cunningham, and Dirks (2010) showed that life science faculty members think that it is important for undergraduate science students to acquire skills such as the ability to interpret data, design experiments, solve problems, write reports, and critically analyze research articles. However, they are also of the opinion that it is very time-consuming to teach these skills. Since they feel a pressing need to cover content in their introductory courses, they are not able to pay much attention to the before-mentioned skills.

1.3 Using Research Articles to Connect Research and Teaching

To master generic skills (such as reading research articles), it is important that higher education students are trained in an environment where teaching and research take place. Universities have a long tradition in combining teaching and research. Many scholars from the past taught at universities, with Galileo and Newton as probably the most well-known examples.[3] The philosopher and Prussian minister of education Wilhelm von Humboldt is considered as one of the founding fathers of the modern university. In the beginning of the nineteenth century, he wrote a memorandum in which he laid down his views on the role of higher education institutions in society. According to von Humboldt, founder of the Humboldt University of Berlin, both students and teachers in higher education institutions "*sind für die Wissenschaft da*" (von Humboldt, 1903, p. 252). Teachers and students complement each other and by working together they can advance science.

Von Humboldt's idea that research and teaching should be connected to each other has in more recent times been called the research-teaching nexus (Neumann & Lindsay, 1987). In this nexus, research and teaching activities are blended (Clark, 1997). In the twentieth century, several devel-

[3] Galileo was by all accounts a popular lecturer (Reston, 1994). In contrast, Newton sometimes lectured to an empty classroom. As Humphrey Newton, his assistant, noted, "so few went to hear Him, & fewer y^{t} understood him, y^{t} oftimes he did in a manner, for want of Hearers, read to y^{e} Walls" (as cited in Westfall, 1983, p. 209).

opments in higher education negatively influenced the conditions that are necessary for the implementation of the research-teaching nexus (Clark, 1997). For example, the knowledge needed in order to do research became increasingly complex. This means that students need to be very knowledgeable before they can participate in research. This has led to an emphasis on knowledge acquisition during the first period of university training, resulting in a separation between teaching and research.

Nevertheless, many educators and policy makers have argued that universities should strive to realize von Humboldt's ideas. Not because they necessarily share von Humboldt's Enlightenment ideals, but rather because they believe that linking teaching and research will enhance students' personal development and employment perspectives (Hammond, 2007).

A number of European and North-American universities have addressed the research-teaching nexus in their policy documents. In these documents, it is stated that students should acquire an understanding of recent research results, the curriculum should be inspired by research done in the institute, and students should be given opportunities to conduct research and become members of a disciplinary community as early as possible (Elsen, Visser-Wijnveen, van der Rijst, & van Driel, 2009). Strengthening the link between research, teaching, and learning is also a priority of the Bologna Process: "Study programmes must reflect changing research priorities and emerging disciplines, and research should underpin teaching and learning" (European Ministers of Higher Education, 2012, p. 2).

Connecting teaching and research can be done in different ways. According to Griffiths (2004), teaching where this connection is made can be subdivided into four categories:

- Research-led teaching: There is an emphasis on subject content and understanding research findings (instead of research processes). The selection of content is based on the lecturers' research interests. Teaching is seen as the transmission of information.
- Research-oriented teaching: There is an emphasis on knowledge and on understanding how this knowledge is produced. The development of inquiry skills and a research ethos is stimulated.
- Research-based teaching: There is an emphasis on inquiry-based activities. Students and teachers are working together as much as possible during these activities.
- Research-informed teaching: Results of research into education and learning are used during teaching.

Healey (2005) has represented these different categories of teaching schematically along two axes (leaving out research-informed teaching, probably because it applies to educational research and not necessarily to research in which the teacher is a specialist, as in the other types). One axis represents the role of students (as participants or audience members) and the other axis represents if the emphasis is on research content or research processes and problems. Using this diagram, Healey identified a fourth category, research-tutored teaching, where students are participants and the emphasis is on research content.

Jenkins and Healey (2010) presented a slightly adapted version of this diagram (Figure 1.1). They describe the different categories of teaching as follows: "Research-led: learning about current research in the discipline; research-oriented: developing research skills and techniques; research-based: undertaking research and inquiry; and research-tutored: engaging in research discussions" (p. 38).

These four categories of teaching are not mutually exclusive (Jenkins & Healey, 2010). For example, students may learn about research skills and techniques (research-oriented strategy) and engage in research discussions (research-tutored strategy) at the same time.

Jenkins and Healey (2010) note that a growing number of higher education institutions have increased the share of research-tutored and research-based teaching by developing inquiry-based curricula, letting students work on research projects, offering research-based courses, or by focusing on problem-based learning. Despite these initiatives, the potential connections between teaching and research are often not sought or realized – even though national and international policy statements emphasize the value of learning in research environments. As Jenkins and Healey (2010) note, students in higher education are too often mere audience members instead of participants. According to them, curricula

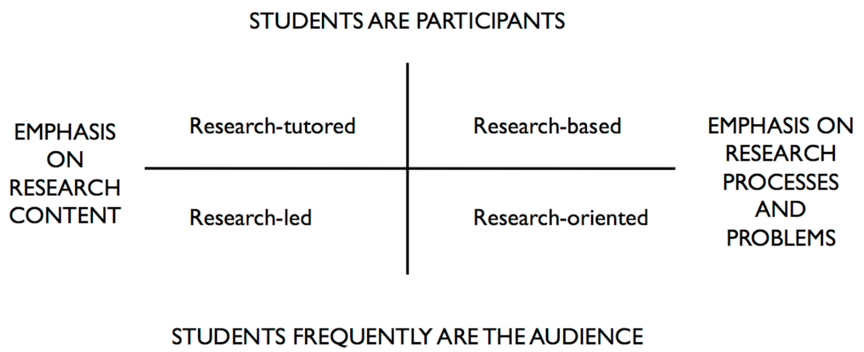


Figure 1.1. Categories of teaching that link teaching and research (adapted from Jenkins & Healey, 2010).

should contain a balanced mixture of all four ways of teaching: "All four ways of engaging students are valid and valuable, and we think curricula should contain elements of all of them" (p. 38). In this way, the variance in students' learning styles is catered for. Ultimately, a balanced mixture of teaching categories in a curriculum will strengthen the research-teaching nexus (Elsen et al., 2009).

In this thesis, we let students read, interpret, and discuss research articles while focusing on argumentation. Looking at Figure 1.1, this activity is best qualified as research-tutored and research-based as students respectively engage in research discussions and undertake inquiry. In many reform-oriented classrooms in the United States, there is less emphasis on texts because educators prefer "doing science" above "reading science" (Pearson, Moje, & Greenleaf, 2010). This preference is based on the assumption that reading is a process in which information is passively received. According to Norris and Phillips (2003), this can be classified as a simple view on reading. In this view, reading is nothing more than knowing the words and locating information in the text. However, reading can enhance rather than diminish the acquisition of knowledge and inquiry in science (Pearson et al., 2010). Inquiry-driven literate practice is "a process of actively making meaning of science; these inquiries are best regarded as investigations in their own right" (Pearson et al., 2010, p. 460).

In the next two sections we will discuss the nature of research articles and explain why certain features of research articles make it difficult for novice readers to read texts of this genre.

1.4 The Nature of Research Articles

The first scientific journals appeared at the end of the 17th century. Since then, the genre of the research article has gone through many changes (Gross et al., 2002). Bazerman (1988) analyzed research articles that were published in the *Philosophical Transactions of the Royal Society of London* between 1665 and 1800. In the beginning of this period, Bazerman claims, the experiments in these research articles were not more than news: "For those actively pursuing nature, nature was portrayed as speaking for herself. The scientific report was simply a matter of news. Just as an earthquake or passage of a meteor needed to be reported, so did experiments" (p. 77). Later on, nature began to be treated more and more as a matter of discussion. Authors of research articles recognized that they were not simply reporting the truth, but were "telling a story that can be questioned and that has a meaning which itself can be mooted" (p. 78). In effect, descriptions of experiments began to be used as persuasive means. This led to a number of changes in the genre. For example, more emphasis was placed on the methodology and the results of experiments: "As disputes arise over reported results, writers become more careful about reporting what they see, and measurement takes a greater role" (p. 72).

Bazerman (1988) also analyzed research articles published between 1893 and 1980 in *Physical Review* on the subject of spectroscopy. In this case, he looked specifically at the integration of theory into the text. In the beginning of this period, references were located primarily at the beginning of articles. These references were merely a "roll-call" of related articles about the same subject; they did not refer to specific findings and they did not identify a specific relation to the article. Later on, the references spread to other parts of the article and were used to discuss specific results from other studies. According to Bazerman, this indicates that journal articles became increasingly embedded in literature. Furthermore, articles gained specific theory and discussion sections. Before the rise of the Discussion section, articles ended with the results, "as though the results could stand alone and complete in their meaning" (p. 174).

To study the development of text features of journal articles over time, Berkenkotter and Huckin (1995) studied 350 research articles from 12 different journals that were published between 1944 and 1989. They observed that titles became more informative (e.g. by announcing the main results). They also saw that between 1944 and 1989, 11 journals started with the inclusion of Abstracts (the other journal started in 1920 with this feature). Furthermore, Abstracts became longer and contained more results statements. The Introduction section also underwent changes: authors were more likely to include a statement of their main finding in the Introduction section. Interestingly, Method sections were increasingly de-emphasized by moving them to a non-traditional location or by decreasing the font size. Berkenkotter and Huckin also observed that background information and raw data became less prominent. They conclude that all these changes aimed to emphasize the most important, newsworthy information of a research article. This is necessary, because scientists are nowadays "deluged with information" (p. 39) and do not have much time to read a research article.

These studies show that the features of research articles are neither arbitrary nor inevitable. Instead, these features are socially negotiated and reflect the changes in scientists' views on science, the increasing professionalization of science, and the growth of knowledge (Bhatia, 1997; Goldman & Bisanz, 2002).

Several authors have written about the role of language in science and how authors of research articles use rhetorical strategies (i.e. the effective use of language to persuade readers). Gould (1993) distinguishes two aspects of scientific persuasion: (1) form and language and (2) logic and empirical content. (In practice, these two aspects are difficult to separate. Form and language shape the logic and empirical content and vice versa.) Scientists think, Gould argues, that the logic and empirical content of a research article determine the strength of the claims that are presented; these aspects will persuade the reader to accept or reject a certain claim. In contrast, the first aspect is not really appreciated by scientists:

> Scientists, for the most part, simply do not acknowledge that the form and language of an argument (as opposed to its logic and empirical content) could have anything to do with its effectiveness. Humanists probably do not know that scientists even define the word *rhetoric* in an exclusively pejorative way – as an attempt to bamboozle by words alone, when you don't have the goods in logic or data.[4] (Gould, 1993, p. 323)

According to Miller (1979), scientists believe that language is a kind of windowpane. Scientists think that "if language is clear, then we see reality accurately; if language is highly decorative or opaque, then we see what is not really there or we see it with difficulty" (Miller, 1979, p. 612). According to Miller, this view of language as a windowpane is strongly rooted in positivism. According to this philosophical viewpoint, observations are neutral and independent of background beliefs and theory (Brown, 2001). Miller (1979): "In this [positivistic] epistemology, language, based as it is in personal psychology, is largely a distraction for science; and rhetoric is just irrelevant, because conclusions follow necessarily from the data of observation and the procedure of logic" (p. 612).

In the late 1970s and the 1980s a number of authors have studied how a research article is created and how its content is shaped by rhetorical considerations (i.e. considerations about the effective use of language) in real-life settings (Latour & Woolgar, 1979; Knorr-Cetina, 1981; Gilbert & Mulkay, 1984). These authors show why it is spurious to view a research article – in its ideal form – as a rhetoric-free and objective account of laboratory proceedings.

For example, Knorr-Cetina (1981) described in a detailed case-study how processes of recontextualisation and decontextualisation contribute to the creation of a chemistry research paper about the recovery of proteins. In the Introduction section of this paper, the authors state that they instigated their study because there is a problem that needs a solution.

4 As Charney (1993) has shown, some scientists are wary of Gould's own mastery of form and language. These scientists view Gould's style as a smoke-screen: he uses his eloquence to obfuscate his arguments.

However, this did not correspond with the actual reality of the research process: the authors encountered by chance a solution for a certain problem on which they originally did not work. In contrast to the impression that is given in the Introduction section, it was not the authors' intention to solve this specific problem. This is not a misrepresentation, but a recontextualisation that serves as a literary strategy to persuade readers (i.e. transforming context). Furthermore, the Introduction section is also decontextualized (i.e. leaving out context). Knorr-Cetina describes how the authors' reasons to instigate their research were sometimes quite personal. For example, one of the authors wanted to do the study because he thought it was good for the advancement of his career. However, all reasons other than the recontextualized one were left out in the Introduction section by the authors.

In the Method and Results sections no recontextualisation takes place, only decontextualisation occurs. According to Knorr-Cetina, these sections "are marked by a conspicuous avoidance of arguments which might ground their assertions" (p. 113). For example, the Method section does not dwell on why certain materials and techniques were used. Instead, there is a focus on laboratory equipment and the origins of certain techniques. This strategy constitutes to the readers' "tendency to believe that the research topic alone, and not its 'internal' execution, is a matter of social choice and negotiation" (p. 118). In other words: it is made to appear that the execution of a study (the choice of materials and methods) always follows logically from the research topic.

Knorr-Cetina also states that the distinction in research articles between the Method, Results, and Discussion section is rather artificial: "The scientists observed did not first perform experiments, then obtain results and finally interpret the outcome" (p. 121). Some authors go even further. For example, the immunologist Peter Medawar once famously described the scientific paper as a fraud (Medawar, 1996). According to him, research articles give a misleading representation of scientific discoveries. With some exaggeration, Medawar describes research articles as follows:

> The section called 'results' consists of a stream of factual information in which it is considered extremely bad form to discuss the significance of the results you are getting. You have to pretend firmly that your mind is, so to speak, a virgin receptacle, an empty vessel, for information which floods into it from the external world for no reason which you yourself have revealed. You reserve all appraisal of the scientific evidence until the 'discussion' section, and in the discussion you adopt the ludicrous pretence of asking yourself if the information you have collected actually means anything; of asking yourself if any general

> truths are going to emerge from the contemplation of all the evidence you brandished in the section called 'results'. (p. 33–34)

As Markel (1993) states, the research article is a reconstruction of research, rather than a straightforward account. Therefore, it is not correct to speak of fraud. A more useful approach may be Gilbert and Mulkay's (1984) distinction between an empiricist repertoire and a contingent repertoire. In research articles, the empiricist repertoire is dominant. This repertoire may be characterized as formal and impersonal: the authors' judgments and actions are glossed over as much as possible. This gives the reader the impression that "the physical world seems regularly to speak, and sometimes to act, for itself" (p. 56). When scientists are engaged in conversations about their research, they often use the contingent repertoire. Then, they present "their actions and beliefs as heavily dependent on speculative insights, prior intellectual commitments, personal characteristics, indescribable skills, social ties and group membership" (p. 56).

1.5 Learning to Read Research Articles

Language

There are several reasons why novice readers may have difficulty with reading research articles. The first of these reasons is the widespread use of Scientific English, a register of the modern English language (Halliday, 1993). Fang (2005) describes four register features that can cause reading difficulties with novice readers.[5] Firstly, scientific text has a high density of information. Sentences often contain a high number of nouns and extended noun phrases. This is an efficient way of communication for experts, but it may also lead to confusion with novice readers, because the syntactic clues that specify the semantic relationships between the nouns are left out (Gross et al., 2002). Secondly, scientific text is often abstract. Concrete experiences are turned into abstract entities via the conversion of verbs and adjectives into nouns. For example: *wheeze and become short of breath* are synthesized into *asthma attacks*. This process, called nominalization, can "create problems for readers, because it tends to neutralize or obscure meanings and construct an ideology that is often not transparent to naïve readers" (Fang, 2005, p. 340). Thirdly, scientific texts contain a high number of technical terms. Students who are not familiar with these specialized terms will struggle with comprehending the text. Fourthly, scientific text is authoritative (Meadows, 1985). Authors try to distance themselves from the text, for instance by avoiding

[5] See also Shymansky, Yore, and Good (1991), who describe some additional features of scientific language that could hinder novice readers.

first-person references. This can alienate students, who are often used to more personal and informal texts.[6]

Persuasiveness

Persuasiveness is the second reason why novice readers may have difficulty with reading research articles. As indicated in Section 1.4, research articles are persuasive in nature. Authors use experimental data to convince readers that the presented claims are justified. Sometimes they argue vehemently why their interpretation of the data is the only correct one. Additionally, they use references to other studies to consolidate their conclusions (Latour, 1987).

As Gould (1993) states, scientific persuasion involves (1) form and language and (2) logic and empirical content. The use of persuasive language in research articles demands of readers specific skills helping them with decoding the intentions of authors. For example, one of the features authors of research articles use, are so-called hedges: words such as *may*, *might*, or *possibly*, which indicate the (un)certainty of statements. As Hyland (1998b) remarks:

> Hedges are a crucial means of presenting new claims for ratification and are among the primary features which shape the research article (…) as the principal vehicle for new knowledge. Hedging enables writers to express a perspective on their statements, to present unproven claims with caution and to enter into a dialogue with their audiences. (p. 6)

Students do not have much experience with hedges, because they are mostly used to reading textbooks, which are written with the aim to present factual statements that are widely agreed upon in the scientific community (Goldman & Bisanz, 2002; Gillen, 2006). In general, textbooks tend to neglect the processes by which scientific knowledge is produced (Duncan, Lubman, & Hoskins, 2011). The style of textbooks "sweeps away

[6] Interestingly, some high-profile scientists have criticized the impersonal style of research articles. David Mermin (1990) notes that the "insistence on bland impersonality and the widespread indifference to anything like the display of a unique human author in scientific exposition, have not only transformed the reading of most scientific papers into an act of tedious drudgery, but have also deprived scientists of some powerful tools for enhancing their clarity in communicating matters of great complexity" (p. xi–xii). Roald Hoffmann (2002) describes the scientific article as a "mechanical, ritualized product that (...) propagates the notion that scientists are dry and insensitive, that they respond only to wriggles in a spectrum" (p. 48). This is why Hoffmann argues for a humanization of the publication process and relaxing the strictures on expressing emotion, personal motivation, and historicity.

the memories of doubt or difficulty over what might be taken as true, and makes 'the facts' appear to be completely outside human agency" (Sutton, 2003, p. 30). As a result, academic texts are viewed by students as "autonomous, uncontested and unnegotiated, unencumbered by the values and oppositions that they may freely recognize in their out-of-school lives and textual experiences" (Johns, 2002, p. 239–240). However, for the interpretation of a research article, one needs to develop a deep understanding of hedges and other specific rhetorical strategies used by authors.

Contextual Features

The third reason why novice readers may have difficulty with reading research articles is that to fully understand a research article one has to be knowledgeable about its context. Research articles are no self-contained entities; they are part of a specific scientific discourse. For example, it is difficult for novice readers to judge the relevance of a research article because much of this relevance is contextual (Fahnestock, 1986). Relevance is not explicitly mentioned, but "supplied by context, by the assumed inferences the intended audience will make" (p. 278). As an example Fahnestock quotes Watson and Crick's seminal article about the structure of DNA (Watson & Crick, 1953). Watson and Crick did not spell out the consequences of their revolutionary discovery because readers were expected to figure this out themselves.

The language, persuasiveness, and contextual features of research articles suggest to us that higher education institutions should use specific teaching strategies that will introduce science students to primary literature and will improve their level of scientific literacy. Before we will describe the development of such a teaching strategy, we will discuss a number of higher education courses with similar aims.

1.6 Courses in Higher Education Introducing Primary Literature

We found a small number of studies describing higher education courses that aim to introduce students to scientific primary literature.[7] We have summarized these 15 studies in Table 1.1. These studies provide us with use-

7 We used the search engines ERIC (Education Resources Information Centre, available via http://www.eric.ed.gov) and Google Scholar (http://scholar.google.com) to get an overview of studies describing courses that aim to introduce higher education science students to primary literature. We used combinations of the following keywords: "reading," "literacy," "primary literature," "science," "scientific," "undergraduate," "higher education," and "college." Studies that described courses that only used primary literature in a minor way (e.g. as an additional information source for a lab assignment) were left out.

Table 1.1. Overview of studies describing primary literature courses.

Study	Subject	Participants	Teaching strategy	Measurement of effectiveness
Janick-Buckner (1997)	Cell biology	Majors (*n* = 16)	Writing critical review, group discussion	Evaluation
Herman (1999)	Microbial genetics	Undergraduate students (unknown number)	Guided reading, group discussion	Classroom observations
Houde (2000)	Animal behavior	College students (*n* = 14)	Participating in a symposium	Evaluation
Levine (2001)	Molecular genetics	Junior biology majors (unknown number)	Guided reading, group discussion (jigsaw approach)	Classroom observations, evaluation
Smith (2001)	Evolution and ecology	Biology majors (unknown number)	Guided reading (in groups)	Evaluation
Peck (2004)	Petrology	Undergraduate geology students (unknown number)	Guided reading, group discussion	Evaluation
Kuldell (2003)	Molecular genetics	College students (*n* = 39)	Guided reading, group discussion (jigsaw approach)	Classroom observations, evaluation
Mulnix (2003)	Cell physiology	Biology and biochemistry majors (*n* = 77)	Making a poster (in pairs or triplets)	Evaluation
Almeida & Liotta (2005)	Organic chemistry of the cell	Sophomore biochemistry and biology majors (*n* = 55)	Guided reading, group discussion	Evaluation, academic performance
Kozeracki, Carey, Colicelli, & Levis-Fitzgerald (2006)	Biology	Science majors (*n* = 40)	Giving presentations, group discussion	Evaluation, academic performance
Hoskins, Stevens, & Nehm (2007)	Biology	Junior or senior biology majors (*n* = 51)	Guided reading, group discussion	Test (pre and post)
Jacques-Fricke, Hubert, & Miller (2009)	Molecular biology	Third or fourth year biology majors (*n* = 70)	Group discussion (jigsaw approach)	Evaluation, academic performance
Roberts (2009)	Molecular and nutritional biochemistry	Undergraduate students (*n* = 27)	Giving presentations	Evaluation
Hoskins, Lopatto, & Stevens (2011)	Biology	Junior or senior biology majors (*n* = 189)	Guided reading, group discussion	Test (pre and post)
Wenk & Tronsky (2011)	Drugs in the nervous system	First-year college students (*n* = 41)	Guided reading, group discussion (jigsaw approach), giving presentations	Test (pre and post)

ful information about introducing primary literature to students. In 13 of the 15 studies the participants were undergraduate students (Kuldell, 2003, and Houde, 2000, do not specify if their participants are undergraduate or graduate students). The studies we found frequently described courses using a combination of individual guided reading (e.g. answering guiding questions about certain aspects of the research article) and group discussions.

In some cases, the guiding questions were quite straightforward and required only basic comprehension of the research article, as they were designed to make sure that students read the articles before the group discussion. For example, Levine (2001) asked: "How does the data in figure 2 provide an explanation for the differences seen in treatment effectiveness for patients 1 and 2?" (p. 124). In other cases, the guiding questions were more ambitious. For example, Peck's (2004) questions were "designed to step the student through the approach and logic used in the paper requiring progressively more sophisticated understanding and reflection on the methods, assumptions, and wider implications of the research" (p. 284). In the studies where no guiding questions were used, other methods were used to engage students with the articles, such as writing critical reviews (Janick-Bucker, 1997) or making a poster (Mulnix, 2003). In these cases, students were provided with guidelines by the course instructors.

In 11 of the 15 studies group discussions were used. Usually, the group discussions were led or moderated by a teacher. For example, Janick-Buckner (1997) discussed a research article during three sessions of 80 minutes. During these sessions, she discussed with the students the reasons why the authors conducted their study, the experimentation, the figures and tables, the results, the authors' interpretation of the results, and what should be done next to further this line of research. She observed that students initially addressed their questions and comments to her. Later on, students directed their questions and comments to the entire group. Consequently, Janick-Buckner tried to act more like a moderator and less like a leader of the discussion. She also noted that it was important to create an atmosphere of trust, so that students felt encouraged to discuss comprehension difficulties. For this purpose, she did not ask a student a question unless she was sure he or she could provide an answer. In other studies a jigsaw approach was used. For example, in Levine's (2001) study, students were divided into groups and each of the group members had to answer questions about a different part of the research article at home. During class, the students discussed the article in groups and made an assignment for which they used their combined knowledge.

The studies presented in Table 1.1 have a number of limitations. For example, the effectiveness of these courses is often determined by evaluations in which the students assess the course and/or their own abilities via questionnaires; a systematic evaluation of learning outcomes is lacking. One of the exceptions is Wenk and Tronsky's (2011) study. By analyzing students' answers, they measured before and after an introductory natural science course how well they are able to explain a research article's (1) research question/importance, (2) hypothesis, (3) setup (design), (4) collected data, (5) results supporting the hypothesis, (6) alternative explanations, and (7) future research suggestions. After the course, students were better able to explain the hypothesis, setup, collected data, and results. Students still had difficulty with proposing alternative explanations and future research suggestions.

Furthermore, the educational aims are often poorly defined in the studies we found. Therefore, it is unclear what kind of skills or abilities the authors want to improve. The studies may be characterized as good-practices: they describe the content of the course, but students' products and the group discussions are not systematically analyzed. The studies also often lack descriptions of a theoretical framework: it is not clear which teaching models were used to design the course.

The results of our literature study suggest to us that studies describing primary literature courses are relatively rare. This is noteworthy, because reading research articles seems to us an essential skill for science students in higher education. It is our impression that in most science curricula there are no specific courses aimed at teaching students to read primary literature. It seems that the usual pedagogy with regard to reading research articles is "learning by doing." Taken together, our literature study shows the need for the development and systematic evaluation of specific teaching strategies for introducing students to primary literature.

1.7 Reading as Activity in the Scientific Community

As shown above, studies conducted into the reading of scientific texts by novice readers, lack theoretical frameworks. Gaining experience with reading research articles can be described with concepts from activity theory. Using this framework, the scientific community can be defined as a multi-voiced "activity system": a community of multiple points of view, traditions and interests (Engeström, 2001). The idea of an activity system is derived from cultural historical activity theory, a powerful framework which can be used to analyze complex (educational) settings and which was developed by Alexei N. Leont'ev and Sergei Rubinstein (based on the ideas of Vygotsky). Russell (1997) describes an activity system as follows:

> An activity system is any ongoing, object-directed, historically-conditioned, dialectically-structured, tool-mediated human interaction: a family, a religious organization, an advocacy group, a political move-

> ment, a course of study, a school, a discipline, a research laboratory, a profession, and so on. These activity systems are mutually (re)constructed by participants using certain tools and not others (including discursive tools such as speech sounds and inscriptions). (p. 510)

From the perspective of activity theory Russell (1997) describes how people learn to write in a certain genre. According to Russell, writing texts requires certain tools (e.g. lexical and syntactic items). These tools can be used again: “If the subject perceives conditions as the same or similar, the subject may act in the same way again” (p. 515). In this way, the use of these tools may become routine: the first time one has to think consciously about the tools he or she uses. Later on, it may become an unconscious routine operation. “For experienced insiders in a profession, for example, their ways of writing may be so routine that they come to seem natural. In this sense, genre helps account for social-psychological stability, identity, and predictability” (p. 515). When newcomers write in a certain genre, a process of appropriation – the picking-up of a tool-in-use from others – will be occurring:

> For newcomers to an activity system – like students learning to write a new genre in a new discipline or profession (to them) – the new ways they use these tools called words are encountered at the level of conscious actions. Through continued interaction with others in the activity system, the ways of using the tool (say, the introduction, methods, results, discussion [IMRD] structure in science) becomes a routine operation, often unconscious. Moreover, as an individual appropriates (learns to use) the ways with words of others, they may (or may not) also appropriate the object/motive, and subjectivity (identity) of the collective, of a new activity system. (p. 516)

It is not difficult to imagine that a similar process takes place when newcomers learn to read texts belonging to a certain genre. Via the process of appropriation the reader learns what features of a text are important, how to interpret specific semiotic constructions, how to evaluate the content, and so forth. This all takes place through interactions with the author (via the text), peers, and teachers.

Appropriation can be described as the transformation of reading strategies into reading skills. Afflerbach, Pearson, and Paris (2008) define reading skills as “automatic actions that result in decoding and comprehension with speed, efficiency, and fluency and usually occur without awareness of the components or control involved” (p. 368). In contrast, they define reading strategies as “deliberate, goal-directed attempts to control and modify the reader’s efforts to decode text, understand words, and construct meanings of text" (p. 368). When a reader gets more experienced, the use of specific actions while reading will become more automatic and less deliberate and effortful. However, every reader (no matter how experienced) may come across texts that are difficult to read. Then, he or she is forced to use reading strategies.

By reading research articles, students gain generic practice: "…the ability to respond to recurrent and novel rhetorical situations by constructing, interpreting, using and often exploiting generic conventions embedded in specific disciplinary cultures and practices to achieve professional ends" (Bhatia, 2004, p. 144). In this way, students are slowly becoming part of a community of practice. Enculturation into a community of practice is described by Lave and Wenger (1991) as "legitimate peripheral participation."

1.8 Research Questions and Design

In the previous section we have described the importance of enhancing higher education students' scientific literacy (in its fundamental sense). We also have sketched the historical development of research articles and listed a number of reasons why novice readers have difficulty with reading research articles. We described the studies in which primary literature courses are presented and concluded that there is need for more research into the development and evaluation of effective teaching strategies that aim to introduce students to primary literature (i.e. the appropriation of a genre). This leads to the primary research question of this thesis:

How can undergraduate life science students be taught to read research articles?

In our research we were primarily concerned with students' grasp of how research articles are structured around rhetorical moves. Rhetorical moves are used to describe genres in genre analysis studies (described in more detail in Chapter 3). A rhetorical move may be defined as "a section of a text that performs a specific communicative function" (Connor, Upton, & Kanoksilapatham, 2007, p. 23). Our reason for following this approach was that several authors have suggested a clear relation between reading ability and knowledge about the structural characteristics of a text (Hill, Soppelsa, & West, 1982; Samuels et al., 1988; Blanton, 1990; Swales, 1990; du Boulay, 1999). 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).[8] Swales

[8] It has to be noted that some authors are rather critical of using this so-called "genreist" approach for the purpose of teaching students how to write a certain type

calls the recognition of the rhetorical structure (i.e. the arrangement of rhetorical moves) of a genre-specific text *rhetorical consciousness*. Because rhetorical consciousness is very important for reading, we use this concept as a tool to teach students how to read a research article. In other words: we want to teach them reading strategies in which they learn how to identify certain rhetorical moves. This is also in line with Norris and Phillips (2003), who stated that a scientific literate person (see Section 1.1) should be able to determine the different types of statements in a scientific text (hypothesis, evidence, conclusion, expressed doubt, et cetera). As research articles are persuasive in nature, we want to focus especially on rhetorical moves that play an important role in the authors' argumentation. This will ensure that students will become familiar with the persuasive aspects of research articles – as explained in Section 1.5, persuasiveness is one of the possible reasons why novice readers find it difficult to read research articles.

To answer our primary research question, we followed an educational design research approach in which we used so-called design experiments. Design experiments are particularly useful for studying students' learning processes in complex, real world situations (Collins, Joseph, & Bielaczyc, 2004) and evaluating a teaching strategy in a systematic way. Van den Akker (1999) lists two important goals of educational design research: (1) providing ideas for the optimization of the quality of an intervention and (2) generating, articulating, and testing design principles. In our research, we were mostly led by the second goal: determining the effectiveness of our design principles.

> [Design] principles are usually heuristic statements of a format such as: 'If you want to design intervention X [for the purpose/function Y in context Z], then you are best advised to give that intervention the characteristics A, B, and C [substantive emphasis], and to do that via procedures K, L, and M [procedural emphasis], because of arguments P, Q, and R.' (van den Akker, 1999, p. 9)

As is customary in educational design research, we began with a preliminary investigation (van den Akker, 1999). During this investigation, described in Chapter 2, we have determined to what extent students are able to identify two important rhetorical moves in research articles: conclusions and supports. We also studied the features students used to identify these

of text. These opponents argue that students should not be coerced into using stilted forms of text (for an overview of this discussion, see Purcell-Gates, Duke, & Martineau, 2007).

moves and described their reading behaviors. Additionally, we compared students' performance with expert readers. We developed a set of design principles based on Chapter 2's findings and ideas from argumentation theory, genre analysis, and cognitive apprenticeship. The cognitive apprenticeship perspective was chosen because learning to read research articles is a process of enculturation: students will adopt the cultural practices of a new social group (Brown, Collins, & Duguid, 1989). Cognitive apprenticeship is a useful teaching method to stimulate enculturation (Brown et al., 1989; Collins, Brown, & Holum, 1991). The design principles (described in Chapter 4) were used to develop a teaching strategy that aimed to improve students' ability to read research articles. For this purpose, we developed a heuristic that could be used for the teaching strategy. This heuristic, the Scientific Argumentation Model, depicts and describes the different rhetorical moves of a research article (Chapter 3). Our teaching strategy was implemented in Reading Research Articles, a module for undergraduate life science students. The effects of this module on students' ability to identify rhetorical moves were measured via a pre- and post-test instrument (Chapter 4). Furthermore, students characterized the change in their own reading behaviors, so that we could determine if they read more like an expert. In Chapter 5, we described the features students use to identify rhetorical moves. This leads to the first three of our five secondary research questions:

1. How does our teaching strategy influence students' ability to identify rhetorical moves in research articles?
2. Which features of rhetorical moves do students use for their identification?
3. How do students' reading behaviors change during the implementation of our teaching strategy?

An integral part of the Scientific Argumentation Model is a scheme that helps students to understand the relations between rhetorical moves in research articles. In Chapter 6 we described students' ability to construct such an argumentation scheme. So, our fourth secondary research question is:

4. What is students' ability to construct an argumentation scheme of a research article?

An important aspect of reading research articles (at least for experts) is the ability to read critically. In Chapter 7, we described to what extent undergraduate life science students are able to critically evaluate research articles. So, our fifth secondary research question is:

5. What is students' ability to critically evaluate research articles?

To summarize, the first secondary research question is answered in Chapters 4, 5, 6, and 7. The second secondary research question is answered in Chapters 2 and 5. The third secondary research question is answered in Chapter 4. The fourth secondary research question is answered in Chapter 6, while the fifth secondary research question is answered in Chapter 7.

1.9 Educational Setting

Our preliminary research and implementation of the module took place in the undergraduate course Biomedical Research. This course was part of the last quarter of the first-year bachelor programs of Biology and Life Science and Technology at the University of Groningen (a research university located in the Netherlands).[9] Students who followed the course had chosen to major in biomedical sciences or behavioral and neurosciences. The subjects of the course were related to physiology and pharmacology (with a focus on the cardiovascular system). The course was designed in 2007 by two medical faculty members from the department of Clinical Pharmacology. When designing the course, they wanted to stimulate the development of students' generic skills. One of the generic skills they decided to focus on was reading research articles. The rationale for this approach was their viewpoint that students should be introduced with research articles as early as possible during their academic training. Educational aims of the course were:

- Gaining knowledge about the (patho)physiology and pharmacology of the cardiovascular system.
- Being introduced to the possibilities and limitations of in vitro animal research.
- Gaining research skills during lab assignments.
- Gaining academic skills: summarizing, question and problem statement, debating, evaluation of methodology, and abstracting. ("Biomedisch onderzoek," n.d., own translation)

The course was built around lectures (related to the first educational aim), lab work (related to the second and third educational aims), and tutor group meetings (related to the first, second, and fourth educational aims).

[9] In the Netherlands there are professional and academic bachelor degrees. The professional bachelor program takes four years and is followed at a university of applied sciences (*Hogeschool*). The academic bachelor program takes three years and is followed at a research university (*Universiteit*). Typically, students with a professional bachelor degree enter the labor market, while students with an academic bachelor degree enter a master program of two years.

The tutor group meetings were aimed to give students experience in reading research articles about the cardiovascular system. Before each tutor group meeting, students read a research article and made an assignment. This assignment was discussed during the next tutor group meeting. At the end of the course, students had to pass a multiple-choice examination and an oral examination. During the oral examination, students abstracted a research article. Although the research articles used in the course were in English, the assignments and lectures were in Dutch.

In courses followed before Biomedical Research, the undergraduate students only incidentally read a research article. At the beginning of their bachelor program they were provided with an instruction manual (Ossevoort & Voskamp, 2008) containing information about how to find literature and the use of references in scientific reports. This guide did not contain information about how to read research articles.

After Biomedical Research, students regularly read research articles for the courses they follow during the second and third years of their bachelor program. At the end of the bachelor program, students write a bachelor thesis of maximal 20 pages in four weeks (Opleidingscommissie Biologie, 2005). This thesis follows the format of a literature review article. On average, a student's bachelor thesis contains approximately 30 references (M. Ossevoort, personal communication, June 22, 2012). So, at the end of the bachelor program students have to be able to read and abstract 30 research articles in four weeks.

In 2008 we became involved with the course. The course had a clear focus on reading research articles (which were discussed during the tutor group meetings), so it seemed a good starting point for our research. Our research centered upon the tutor group meetings; no data collection took place during lectures and lab work. In 2008, we conducted our preliminary investigation (described in Chapter 2) by adapting the assignments – originally developed by the two faculty members – which students made for the tutor group meetings. The next year, in 2009, we implemented our module by further adapting the assignments and the instructions for the tutors (described in Chapters 4, 5, and 6). In 2010, we slightly adapted the assignments to study students' critical evaluation skills (described in Chapter 7).

Chapter 2
Students' First Experiences with Reading Primary Literature[1]

2.1 Introduction

This chapter describes a preliminary investigation in which we have characterized undergraduate students' difficulties with reading research articles. In this investigation, we focused on reading behavior and identifying authors' conclusion and supports. There is relatively little known about this subject. For example, the studies listed in Section 1.6 (containing descriptions of courses that aim to introduce undergraduate students to research articles) lack in-depth information about students' reading abilities. However, several studies describe the reading behaviors of scientists, graduate students, and high school students.

Bazerman (1988) has probably published the most extensive study on the reading behaviors of scientists. He interviewed seven physicists about their reading behaviors and observed four of them while reading research articles. Bazerman found that the physicists read articles selectively (not reading all parts of an article). Often they skip whole parts of the text, only reading the newsworthy parts. What was considered the news depended on the scientist's interests and purposes. For instance, if the reader was very familiar with the topic of the article, he tended to skip most of the text and only stopped at new techniques, tricks, or equations. Furthermore, the physicists read articles non-sequentially (not reading the parts of an article in order). They often read backwards and jumped back and forth (when, for example, a certain section raised questions about an earlier one). A similar study by Berkenkotter and Huckin (1995), who interviewed seven researchers (four physicists and three biologists) about their reading behaviors, confirmed these results.

[1] This chapter is based on van Lacum, E., Ossevoort, M., Buikema, H., & Goedhart, M. (2012). First Experiences with Reading Primary Literature by Undergraduate Life Science Students. *International Journal of Science Education*, *34*(12), 1795–1821.

According to Bazerman (1988), the reading behaviors of scientists are very much dependent on what he calls the "purpose-laden schema." This is a personal map of the field, which "provides the framework against which the reader comes to understand an article. The reader will process information that has significance for the existing schema and will view that information from the perspective of the schema" (p. 243). So, the purpose-laden schema will influence how readers approach a research article. Bazerman's results correspond with a study by Charney (1993), whose research suggests 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, Charney found, graduate students tended to do no more than understand the text and integrate it with their prior knowledge.

Brill, Falk, and Yarden (2004) describe how novice readers (in this case high school biology students) try to comprehend an adapted research article.[2] They observed that students encountered a number of comprehension difficulties, probably due to a lack of schemas (cognitive structures that give meaning to the information in the text) and automation (unconsciously applying the appropriate schema to the situation). For example, one of their students was not able to create a coherent understanding of the text when reading the Abstract. This is in contrast to expert readers, the authors surmised: "Expert readers can apply preformed schemas to construct the main outline of the research that is about to be read from the abstract, although they have never read about it before" (p. 508).

We can probably expect some notable differences between students and experts regarding the reading of research articles:

1. Because of their experience in reading and writing research articles, experts have a better grasp of a research article's rhetorical structure. Experts know how authors' conclusions are interwoven in the Discussion section of an article. They know which semantic structures authors use to connect supports with conclusions. Students do not possess this kind of knowledge, since they are less familiar with reading and writing research articles.
2. Experts have a better understanding of the conceptual/epistemological nature of rhetorical moves such as conclusions and supports (see be-

[2] The article was translated into Hebrew, essential information was added to the Introduction section, the Method section was simplified, the Results section was shortened, and a section was added about the contribution of the study to the understanding of processes in humans.

low). Experts know better which conclusions and which supports are valid or not.

3. Experts have more prior knowledge than students. Experts usually are familiar with the technical terms mentioned in research articles and understand immediately why and how certain experiments have been done and how to interpret the presented data.
4. Experts read with a different goal than students, because they function in a different community (see Section 1.7). Students' reading is assignment-driven, while experts read because they want to incorporate the information into their research.

In this chapter we study the reading behaviors of undergraduate students and experts from the field of life sciences. We limit ourselves to students' ability to identify conclusions (or claims/knowledge claims) and the supports for these conclusions (or evidence). We consider these as central elements in a research article.

To identify conclusions and supports, we made use of the work of Dudley-Evans (1994), who identified nine rhetorical moves in Discussion sections of research articles. Dudley-Evans' nine moves are:

1. Information move: background information about theory, research aim, methodology, or previous research.
2. Statement of result: presents a numerical value or refers to a graph or table.
3. Finding: observation arising from research; contains no reference to a graph or table.
4. (Un)expected outcome: comment on an expected or unexpected/surprising result.
5. Reference to previous research: used to compare results or as support for claim.
6. Explanation: reasons for an unexpected result.
7. Claim: a generalization arising from the results.
8. Limitation: caveats about the findings, methodology, or claims.
9. Recommendation: suggestions for future lines of research or methodology.

We define a conclusion as Move 7: a generalization based on the results presented in the article. Moves 2, 3, and 5 form the supports. Supports may be results from authors' own research (Moves 2 and 3) or results from other studies (Move 5). Moves 2 and 3 represent supports from different epistemological levels: Move 2 presents a numerical value or refers to a graph or table while Move 3 contains an observation arising from research. It has to be noted that Moves 2 and 3 are not only located in the Discussion section but also in the Results section.

2.2 Research Questions

In this chapter we describe to what extent undergraduate students are able to identify conclusions and supports in experimental research articles, which features they use for identification of these moves, and what their reading behaviors are. In this way, we could get a clearer picture of the undergraduate students' abilities regarding reading primary literature. We also compared students' performances with those of expert readers. The performances of these expert readers indicated to us the level that students ultimately should reach in their academic career.

Our research questions were:

1. Which types of statements in a research article do undergraduate life science students identify as conclusions and supports and do these statements correspond with the statements found by expert readers?
2. Which textual features do undergraduate life science students use to identify conclusions and supports and do these features correspond with experts?
3. What kinds of reading behaviors do undergraduate life science students exhibit when they read a research article and identify conclusions and supports?

With the answers to these research questions, we want to develop a new teaching strategy that aims to improve students' ability to read primary literature.

2.3 Method

Educational Setting and Participants

The data collection took place during the 2008 edition of Biomedical Research. The total number of participants who entered the course was 138. For the tutor group meetings, all students were randomly divided into 12 groups. The tutors were senior medicine students and life science students. For our data collection, we randomly chose two tutor groups, consisting of 9 and 11 students. Eight students were male and 12 students were female. They were approximately 18–20 years old. All students were novice readers of primary literature and their native language was Dutch. At the beginning of the course, these students (minus one) filled out a questionnaire about their reading experiences regarding primary literature. The majority of the students (13 out of 19) said they had read 4–6 research articles in previous courses. Three students had read less than four articles and three students had read more (up to 7–12 articles).

The two experts (called Expert A and Expert B) were a professor and an assistant professor who were both lecturers of the course. They frequently write research articles and are experienced readers of primary literature in the area of cardiovascular research.

Intervention

Assignments. In this chapter we present the written answers to the three assignments of the abovementioned 20 students. These assignments were given in Weeks 3, 4, and 5 of the course as homework and consisted of individually reading a research article and answering questions about the conclusions and supports in the article. In each week a different research article was used. The concepts presented in the articles had been discussed during lectures beforehand. The assignments, including the research articles, were handed out and explained by the tutor at the end of each tutor meeting. Students did not receive explicit instructions about the nature of conclusions and supports. Approximately six days later, the students handed in their answers via e-mail. In the following tutor group meeting, the answers of the students on the assignments were discussed. The assignments were:

- Assignment 1 (Article 1): Write down verbatim the sentence or sentences that are according to you conclusions and point out which conclusion or conclusions is/ are the most important.
- Assignment 2 (Article 2): Write down verbatim the sentence or sentences that are according to you conclusions and point out which conclusion or conclusions is/ are the most important. Furthermore, write down verbatim the support(s) of each conclusion.
- Assignment 3 (Article 3): Write down the sentence or sentences that are according to you conclusions and point out which conclusion or conclusions is/are the most important. Furthermore, write down the support(s) of each conclusion. You are allowed to paraphrase the article's sentences.

Research articles. The three research articles used subsequently for Assignments 1, 2, and 3 were:

1. Miranda et al. (2007), in which the authors established if the electrocardiogram (ECG, a recording of the electrical activity of the heart) and/ or echocardiogram (ECHO, a moving ultrasound picture of the heart) are able to adequately predict the size of a myocardial infarction in rats.

2. Kleiger, Miller, Bigger, and Moss (1987), in which the authors test the hypothesis that the degree of heart rate variability can predict long-term survival after a (acute) myocardial infarction.
3. Prunier et al. (2007), in which the authors determined if treatment with erythropoietin (EPO) has a positive influence on rats who suffered a myocardial infarction.

The research articles were selected by the course instructors and were not adapted or translated. All three articles had the conventional Introduction, Method, Results, and Discussion structure. Article 1 contained approximately 4,000 words, one table, and two figures. Article 2 contained approximately 5,000 words, six tables, and five figures. Article 3 contained approximately 6,000 words, three tables, and six figures. Readability of the articles was measured using the Flesch Reading Ease Score (Flesch, 1948). This score puts texts on a scale between 0 (*very difficult to read*) and 100 (*very easy to read*). Article 1 had a Flesch Reading Ease Score of 62, Article 2 of 27, and Article 3 of 47. This means that Article 1 had a *standard readability* and that Articles 2 and 3 were, respectively, *very difficult* and *difficult to read*. Although the readability varied, all research articles satisfied the criteria set out by Muench (2000) for selecting suitable primary literature for novice readers: the experiments in the article could easily be visualized and the given results were unambiguous. Furthermore, the relationship between the conclusion and the data was relatively simple in all three articles.

Data Sources

Assignments. All students' written answers were collected. In addition, the two experts made Assignments 1 and 3, and handed them in via e-mail.

Student interviews. Four students (Jessica, Bill, Mary, and Leah – pseudonyms) were asked to make the assignments during a task-based/think-aloud interview (van Someren, Barnard, & Sandberg, 1994; Goldin, 2000). The interviews took place in a lecture room at the university, where the students had access to an internet-connected computer. Jessica and Bill were observed while making Assignments 1 and 3, Leah was observed while making Assignment 2, and Mary was observed while making Assignment 3. We provided them with paper versions of the research articles and the assignments.

We asked the students to articulate their thoughts as much as possible while reading the research articles and making the assignments. During this process we followed the guidelines of van Someren et al. (1994). In the semi-structured interview afterwards we asked them general questions about the article and assignment (What did you think of the article? Was it easy to read? What did you think of the assignment? What is a conclusion and how can you recognize it? Etc.). In addition, we asked them to tell us how they came to their answers. These sessions were audio-recorded and transcribed verbatim. Also, students' search history on the computer was registered. We timed how long it took the students to read the article and make the assignment.

Experts interview. The interview with the experts took place after the students and experts had handed in the written answers to their assignments. To prepare for the expert interview, we analyzed the written answers of students and experts to the assignments. Then we interviewed the two experts together and asked them to compare each other's written answers and to comment on the students' answers. In addition, we asked them some general questions about conclusions and supports in research articles and their reading behaviors. The duration of this interview session was approximately 60 minutes. The session was audio-recorded and transcribed verbatim.

Data Analysis

Research articles. We performed a move analysis of the Results and Discussion sections of the three articles by determining per sentence the type of move (based on the definitions by Dudley-Evans, 1994 – see Section 2.1). We used this analysis to classify the sentences in Tables 2.1, 2.2, and 2.3 (see the last column).

Assignments. For the written answers' analysis, we chose the sentence as unit of analysis because, in our experience, sentences (in contrast to paragraphs) seldom contain more than one rhetorical move (Holmes, 1997). For Assignments 1, 2, and 3, we determined how many sentences each student and expert identified as conclusion and in which section of the article (Abstract, Introduction, Method, Results, or Discussion) they found them. We made a list of the sentences that were most frequently mentioned as the most important ("main") conclusion by the students. Reporting verbs and transition words/phrases were underlined (Bloch, 2010). Then, we compared the student results with the expert results. One student did not hand in Assignment 1. Two students did not hand in Assignment 2.

For Assignments 2 and 3, we analyzed the supports students mentioned. We performed two different analyses. In the first analysis we grouped all students' supports together and determined, based on the definitions by Dudley-Evans (1994), for each conclusion or set of conclusions if the supports contained a (a) statement of result, (b) finding, (c) reference to previous research, or (d) another conclusion (i.e. claim). We added one category to this list. Students could also refer to non-textual elements such as (e) inscriptions.[3] For this analysis, we used the answers of 17 students; three students did not hand in Assignments 2 and/or 3. In the second analysis we determined if students' supports for three selected conclusions in each article (assigned by us or the experts as important) were correct (i.e. referred to matching data or data interpretations) and complete (i.e. referred to the most important data or data interpretations mentioned in the text). Correctness and completeness were determined by the experts' comments in the interview and our own analysis. In this analysis, we used answers from 15 (Assignment 2) and 15 (Assignment 3) students. The other students were left out because they did not mention one of the three conclusions or because of missing data.

Student and expert interviews. First, we marked fragments in the transcripts related to the identification of conclusions or supports (for example: "This is a conclusion because…"). Then, the two researchers compared and discussed the selected fragments.

The marked fragments were used to determine how the interviewees identified rhetorical moves. We surmised that rhetorical moves can possibly be identified by content-based, organizational, and lexical features. For example, conclusions can be described as answers to research questions (content-based feature), occur in the Discussion section (organizational feature), and often contain reporting verbs – e.g. suggests, found, show – and transition words/phrases – e.g. overall, so, in summary (lexical features). Supports consist of experimental results (content-based feature), occur in the Results and Discussion sections (organizational feature), and can contain reporting verbs and references to graphs or tables (lexical features).[4]

We also determined what kinds of reading behaviors students exhibited by analyzing the transcripts and observation notes. Two researchers read the transcripts and observation notes independently and marked fragments that related to reading behaviors. For further analysis, we adapted a subdivision made by Brill et al. (2004): (a) connecting prior knowledge (recognizing a technical term from a lecture), (b) using illustrations (looking at a graph to understand the experimental procedures), (c) making predictions (interpreting a graph before reading the interpretation of the

[3] Roth, Bowen, and McGinn (1999) describe inscriptions (following Latour, 1987) as "representations other than text" (p. 977). Inscriptions can be "readings from simple devices, recordings from automated devices, computer screen output, photographs, micrographs, data tables, graphs, and equations" (p. 978).

[4] These features will be discussed in more detail in Chapter 5.

author), (d) ignoring technical terms, (e) using the internet (looking up technical terms), (f) repeated reading (rereading certain parts of the article), (g) sequential reading, and (h) selective reading.

2.4 Results

Identification of Conclusions by Experts and Students

The experts independently identified the conclusions in Articles 1 and 3. The sentences students identified as main conclusion in Articles 1, 2, and 3 are shown in Tables 2.1, 2.2, and 2.3, respectively. With superscript letters we show which of these sentences the experts also identified.

Regarding Article 1, the experts both identified eight sentences in the text. Five of these sentences were the same with both experts (the experts identified in total 11 sentences), so there was considerable agreement between the experts. Both experts agreed that Sentence 1A (Table 2.1) was the most important conclusion. All the sentences identified as conclusions by our own move analysis in Table 2.1 were also seen as conclusions by one or both experts, except for Sentence 1J (Table 2.1). It has to be noted that not all experts' conclusions are included in Tables 2.1 and 2.3, because the tables only contain sentences identified by students as main conclusion.

Regarding Article 3, Expert A phrased as main conclusion (because he paraphrased his answer, we could not link his conclusions with one of the sentences in the tables): "Of the 2 doses of darbepoetin (0.75 and 1.5 mg/kg, injected once a week for 8 weeks) only the highest dose improves cardiac function and architecture in a rat model of myocardial infarction. The mechanism is unknown." In addition, he phrased three sub-conclusions:

- "The hematocrit was dose dependent increased by both doses of darbepoetin."
- "Only the highest dose increases the number of circulating EPC's."
- "Only the highest dose increases the capillary density."

Expert B identified sentence 3E (Table 2.3) from the Abstract as a main conclusion. Furthermore, he identified three conclusions in the Discussion section (which together form the integrated conclusion): sentences 3A, 3B, and 3C (Table 2.3). The sentences identified as conclusions by Expert B were also conclusions according to our own move analysis. In contrast, the sentences of Expert A are probably more close to findings than to conclusions.

Students predominantly identified sentences from the Discussion section as conclusions. In total, students identified 36, 34, and 24 sentences

Table 2.1. Sentences identified as main conclusion by students (n = 19) and experts (n = 2) in Article 1.

I	II	III	IV	V	VI
1A	D	In summary our data show that ECG and ECHO detect non-invasively MI in rats and that these methods can be used to estimate infarct size.[ab]	17	16	Conclusion
1B	D	The most significant correlations were found between echocardiographic parameters and infarct size as measured by histopathology; among these, LV diastolic and systolic volumes 7 days after MI, and M-mode SF% and EF% at 7 and 28 days post-MI.	10	8	Finding
1C	D	Our data further suggest that the 7 day interval is actually the most accurate for estimation of infarct size by echocardiography. [ab]	8	5	Conclusion
1D	D	These data suggest that the best time point to estimate infarct size by M-mode CHO is at 7 days post-surgery. [ab]	17	3	Conclusion
1E	A	In summary we show that conventional ECG and ECHO methods can be used to estimate infarct size in rats. [b]	3	3	Conclusion
1F	D	Overall, the conventional ECG is an excellent marker for MI but a poor predictor of infarct size. [ab]	15	2	Conclusion
1G	D	By ECHO we found a correlation between infarct size and LV dilatation only at 7 days post MI.	9	2	Finding
1H	D	A negative correlation between infarct size and ejection fraction and shortening fraction by M-mode was found at 7 and 28 days post MI (Figs. 2-C and 2-D).	9	2	Statement of result
1I	D	Thus, in the examinations at 28 days post-infarct, a positive correlation between E/A ratio and infarct size was found, suggesting that animals with a larger infarct have greater diastolic dysfunction. [a]	13	1	Finding/ Conclusion
1J	D	These findings imply that besides systolic dysfunction, the infarcted group also has important diastolic dysfunction.	13	1	Conclusion
1K	R	This implies that M-mode echocardiography can be used as an estimate of infarct size in this time interval. [b]	6	1	Conclusion
1L	D	The echocardiogram is a useful tool in the evaluation of cardiac function and alterations in the morphology of the heart in a noninvasive form.	6	1	Information move
1M	R	Importantly, on day 7 post-infarction, there was a negative correlation between infarct size, ejection and shortening fraction determined by M-mode as shown in Figures 2-C, 2-D.	3	1	Statement of result
1N	R	Negative correlation was also found between infarct size, ejection and shortening fraction on day 28 post-surgery $R = -0.63$, $p < 0.05$, $N = 11$; $R = -0.65$, $p < 0.05$, $N = 11$; respectively) and [t]here is correlation between E/A ratio and infarct size at 28 days post infarction ($R = 0.71$, $N = 10$, $p = 0.02$).	3	1	Statement of result

(Continued)

Table 2.1. Continued.

I	II	III	IV	V	VI
1O	R	Our data suggest that the 7-day interval is actually the most accurate for estimation of infarct size by ECHO. [b]	2	1	Conclusion

Note. Column I: Sentence code. Column II: Location of the sentence (A=Abstract, R=Results, D= Discussion). Column III: Sentence identified as main conclusion. Column IV: Number of students who mentioned this sentence as conclusion. Column V: Number of students who mentioned this sentence as main conclusion. Column VI: Rhetorical move of the sentence according to authors, following Dudley-Evans (1994). Reporting verbs and transition words/phrases are underlined.
[a]Sentences identified as conclusion by Expert A.
[b]Sentences identified as conclusion by Expert B.

as conclusions in Articles 1, 2, and 3, but only the sentences mentioned as main conclusions are shown in the tables. Not all identified sentences are actual conclusions; students also identified findings and statements of result as conclusions. This was especially the case for Article 2. The average number of sentences identified as conclusions was 10.9 (Article 1), 7.1 (Article 2), and 6.2 (Article 3). So, as the course progressed, students identified on average less sentences as conclusions (Figure 2.1). With 16 of the 20 students we saw a reduction in the number of sentences identified as conclusion (comparing Assignment 1 with Assignment 3) – even though Article 3 has approximately 2,000 more words than Article 1. It has to be noted that the experts showed a similar decrease in identified conclusions.

In Tables 2.1–2.3 we underlined the reporting verbs and transition words/phrases that could signal a conclusion. It is interesting to observe that Table 2.2, in stark contrast to Tables 2.1 and 2.3, contains only two sentences with a reporting verb or transition phrase.

We observed that students identified on average more sentences as conclusions than the experts did. The students identified on average 10.9 sentences in Article 1 (with very varied scores – 4 students identified 5 or less sentences, five students identified 6–9 sentences, and 10 students identified 10 sentences or more), while the experts identified only eight each. The same difference, although less pronounced, was seen in Article 3. Nevertheless, as the course progressed, students identified on average fewer sentences as (main) conclusions. It is interesting to observe that the sentences the students identified in Articles 1 and 3 almost always contained a reporting verb or transition phrase. This is in contrast to the sentences identified in Article 2. It also seems that there was (compared with Articles 1 and 3) less consensus between the students about the main conclusion of Article 2: there was not a single sentence that was identified as main conclusion by a majority of the students. However, there was considerable agreement between students and experts regarding the main conclusion of Articles 1 and 3. With the exception of Article 3, students identified a wide range of

Table 2.2. Sentences identified as main conclusion by students (n = 18) in Article 2.

I	II	III	IV	V	VI
2A	R	Decreased HR variability increases the risk of death irrespective of average HR (Fig. 2), variables reflecting left ventricular function (Fig. 3), those measuring ventricular ectopic activity (Fig. 4 and 5), clinical or demographic variables (Table V) or drug treatment (Table VI).	12	7	Statement of result
2B	D	The present study suggests that patients with decreased HR variability have decreased vagal tone or increased sympathetic tone and may have higher risk of ventricular fibrillation.	15	5	Conclusion
2C	D	Furthermore, HR variability has a significant and strong association with mortality during follow-up even after adjusting statistically for ventricular arrhythmias detected in the same Holter recording.	13	5	Finding
2D	D	Thus, the strong univariate association with mortality (the strongest of any Holter variable) along with the ease and low cost of measuring HR variability make it an important advance in postinfarction risk stratification.	12	4	Conclusion
2E	D	Using HR variability together with information about ventricular arrhythias will improve substantially the prediction of outcome.	9	3	Conclusion
2F	D	These speculations have therapeutic implications: agents that blunt sympathetic influence or agents that promote vagal influence may have therapeutic value after infarction, particularly in patients with low HR variability.	4	2	Conclusion
2G	A	HR variability remained a significant predictor of mortality after adjusting for clinical, demographic, other Holter features and ejection fraction.	2	2	Finding
2H	D	In the present study, HR variability computed from all of the RR intervals in a 24-hour continuous electrocardiographic recording made 11 ± 3 days after infarction was significantly and strongly associated with subsequent mortality.	9	1	Finding/ Conclusion
2I	D	Beta-blocking drugs reduce the risk of mortality after myocardial infarction, in part because of their antiarrhythmic and antifibrillatory actions.[33, 34]	2	1	Reference to previous research
2J	A	HR variability is a predictor of long-term survival after acute myocardial infarction.	1	1	Conclusion
2K	A	Of all Holter variables measured, HR variability had the strongest univariate correlation with mortality.	1	1	Finding
2L	R	The association of HR variability with survival is statistically significant when evaluated with the proportional hazards model, using HR variability as measured (...).	1	1	Finding

Note. Column I: Sentence code. Column II: Location of the sentence (A=Abstract, R=Results, D= Discussion). Column III: Sentence identified as main conclusion. Column IV: Number of students who mentioned this sentence as conclusion. Column V: Number of students who mentioned this sentence as main conclusion. Column VI: Rhetorical move of the sentence according to authors, following Dudley-Evans (1994). Reporting verbs and transition words/phrases are underlined.

Table 2.3. Sentences identified as main conclusion by students (n = 20) and expert (n = 1) in Article 3.

I	II	III	IV	V	VI
3A	D	These results show that chronic EPO treatment beginning 7 days after MI reperfusion in rats attenuates cardiac remodeling and improves cardiac function. [b]	18	15	Conclusion
3B	D	These effects occurred only with an EPO dose that induced EPC mobilization in blood and increased capillary density in the MI border zone. [b]	12	8	Conclusion
3C	D	The improvement in cardiac contractility was clearly related to EPC mobilization. [b]	8	4	Conclusion
3D	D	To our knowledge, this is the first study suggesting beneficial effects of chronic EPO therapy at standard doses after I/R.	4	2	Conclusion
3E	A	We found that chronic EPO treatment reduces MI size and improves cardiac function only at a dose that induces EPC mobilization in blood and that increases capillary density in the infarct border zone.	2	2	Conclusion
3F	R	No significant difference was observed between untreated MI rats and infarcted rats receiving the lower dose of EPO.	1	1	Finding
3G	R	In contrast, the higher dose of EPO prevented anterior wall thinning and LV dilatation and preserved LV systolic function (Table 1).	1	1	Statement of result

Note. Column I: Sentence code. Column II: Location of the sentence (A=Abstract, R=Results, D= Discussion). Column III: Sentence identified as main conclusion. Column IV: Number of students who mentioned this sentence as conclusion. Column V: Number of students who mentioned this sentence as main conclusion. Column VI: Rhetorical move of the sentence according to authors, following Dudley-Evans (1994). Reporting verbs and transition words/phrases are underlined.
[b]Sentences identified as conclusion by Expert B.

sentences that were not seen as conclusions by the experts or by our own move analysis. Instead, the students identified a great number of findings and statements of results as conclusions. Furthermore, it should be noted that not all conclusions found by our move analysis were recognized as such by our experts.

Regarding Article 3, it is interesting to observe that 10 of the 20 students combined two conclusions: the one that described the effects of the treatment (attenuation of cardiac modeling and an improvement of cardiac function) and the one that described a possible mechanism (EPC mobilization and increasing capillary density) for these effects. In contrast, the experts did not cluster these conclusions: they made a clear distinction between these two.

Features Used by Experts and Students for Identifying Conclusions

We asked the experts how they identified conclusions. As Expert A stated in the interview, a conclusion is the answer to the research question:

> And one has data to justify it. But data are only data. They only have meaning in a context. And that is the conclusion. (…) In the research question you have the context taken into account and you have tried to manipulate it so that you can use the data to reach a certain conclusion.

According to Expert B, the conclusion is the summarized answer to the research questions of the authors, justified with data.

Expert A stated that a conclusion should really stay close to the data. Additionally, there are conclusions that place the research into perspective ("like…what does this mean for humankind, for the climate problem, for the expansion of the universe"). Expert B made another distinction: there is a main conclusion (which is rather broad) and there are sub-conclusions that point to the main conclusion.

Expert A added to this remark that there are two types of sub-conclusions. There are sub-conclusions that "specify the broad, big thing" and there are sub-conclusions that consist of information that can be useful to other researchers. For example, "method A is superior to method B, C, and D because of…"

When they determined the importance of conclusions, the experts used three criteria: (1) does it answer the research question, (2) how "big" is the conclusion, and (3) how much evidence is given? Furthermore, the experts used domain knowledge while they determined the importance of conclusions. Take, for example, the following conclusion which was identified

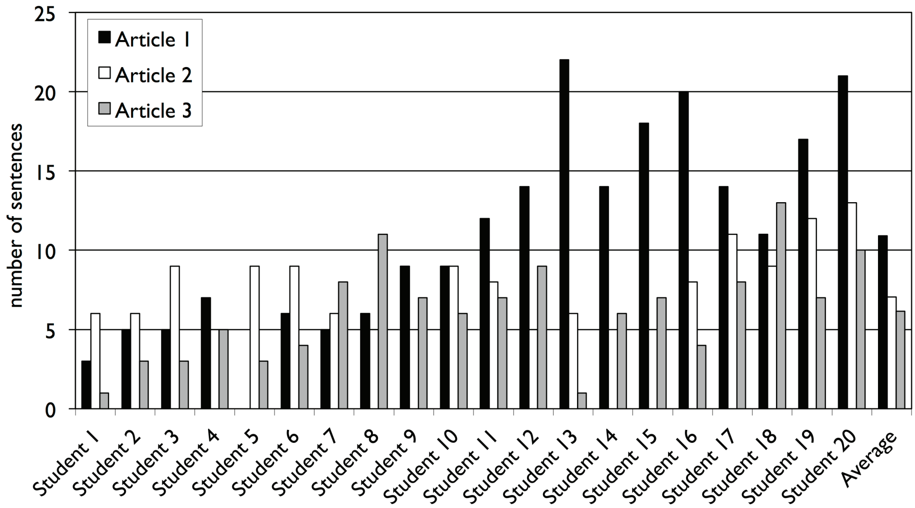


Figure 2.1. The number of sentences identified as conclusions by 20 students in Article 1 (black bar), Article 2 (white bar), and Article 3 (grey bar). Bill is student 1, Mary is student 7, Leah is student 10, and Jessica is student 18.

by Expert A: "It is interesting to observe that these values were already modified within 24 hours of the infarction and the ECHO was sensitive enough to detect these alterations precociously." According to him this was a reasonably important conclusion, because the researchers "themselves, I think, had never expected that the first time point was suitable."

When the experts examined the students' answers (Table 3.1), they noted that some of the students' sentences were not conclusions but results. An example is Sentence 1B (Table 2.1). As Expert A said: "This is in my view not a conclusion, because it is a description of the data. (…) This is a nice sentence which describes relations between data, but it does not give an interpretation." Expert B added: "It describes something and it is maybe important data (…) on the basis of this you (…) have to make a conclusion."

The students also identified Sentence 1J (Table 2.1) as conclusion. According to Expert A, this was not a conclusion because the sentence contained no new information. "This is intrinsic to the model: you always see this. (…) This is known for years."

We also asked the four interviewed students how they identified conclusions. According to Bill and Jessica, a conclusion is the answer to the research question. Bill remarked that the research question and the conclusion are sometimes formulated in the same way: "The question was: can an ECG do this. And if you read: this ECG can do or can't do this… then… so it is formulated a little bit the same as the question. Then... I think: hey, that's a conclusion." Bill, like Mary, also used lexical features (such as "we conclude… ") to identify conclusions. Nevertheless, Mary found it difficult to distinguish conclusions from other statements.

Interestingly, Jessica also said that a conclusion could be a sort of summary. She stated: "Here they really state the conclusion. At this point they summarize it shortly." Leah also thought that conclusions could be a summary: "They summarize a little bit the previous text and they tell what you can do with it (…) and it is an important advance [in the research field] (…) so it seems a conclusion to me." Later on, Leah stated that a conclusion could also be an explanation: "…and then they explain why, so I think this is also a conclusion."

How did the students decide how important each conclusion was? Jessica reasoned that the most important conclusion is the one that answers the most important research question. And, according to Jessica, the most important research question is the one that is mentioned first by the authors. As Jessica put it, when asked why a certain conclusion was the most important one: "Because… that one… that one is the answer to the first research question (…) that is the first thing they want to know." Leah used another method. According to her, the most important conclusions were the ones that summarize the research and tell what you can do with it. The conclusions that were "directly deduced from the results" were less important.

We observed a number of differences between experts and students on the identification of conclusions. The experts used rather specific content knowledge when they identified conclusions and rated their importance. For example, they saw a certain sentence not as conclusion, because the information was not new. This is an example of a content-based feature. Students also used lexical features (such as reporting verbs) and organizational features (such as the place of the sentence in the paragraph). Furthermore, the students formulated two different definitions of a conclusion: it answers the research question and it summarizes the results. The experts defined the conclusion in a slightly different way: it is a summarized answer to the research question(s).

Identification of Supports by Experts and Students

Experts' written answers regarding the identification of supports were rather different from students' answers. For example, the conclusions for which Expert A gave supports were phrased by himself. This made student–expert comparisons difficult. So, for the analysis of the students' answers we used experts' comments in the interview and our own content knowledge. We observed that the experts preferred to use data sources as supports (e.g. they only wrote "CD31 data"). Both experts did not mention references to previous research as supports.

We analyzed the written answers to Assignments 2 and 3 to elucidate which types of supports students identified. We grouped 17 students' supports and determined how many times an inscription, statement of result, finding, reference to previous research, and another conclusion were mentioned (Table 2.4). In total, the 17 students gave supports for 109 (Assignment 2) and 74 (Assignment 3) text fragments which were identified as conclusions. These fragments were not necessarily unique and often contained multiple sentences; students tended to give supports for aggregations of conclusions. Students most often mentioned inscriptions as supports. References to previous research were also frequently mentioned. Conclusions were least frequently mentioned as supports (Table 2.4).

We also analyzed the completeness and correctness of supports for Assignments 2 (Table 2.5) and 3 (Table 2.6). To this end, we classified the supports on the basis of the data sources used in the articles (and not according to Dudley-Evans, 1994). This analysis was done for a selection of

Table 2.4. Types of supports identified by students (n = 17).

	Assignment 2	Assignment 3
Inscription	49%	39%
Statement of result	18%	27%
Finding	15%	12%
Conclusion	5%	9%
Reference to previous research	23%	26%

Note. In total, the students gave supports for 109 (Assignment 2) and 74 (Assignment 3) text fragments that were identified as conclusions. The percentages denote in how many instances students mentioned an inscription, statement of result, and so forth, as support for their text fragments.

Table 2.5. The supports of three selected conclusions in Article 2 as given by 14 students.

Sentences identified as conclusion	Number of students	Supports						
		Other conclusion	Mortality	Correlation with other variables	Univariate/ independent relations	Effects of drugs	HRV & runs/ couplets	Reference
2H	8	0	5	***1***	2	***1***	0	3
2D	10	2	1	***1***	7	***0***	0	1
2E	7	1	0	***0***	1	***0***	5	1

Note. Sentences can be found in Table 2. Bold and italic values indicate that the support is incorrect according to the experts.

Table 2.6. The supports of three selected conclusions in Article 3 as given by 15 students.

Sentences identified as conclusion	Number of students	Supports							
		Other conclusion	Heart function	Infarct size	Capilary density	CD31	Hematocrite	Apoptosis	Reference
3A	8	2	2	***3***	2	***1***	1	***0***	1
3B	2	0	1	***1***	2	0	0	***0***	0
3C	2	0	0	***0***	1	1	0	***0***	0
3A+3B	3	1	2	***0***	1	2	0	***1***	1
3A+3B +3C	4	1	2	***2***	2	2	0	***0***	1

Note. Sentences can be found in Table 3. Bold and italic values indicate that the support is incorrect according to the experts.

three conclusions in each article. It is noteworthy to observe that the students often did not mention the most important supports for a conclusion. For example, in Assignment 3, only two of the eight students mentioned the (rather important) heart function data as support for Sentence 3A. However, most students' supports were correct; i.e. they referred to the right data sources. Of all the 33 supports mentioned by students in Table 2.5, only 3 were incorrect. The same applies for Table 2.6: of all the 36 supports, only 8 were incorrect.

Features Used by Experts and Students for Identifying Supports

What features do experts use to identify supports in a research article? Experts saw the supports (or evidence, as they call it) as among the most important aspects of an article. As Expert A said: "The body of evidence in an article is very important and you continually look back at it. In fact, I'm interested in the conclusions of the authors, but more to direct my own thinking process." According to Expert A, it is necessary to form your own conclusions, because the conclusions of authors are not always uncontested. However, the most important evidence can be found in the Discussion section. As Expert A said about writing the Discussion section: "If you just say: well, this does it… look, go to the data… then it will be a little bit of a puzzle. So usually, I try to mention the two or three most important supports before [the conclusion]." When asked what they expected from students regarding the identification of supports, they stated that they want students to mention at least one reference to the results per conclusion. They rather see that students refer to original data than to cited articles. References to previous research are, according to the experts, weak evidence. Expert B: "You have to show it yourself."

We also interviewed our four students about the identification of supports. When Leah tried to find the supports of the conclusions of Article 2, she admitted that she found it rather difficult: "Sometimes I didn't have a clue where [the authors] got it from." She often looked for textual clues (rather than via inferring) when deciding which support belonged to a certain conclusion. Take, for example, Sentence 2H (Table 2.2). According to Leah, this sentence was justified by a table in which RR-intervals are given. When asked why, she remarked: "Because it's about the RR-interval." Leah distinguished three types of supports: results, references to previous research, and other conclusions. Results are the most important supports, according to Leah: "Because they just measured it, so you know for sure that it is true." The second most important supports are other conclusions. The least important supports are references to previous research, because "literature can be obsolete or superseded."

Jessica was better able to explain the relationship between her supports and conclusions of Article 3. She looked more at the content of the conclusion, before deciding which supports were related to it. She made a distinction between the main conclusion (which was justified by "all the results") and other conclusions (which were justified by authors' results or references to previous research).

Mary had a similar approach: she looked at each figure and table and then decided which conclusion related to it. For example: "Let's see. Figure 5 (…) I think the black bar indicates that it's higher… that there are more EPC's mobilized and that's what the conclusion also says."

In summary, it seems that students tend to stay as close to the original data as possible when they have to give supports for conclusions. As Table 2.4 shows, a table or graph is more likely seen as support than an interpretation of said inscription (e.g. statement of result, finding, or conclusion). Interestingly, the interviews with the students show that these inscriptions were not given much attention during reading (see below). On the whole, the supports given by students were correct. The supports they mentioned mostly matched with their conclusions. However, students often failed to mention important supports for a particular conclusion. Students seemed to prefer the use of inscriptions as supports, but also referred to statements of results and findings. This is in contrast to experts, who did prefer to mention only a general description of the data source (e.g. "CD31 data"). Another difference between experts and students was the use of references to previous research as supports. Experts did not attach much importance to references to previous research and did not mention these as supports, while students occasionally mentioned them. (This is not seen in Tables 2.5 and 2.6. Probably because the sentences in these tables were major conclusions for which the authors provided mainly supports derived from the experimental results instead of references to previous research.)

Reading Behaviors Exhibited by Students

Bill. When reading the article for the first time, Bill used a sequential reading style. He seldom jumped back and forth. However, he adopted a more non-sequential reading style when making the assignment. Bill skimmed his articles, so he read rather quickly. He spent 1 hour and 10 minutes on Assignment 1 and 1 hour and 30 minutes on Assignment 3 (including reading the article). Bill skipped some parts of the article (for example, the paragraph about the statistical analysis, because "for this assignment it is not important"). Bill also found the Discussion section much more important than the Results section. "I understood the conclusions, and that is (…)

what is really important for an article. (…) How they came to it… that is more… more a proof like: it is really the case."

Bill did not pay much attention to inscriptions. Instead he looked for verbal descriptions of tables and figures. As Bill said: "Here they say in words what was presented in the table. That's easy." Bill skipped sentences in which previous research was referenced. As he said: "If you read: those and those have studied this, that's not interesting." Bill tended to ignore unknown technical terms. Bill thought it would be much effort for little gain to look them up: "Now they describe their methods. That's not very interesting. Because… things like the Simpson's method… well, I have to look up who Simpson is. And this is probably more work than it is worth." However, he looked up one technical term on the internet during Assignment 1, and two technical terms during Assignment 3.

Jessica. Jessica read her texts mainly sequentially during the first reading and seldom jumped back and forth. Later on, while making the assignment, she adopted a more non-sequential reading style. She spent 1 hour and 45 minutes on Assignment 1 and 2 hours and 10 minutes on Assignment 3.

While reading, Jessica tried to develop her own explanations and predictions:

> Well, so apparently it is different when they look at the dead rat… when they look what the… yes… what the condition is of the heart. But I think what happens next is that it will get worse or something like that. That you first have an infarct and that after a certain amount of time... that they see that the situation will get even worse.

While reading the results, Jessica tried to interpret the findings of the authors. For example: "There is a reduction in the amplitude (…) so the situation gets worse." Or: "Well, it looks like their hypothesis is correct." Jessica admitted that she reads the Results section quite intensively. "Yes… actually… I spend maybe too much time on it. But I find it interesting to read."

Jessica did not pay much attention to the article's inscriptions. "Everything is in the text. I only look at graphs and tables if I'm confused about something." After reading the Discussion section of Article 1, Jessica said that reading the results had made understanding the Discussion section easier. "It's really valuable to read the results. It makes things clearer. Because… yes, otherwise… I don't think I could have understand quite as good what they were saying."

Jessica seemed to be the only one of the four students who actively compared results from referenced research with the authors' results. Jessica did not frequently look up words on the internet: “First I try to read further as much as possible to see if I can make out [the meaning of a term]. If I can’t I will look them up on the computer.” She looked up two technical terms during Assignment 1. During Assignment 3 she did not look up technical terms.

Leah. Just like Jessica and Bill, Leah read sequentially during her first reading. When making the assignment, she began to jump back and forth. Leah skimmed the article and spent 2 hours and 5 minutes on Assignment 2. She read the whole article (minus the Abstract), but did not “take up” everything (this applied especially to the Method and the Results sections). Instead, she tried to distill the big picture. Leah did not pay much attention to the inscriptions. Leah ignored sentences in which previous research was mentioned, because “you don’t need it.” She remarked: “It is only a brief report of old studies.” She also ignored technical terms and did not look them up on the internet.

Mary. Mary was the only student who read non-sequentially right from the start. She began with reading the Abstract (“to look what kind of research they actually do and what they want to achieve with it”) and then read the Discussion section. She reasoned that reading the Discussion section was the best way to start when you have to identify the conclusions. Then she read the Method section. In hindsight, she did not think that her reading strategy was sensible, because in the beginning she had great difficulty with understanding the content of the article. “If I had read it calmly from the beginning… then I would have understood it better at once.”

Mary tried to interpret the data presented in the article. Although she did not read the text of the Results section, she did study the inscriptions quite thoroughly. “First I try to interpret the figures a little bit (…) and if this fails you read the text, of course. I always read the caption (…) much will be clear on the basis of graphs.” She did this because she found the text in the Results section that described the data “very abstract.” She spent 1 hour 45 minutes on Assignment 3. She looked up five technical terms on the internet.

All four students read selectively and three of them changed their reading from sequential to non-sequential when working on the assignment. Bill and Leah exhibited rather pragmatic reading behaviors. In contrast, the reading behaviors of Jessica and Mary were more active; for instance, they made predictions. All students said they understood the articles. However, we observed some comprehension difficulties, especially regarding the technical terms in the Method section. The students dealt with these difficulties in different ways. Bill and Leah tended to ignore unknown technical terms, while Jessica and Mary tried to find out their meaning on the internet. Overall, we encountered similar behaviors as Brill et al. (2004).

2.5 Discussion

Our results show that students and experts agreed on the most important conclusions of the research articles. However, students identified a wide range of sentences that were not seen as conclusions by the experts. We observed that students used lexical, organizational, and content-based features when identifying conclusions. Experts used content-based features not used by students (e.g. information in a conclusion has to be new). This is probably the reason why some of the sentences we identified as conclusions in our own move analysis (based on Dudley-Evans, 1994) were not identified as such by the experts. For example, Sentence 1J (Table 2.1) was not seen as conclusion by the experts because it contained no new information for them.

Our findings correspond with those of Dee-Lucas and Larkin (1988), who discovered that novice and expert readers used different rules for judging the importance of certain text elements. Novice readers tend to judge the importance of sentences by their form, while expert readers judge the importance of sentences by their content. This so-called “form effect” among novices was not seen in the group of experts. Our findings also correspond with a study by Larson, Britt, and Larson (2004), who have shown that first-year students from an introductory-level psychology class have difficulty with identifying argumentative elements in essays (e.g. they identified general statements and supports as conclusions).

When students identify supports, we see that the mentioned supports are mainly correct (i.e. students referred to data or data interpretations matching the conclusions). However, students often fail to mention the most important supports. Sandoval and Millwood (2005) described similar problems when high school students gave written explanations for two problems of natural selection – using existing data sets that students explored with special computer software. They saw that students had difficulty with citing sufficient evidence for claims and linking specific inscriptions to particular claims. Kelly, Regev, and Prothero (2008) described some criteria for determining the quality of (written) scientific argumentation in students’ reports. One of these criteria is the coordination of evidence across epistemic levels: making explicit “how particular inscriptions or claims provide evidence for higher order, more generalized claims” (p. 133). Our results imply that understanding this coordination of evidence is quite challenging for students when reading a research article.

We found that the interviewed undergraduate students read their articles selectively. For instance, they skipped or skimmed the Method section, because they thought it was not relevant for the assignment. Our students (with one exception) read their articles sequentially at first. But when making the assignments, students read the article non-sequentially. Experts, who have disparate purpose-laden schema, read differently: they read selectively, just like students, but unlike students they read non-sequentially from the start (Bazerman, 1988).

The choice of articles could have influenced our results. For instance, it is possible that certain organizational features could have influenced students' answers. The main conclusions of Articles 1 and 2, for example, were placed at the beginning of the Discussion section, while the main conclusion of Article 3 was placed at the end. However, students seem to be able to find the main conclusion regardless of its location. Some content-based features could also have influenced our results. Article 2, a correlation study, contains a vague research goal ("…to gain insight into the relations of these variables and HR variability to mortality") compared to Articles 1 and 3. This could explain why there was less similarity in students' answers to Assignment 2.

The interviewed students seemed to have no major problems with understanding the concepts described in the research articles (apart from some technical terms; see also Brill et al., 2004). This is probably due to the lectures students attended before reading the articles, in which relevant concepts were presented. This ensured that students had sufficient prior knowledge. It is generally accepted that prior knowledge plays a substantial role in text comprehension. A study of Ozuru, Dempsey, and McNamara (2009) suggests that readers' level of prior knowledge is positively correlated with the ability to comprehend an expository science text.

Our students were non-native speakers of English. We surmise that this factor did not have a significant impact on text comprehension. On the whole, Dutch students are well-versed in the English language. All textbooks that students use 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 & Donin, 1997).

It would be worthwhile if we could advance students' skills to a higher level of competency in such a way that they gradually use more expert-like strategies while reading a research article. To achieve this aim, we will combine ideas from genre analysis, argumentation theory, and cognitive apprenticeship to develop a teaching strategy.

In the next chapters, we want to focus on diminishing the first and second student–expert differences mentioned in Section 2.1: (1) Differences in knowledge about rhetorical moves and (2) differences in understanding the conceptual/epistemological nature of rhetorical moves such as the conclusion and supports. The other two differences (in levels of prior knowledge and in reading goals) are more difficult to bridge. It is impossible to bring the knowledge of students on a par with experts in the short period of a course and the goals of students and experts are often quite different while reading a research article. So, in our future teaching strategy we will direct our attention to the rhetorical structure of scientific text and the conceptual/epistemological nature of rhetorical moves. An integral part of this teaching strategy is a heuristic, the Scientific Argumentation Model, that will help students with identifying rhetorical moves. This heuristic will be described in Chapter 3. In Chapter 4, we will determine if this teaching strategy improves students' reading abilities.

Chapter 3
The Scientific Argumentation Model[1]

3.1 Introduction

As stated in Chapter 1, we want to improve undergraduate students' reading abilities by increasing their understanding of the rhetorical structure of research articles. As Chapter 2 has shown, students should be given support in this process. That is why we want to provide students with a heuristic that helps them to identify rhetorical moves in research articles. As research articles are persuasive in nature (see Sections 1.4 and 1.5), we want to focus especially on rhetorical moves that play an important role in the authors' argumentation. In our view, this heuristic should adhere to the following criteria:

1. The heuristic should describe the research article's rhetorical moves that play an important role in the authors' argumentation.
2. Because novice readers of primary literature should be able to work with the heuristic, the descriptions of the rhetorical moves have to be as simple and as unambiguous as possible.
3. The heuristic should describe the relations between these rhetorical moves.
4. The relations between rhetorical moves should be represented in a visual way. This will make it easier for students to get an overview of these relations.
5. The heuristic should be generic; it should be applicable to a broad range of research articles from the field of life sciences, in all the variations that different journals exhibit.

In this chapter, we will describe how we used ideas from genre analysis and argumentation theory to develop this heuristic. The work in genre analysis

[1] This chapter was written together with Marcel Koeneman.

has provided us with rich descriptions of the rhetorical moves that occur in research articles. It should be noted that genre analysts pay little attention to the relations between rhetorical moves in research articles. These relations are often of an argumentative nature, as moves are often used as a support for other moves. For example, earlier research is used to support a research question, experimental results are used to support a claim, and so forth. To describe these relations, we combined argumentation theory (discussed in Section 3.2) with the concept of rhetorical moves (discussed in Section 3.3).

Our heuristic, called the Scientific Argumentation Model (SAM), consists of two components. The first component is a description of the rhetorical structure (i.e. the arrangement of rhetorical moves in a text) of a research article. Our description of this rhetorical structure is not exhaustive. To make the model manageable for undergraduate students, we constructed a set of seven rhetorical moves. These moves are in some cases aggregations of moves described by other authors and in other cases more specified versions of existing moves. The second component of our heuristic is a scheme that depicts the relations between the seven rhetorical moves. We will use SAM to improve students' understanding of the rhetorical structure of research articles.

In this chapter, we will begin with describing the argumentative aspects of research articles. Then, we will describe several existing argumentation frameworks and discuss which features of these frameworks we will use for our own model (Section 3.2). Next, we will present an overview of rhetorical moves in research articles, based on the work of genre analysts (Section 3.3). Finally, we will describe SAM and its development (Section 3.4).

3.2 Argumentation Frameworks for the Analysis of Research Articles

Andrews (2010) defines argumentation as "the process of developing arguments, the exchange of views, the seeking and provision of good evidence to support claims and propositions – the *choreography* of argument" (p. 39). There is a difference between argumentation and persuasion:

> Most argument would hope to be persuasive, but not all persuasion is argumentative. Although Aristotle characterized rhetoric as 'the art of persuasion' in a general sense, persuasion describes the effect or effectiveness of an approach from one person or a number of persons to another/others. Argument and argumentation, on the other hand, describe the interventions and dialogues that make up human transactions. (Andrews, 2010, p. 39)

As stated in Chapter 1, Gould (1993) distinguished two aspects of scientific persuasion: (1) form and language and (2) logic and empirical content. Following Andrews (2010), form and language in a research article may be persuasive but are not necessarily argumentative. For example, it is perfectly possible to write in a persuasive way without using evidence to support claims. Logic and empirical content are argumentative (and also persuasive if the logic and empirical content are to be accepted by the reader), because they involve the use of evidence to support claims. As Deanna Kuhn (1993) points out: "Not just the theories but even the so-called 'facts' of science become argumentive constructions that must be entered into the arena of public debate" (p. 321). Kuhn compares science with advocacy. Scientific questions are often posed by the proposition of a number of competing theories. Answering a scientific question happens in the form of a debate in which individuals are advocating one of these theories:

> To participate, an individual scientist must analyze the evidence and its bearing on the different theories as a means of argument to the scientific community in support of his or her view. Equally important, this analyzing and weighing process of argument is, in interiorized form, almost certainly an important part of what goes on in the private thought of the individual scientist. Scientists are well aware that explicitly justified arguments are needed to convince the scientific community, and they become accustomed to thinking in such terms. (Kuhn, 1993, p. 321–322)

The four sections in a research article (Introduction, Method, Results, and Discussion) have different persuasive aims. Hunston (1994) used a sample article to characterize these aims. Each of the aims involves a different argument. According to Hunston, the Introduction section persuades the reader that the research described in the article is necessary and worthwhile. The Method section persuades the reader that the research was done well. The Results section persuades the reader that the statistical methods used were useful and informative. The Discussion section persuades the reader that the results make sense and correspond with other examples of research.

Several argumentation frameworks can be used to analyze arguments in scientific texts. In many cases, these frameworks were used to assess students' arguments (for an overview, see Sampson & Clark, 2008). We want to limit ourselves to three frameworks that have been used to analyze arguments in research articles. At the end of this section, we will discuss which features of these frameworks we will use for our own model.

Toulmin (1958)

Stephen Toulmin (1958) devised a "logically candid layout of arguments" (p. 95), which could be used in a variety of cases. This Toulmin scheme (or model), as it became to be known, consists of the following elements: data, warrant, backing, qualifier, rebuttal, and claim (Figure 3.1). The data are the facts that form the foundation of the claim. The warrants are the propositions that are used to make the step from data to claim. These can be rules, principles, and so forth. Warrants are often supported by backings, "without which the warrants themselves would possess neither authority nor currency" (p. 103). The qualifier is "the reference to the degree of force which our data confer on our claim in virtue of our warrant" (p. 101). The rebuttal indicates the circumstances in which the warrant is not applicable. Toulmin gives the following example of a scheme: Harry was born in Bermuda (data). So, presumably (qualifier), Harry is a British subject (claim) since a man born in Bermuda will generally be a British subject (warrant) on account of the following statutes and other legal provisions (backing) unless he has become a naturalized American (rebuttal).

According to Toulmin, Rieke, and Janik (1979), science uses regular and critical arguments. In regular arguments the trustworthiness of warrants is taken for granted. This kind of argument consists of:

- Grounds (called data in Toulmin, 1958): Factual reports. For example the results of a measurement.
- Warrants: The justification of the step from grounds to claim. For example a mathematical formula or a law of nature.
- Backing: The backing supports the warrants. For example the evidence which is accumulated over the years for a certain law of nature.
- Modals (called qualifiers in Toulmin, 1958) and rebuttals: The modal gives an indication of the strength of the argument. Rebuttals are counterarguments or exceptions.
- Claim: A more or less straightforward factual conclusion.

In critical arguments, the credentials of certain ideas are challenged. In these cases one cannot rely on warrants, because "the merits of a theory are no longer presupposed but are themselves up for criticism and reappraisal" (Toulmin et al., 1979, p. 247).

Thompson (1993) used a combination of Toulmin's scheme and genre analysis to analyze the Results sections of 16 experimental research articles written by biochemist Arthur Kornberg and 20 biochemistry articles from other scientists as control. For her analysis, she combined two Toulmin schemes (Figure 3.2). This new scheme reflects how a Results section is structured. In the first scheme the grounds consist of raw experimental data. Via warrants (data are generated by scientifically legitimate methods) and backings (methodological justifications) the grounds lead to the claim that the data are valid. This claim justifies the grounds (data as facts) of the second Toulmin scheme. Via warrants (tacit disciplinary knowledge/values) and backings (references to research showing agreement with the current study) these grounds lead to the claim (truth or knowledge claim). This claim may be accompanied by modal qualifiers like hedges (see Section 1.5).

Suppe (1998)

According to Suppe (1998), existing philosophical testing and confirmation models "ignore most of the scientific substance and seriously garble

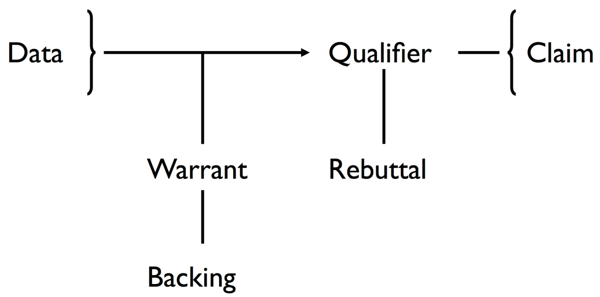


Figure 3.1. The Toulmin scheme (adapted from Toulmin, 1958).

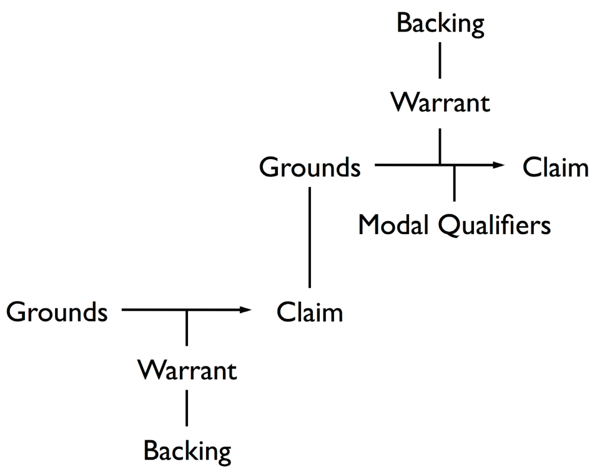


Figure 3.2. Two connected Toulmin schemes (adapted from Thompson, 1993).

Table 3.1. Partial representation of the interpretative-argument structure of a research article (adapted from Suppe, 1998). A letter with a subscripted number denotes a specific text fragment.

Hypothesis (H)	Data (D)	Preferred interpretation (I)	Doubts (Q)	Rejoinders (R)
H			Q_{18}	R_{28}
H_1, H_2	D_1		Q_1	R_1, R_2, R_3
H_3-H_5			Q_3	R_4
C_1: Am-Af	D_2	I_1	I_2	R_6
	D_3	I_2	Q_6	I_4
			Q_4	R_6
	D_2, D_3	I_4	Q_5	R_7
			Q_8	R_9, R_{10}
	D_6		Q_9	R_{11}

the logic of scientific papers" (p. 382). That is why he set out to give a complete description of the argumentative structure of a research paper. According to Suppe, a scientific paper is crafted around the following items: statement which identifies the relevant body of scientific literature and makes the case that the study will add something new, hypothesis (a specific claim), data (which serves as evidence for a claim), methods used in manipulating or interpreting data, statements which specify and motivate the preferred interpretation of data, queries or doubts about interpretations of data, and rejoinders, which impeach as many of the competing interpretations as possible. Suppe developed the model of interpretative-argument structure (Table 3.1), which consists of a table in which each row lists (if available) a hypothesis, the related data, the associated preferred interpretation of this data, doubts and queries about this interpretation, and rejoinders. Suppe also developed the model of graphic/tabular interpretative-argumentative structure, which is similar to the model of interpretative-argument structure but is centered upon the figures and tables of a research article.

Kelly, Regev, and Prothero (2008)

Kelly and Takao (2002) developed a framework in which propositions are arranged according to certain epistemic levels. Although they used their framework for the analysis and assessment of written scientific papers produced by university oceanography students, and not for published research articles, we think that these two genres are close enough related to warrant the inclusion of Kelly and Takao's framework in this section.

Kelly and Takao's framework consists of six stacked epistemic levels that are linked to each other. The first level contains the most specific propositions (such as references to data charts). Then, the propositions become progressively more general and theoretical. Finally, the claims are "so general as to be independent of the specifics of the argument in question" (p. 321).

A more recent publication (Kelly, Regev, & Prothero, 2008) describes a slightly adapted version of this framework. This adapted version has only five epistemic levels: data inscription (see Section 2.3), low inference claim, claim supported by previous claims, theoretical claim, and thesis. Across these levels one can distinguish certain lines of reasoning (Figure 3.3). Kelly et al. (2008) list several criteria that can be used to assess the strength of arguments advanced in students' papers (e.g. students need to pose a solvable research question or thesis statement, lines of reasoning need to converge in a manner that is supportive of the thesis, evidence should be coordinated across epistemic levels, et cetera).

All three frameworks may be used to depict the relations between statements in research articles. The three frameworks all use progressive steps in reasoning: they show how data/grounds (in the case of Toulmin), data inscriptions (in the case of Kelly et al.), and data (in the case of Suppe) advance into higher-level claims.

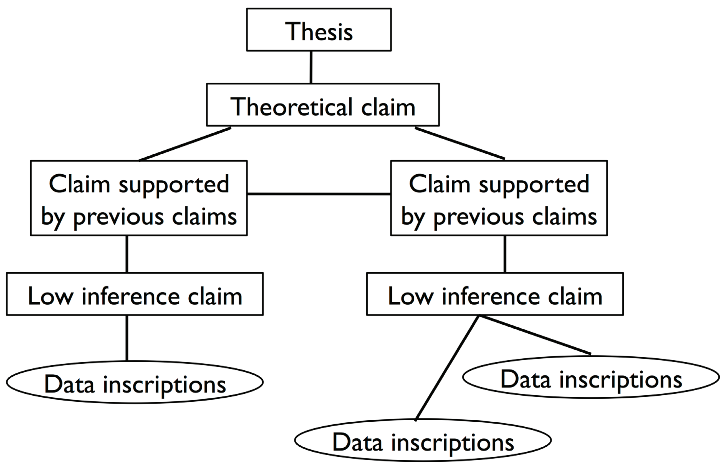


Figure 3.3. An example of a line of reasoning in a research paper (adapted from Kelly et al., 2008).

There are also several differences between the frameworks. The most obvious difference is the way of representation: the frameworks of Toulmin and Kelly et al. use lines and/or arrows to represent the relations between statements, while the framework of Suppe uses a table to group statements. Secondly, Suppe's framework is centered on hypotheses, Kelly et al.'s framework on a thesis, and Toulmin's framework on claims (although admittedly these terms are related to each other). Thirdly, only Thompson's (1993) interpretation of Toulmin's framework explicitly addresses methodological justifications. Fourthly, only the frameworks of Toulmin and Suppe mention statements – doubts and rebuttals – which could potentially weaken the hypothesis/claim. Suppe includes the option of challenging a doubt by a rejoinder. Fifthly, only Toulmin's framework addresses the (often tacit) theoretical assumptions that underlie a study.

In this chapter, we want to construct an argumentation scheme in which we can depict a selection of a research article's rhetorical moves. This scheme has to be relatively easy to work with, because it has to be adopted by undergraduate students as a heuristic. Because of this reason, none of the abovementioned frameworks are completely suitable for our purpose. The Toulmin scheme is very visual and clear-cut. However, students will probably have difficulty with identifying warrants and backings in research articles, as these two argumentative elements are often tacit (Kelly & Takao, 2002; Sampson & Clark, 2008). Authors of research articles presume that readers know and accept the underlying warrants and backings, so they will often not explicitly mention these. For example, scientists do not explicitly state that a widely used experimental procedure is a valid way of gathering data. The framework of Kelly et al. has a relatively clear layout, but the classification of statements into data inscriptions, low inference claims, claims supported by previous claims, theoretical claims, and thesis seems too detailed for our purpose: students will need a great amount of instruction to work with these concepts. However, Kelly et al.'s concept of lines of reasoning seems very promising for describing the supports in a research article. Suppe's argumentative structure is, because of its table form, relatively difficult to understand, although its doubts and rejoinders seem to be useful in describing how authors try to persuade readers.

In summary, we want to use aspects from all three frameworks for the design of our own argumentation scheme, namely:

- The Toulmin scheme's visual, and clear-cut layout.
- Suppe's distinction between doubts and rejoinders.
- Kelly et al.'s lines of reasoning.

3.3 The Rhetorical Moves of Research Articles

Hyon (1996) distinguishes three types of genre theories: (a) English for Specific Purposes (ESP) (b) North-American New Rhetoric Studies, and (c) Australian Genre Theories. ESP studies describe the formal characteristics of a genre by analyzing the frequency of certain rhetorical moves. This can be done within a single discipline (e.g. Williams, 1999), between disciplines (e.g. Peacock, 2002), or between certain time periods (e.g. Li & Ge, 2009). This is called genre analysis. In New Rhetoric studies there is a focus on the situational contexts in which these genres occur. These studies (e.g. the work of Bazerman, 1988) have a more ethnographic character. Furthermore, New Rhetoric is centered on composition studies and professional writing by native speakers, while ESP is directed towards writing by non-native speakers (Flowerdew, 2005). The Australian Genre Theories (also called the "Sydney School") use Halliday's systemic functional linguistics as basis: a framework which "reconceptualizes language as a semiotic tool intimately involved in the negotiation, construction, organization, and reconstrual of human experiences" (Fang, 2005, p. 336). Our work will be partly based on ideas from ESP/genre analysis and the New Rhetoric Studies, because these have paid detailed attention to research articles (Hyon, 1996). We want to use the genre analysts' work to provide students with criteria for identifying conclusions, supports, and other recurring rhetorical moves in research articles.

John Swales, an influential ESP-researcher and one of the founders of genre analysis, defined genre as follows:

> A genre comprises 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. This rationale shapes the schematic structure of the discourse and influences and constrains choice of content and style. (Swales, 1990, p. 58)

The goal of genre analysis is "to describe the communicative purposes of a text by categorizing the various discourse units within the text according to their communicative purposes or rhetorical moves" (Connor, Upton, & Kanoksilapatham, 2007, p. 23). As explained in Chapter 1, 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, Upton, & Kanoksilapatham, 2007, p. 23).[2]

[2] This is quite similar to Roland Barthes' (1977) description of discourse: "A discourse is a long 'sentence' (the units of which are not necessarily sentences), just as a sentence, allowing for certain specifications, is a short 'discourse'" (p. 83).

We will describe the different rhetorical moves that have been identified in the Introduction, Method, Results, and Discussion sections of research articles. These four sections form together the IMRD structure. Although some journals slightly deviate from the IMRD structure (e.g. by placing the Method section at the end), the four different sections are often clearly discernible. Hill, Soppelsa, and West (1982) compare the overall structure of a research article to an hourglass. In the Introduction section, the author advances from the general to the particular, and in the Discussion section the author advances from the particular to the general.

In his influential *How to Write and Publish a Scientific Paper*, Robin Day (1998) summarizes the logic of the IMRD structure as follows: “What question (problem) was studied? The answer is in the Introduction. How was the problem studied? The answer is the Methods. What were the findings? The answer is the Results. What do these findings mean? The answer is the Discussion” (p. 7). This structure is found across a wide variety of disciplines.

According to Burrough-Boenisch (1999), supposed benefits of the IMRD structure are that it allows readers to find information quickly and that it orders scientists’ thoughts. Supposed disadvantages are, according to Burrough-Boenisch, that it does not place enough emphasis on the Results section (which contains the most important information) and that it forces scientists to think and write in a paradigm that does not adequately reflect the process of scientific inquiry.

Research articles also contain references to previous research. These references are concentrated in the Introduction, Method, and Discussion sections and may serve to highlight the relevance of the study, justify the used methods, provide further support for claims, or indicate how the findings can solve a certain problem (Gilbert, 1977). References link the research article to the body of existing scientific knowledge (Amsterdamska & Leydesdorff, 1989). According to Amsterdamska and Leydesdorff (1989), research articles can be viewed as a series of arguments which together serve to introduce and justify a new knowledge claim and demonstrate that the proposed innovation performs a specific function within a particular body of knowledge (e.g. as a solution to a certain problem). References play an important role in the persuasive process. As Latour wrote (1987): “A paper that does not have references is like a child without an escort walking at night in a big city it does not know: isolated, lost, anything may happen to it” (p. 33). By using references, Latour argues, the author ensures that if a reader wants to attack the text, he or she also has to weaken the papers that are referred to. However, literature that is used as reference needs to be transformed by authors. For example, authors “might combine positive and negative modalities, strengthening for instance a paper X in order to weaken a paper Y that would otherwise oppose their claim" (Latour, 1987, p. 37).

Now follows a detailed description of the rhetorical moves identified in the Introduction, Method, Results, and Discussion sections of research articles. For this purpose, we will use several genre analysis studies (Table 3.2). In these studies, the frequency of rhetorical moves was determined in a selection of research articles (with the exception of Swales, 1990, and Dudley-Evans, 1994, whose work has a more theoretical perspective). We will also use writing manuals (Day, 1998; Penrose & Katz, 1998) and articles containing advice for novice and more experienced writers. Because these texts are prescriptive, rather than descriptive, they may not always give an accurate representation of the research article's structure (Crookes, 1986). However, as many scientists and students use these texts, we believe that they contain some valuable insights about the general structure of research articles.

Introduction Section

As Day (1998) notes, the Introduction section should provide background knowledge and the rationale and the purpose of the study. It is important that the purpose of the study (sometimes formulated as a research question or hypothesis) is specific. Day suggests that the Introduction section should contain the main results and conclusion – "Do not keep the reader in suspense" (p. 34) – but this feature is not mentioned in other handbooks on scientific writing. For example, Penrose and Katz (1998) state: "This

Table 3.2. Genre analysis studies that have analyzed rhetorical moves in research articles.

Study	Discipline	Sections analyzed	Size of corpus
Swales (1990)	Not specified	Introduction	Not applicable
Thompson (1993)	Biochemistry	Results	$n = 36$
Dudley-Evans (1994)	Not specified	Discussion	Not applicable
Nwogu (1997)	Biomedical Sciences	All sections	$n = 15$
Williams (1999)	Biomedical Sciences	Results	$n = 8$
Peacock (2002)	Seven different ones	Discussion	$n = 252$
Kanoksilapatham (2005)	Biochemistry	All sections	$n = 60$

[the Introduction] is where you explain your research objectives, argue that the research is important, and place your study in the context of previous research" (p. 40). According to Penrose and Katz, an effective Introduction section has as rhetorical goal to convince readers that the topic is important and that the research described in the article will advance the field's knowledge.

Swales (1990) has looked in detail at the Introduction section of research articles. He identified three rhetorical moves: the first move is establishing a territory (Move 1), in which the significance of the research field is explained. This can be done by so-called steps (or sub-moves): claiming centrality, making topic generalization(s) and/or reviewing items of previous research. Then the second move follows: establishing a niche (Move 2), in which the authors explain the reasons behind their particular research. These reasons can consist of counter-claiming, indicating a gap, question raising or continuing a tradition. Occupying the niche (Move 3) is the final move, in which the authors' research is introduced. This move can consist of specifying the purpose of the study, followed by an announcement of the principal findings and the structure of the article. Nwogu (1997) and Kanoksilapatham (2005) use more or less the same categories, although with some alterations (Table 3.3). Analyses of Introduction sections suggest that Day's recommendation of including the conclusion in this part of the paper is increasingly common (Swales & Najjar, 1987; Berkenkotter & Huckin, 1995).

Method Section

The Method (or Materials and Method) section of a research paper contains a description of the methods and materials used to obtain the results. According to Day (1998), this has to be done in such detail so that other researchers can repeat the procedures. However, this is – in most cases – more of an ideal than a reality. As Gilbert and Mulkay (1984) note, the formal written account of an experimental procedure omits important implicit assumptions which are not necessarily shared by scientists outside the specialist community (or even outside the specific laboratory[3]): "It is often extremely difficult to specify in full the actions relevant to the production of their results" (p. 54). Maybe this is the reason why some journals are nowadays de-emphasizing the Method section (e.g. by placing it at the end): readers who want to replicate the study will generally contact the authors – instead of relying solely on the formal account (Penrose & Katz, 1998). An important additional feature of the Method section is the justification

[3] See also Crabbe, Wahlsten, and Dudek (1999), who have shown how scientific results can be laboratory-specific.

Table 3.3. Moves in the Introduction section of research articles according to Swales (1990), Nwogu (1997), and Kanoksilapatham (2005).

Swales (1990)	Nwogu (1997)	Kanoksilapatham (2005)
Move 1: Establishing a territory by Step 1: Claiming centrality and/or	Move 1: Presenting background information by (1) Reference to established knowledge in the field (2) Reference to main research problems	Move 1: Announcing the importance of the field by Step 1: Claiming the centrality of the topic
Step 2: Making topic generalization(s) and/or		Step 2: Making topic generalizations
Step 3: Reviewing items of previous research	Move 2: Reviewing related research by (1) Reference to previous research	Step 3: Reviewing previous research
Move 2: Establishing a niche by Step 1A: Counter-claiming or		
Step 1B: Indicating a gap or	(2) Reference to limitations of previous research	Move 2: Preparing for the present study by Step 1: Indicating a gap
Step 1C: Question-raising or		Step 2: Raising a question
Step 1D: Continuing a tradition		
Move 3: Occupying the niche by Step 1A: Outlining purposes or	Move 3: Presenting new research by (1) Reference to research purpose	Move 3: Introducing the present study by Step 1: Stating purpose(s)
Step 1B: Announcing present research		
	(2) Reference to main research procedure	Step 2: Describing procedures
Step 2: Announcing principal findings		Step 3: Presenting findings
Step 3: Indicating structure of research article		

of the procedures. The extent of justification depends on the nature of the procedure. Standard procedures (which everyone in the field uses) are generally not justified or explained. Procedures derived from previous studies appear only with references. Procedures are only explained and justified if they are new or adapted (Penrose & Katz, 1998).

Genre analysts have not looked in great detail at the Method section. However, Nwogu (1997) identified three moves in the Method section of medical research articles: Describing data collection procedures, describing experimental procedures and describing data analysis procedures (Table 3.4). Kanoksilapatham (2005) has identified four moves in biochemical research articles: (1) Describing materials, (2) Describing experimental

Table 3.4. Moves in the Method section of research articles according to Nwogu (1997), and Kanoksilapatham (2005).

Nwogu (1997)	Kanoksilapatham (2005)
Move 4: Describing data collection procedures by (1) Indicating source of data	
(2) Indicating data size	
(3) Indicating criteria for data collection.	
Move 5: Describing experimental procedures by (1) Identification of main research apparatus	Move 4: Describing materials by Step 1: Listing materials
	Step 2: Detailing the source of the materials
	Step 3: Providing the background of the materials
(2) Recounting experimental process	Move 5: Describing experimental procedures by Step 1: Documenting established procedures
	Step 2: Detailing procedures
	Step 3: Providing the background of the procedures
(3) Indicating criteria for success	
	Move 6: Detailing equipment (optional)
Move 6: Describing data analysis procedures by (1) Defining terminologies	Move 7: Describing statistical procedures (optional)
(2) Indicating process of data classification	
(3) Identifying analytical instrument/procedure	
(4) Indicating modification to instrument/procedure	

procedures, (3) Detailing equipment, and (4) Describing statistical procedures. Like Nwogu, Kanoksilapatham does not mention a move that justifies procedures or methodology, possibly because his research suggests that in biochemical articles this move occurs in the Results section and not in the Method section.

Results Section

Authors use the Results section to present their collected data. Data are reduced (by calculating averages, for example) and – if necessary – processed into tables and graphs (Penrose & Katz, 1998). Authors use so-called pointers to link tables or graphs with textual statements (Brett, 1994; Penrose & Katz, 1998). These pointers can be integral (*the results are shown in Figure 1*) or non-integral (*there was an increase seen (Figure 1)*). Authors not only present their data; they also comment on them, for example by providing explanations or drawing preliminary conclusions.

Genre analysts (Kanoksilapatham, 2005; Nwogu, 1997; Thompson, 1993; Williams, 1999) have identified a number of rhetorical moves in the Results section (Table 3.5). Nwogu (1997) distinguishes moves that indicate consistent (i.e. expected) observations and moves that indicate non-consistent (i.e. unexpected) observations. Kanoksilapatham (2005) distinguishes the following moves: justifying procedures or methodology, stating results, and stating comments on the results (e.g. explaining the results or making generalizations or interpretations). Williams (1999) makes a distinction between presentational and comment moves. Presentational moves contain procedural information (statements about how and why the data has been produced) and what the findings were, while comment moves are – among other things – explanation of findings, comparisons with literature, or interpretations of findings.

It is interesting to observe that Thompson (1993), just like Kanoksilapatham (2005), identifies methodological justifications in the Results section. Thompson, who studied the research articles of biochemist Arthur Kornberg, explains this as follows:

> Because the validity of experimental data rests entirely on methodology, Kornberg defends technical choices in Results sections rather than in Material and Methods sections. If we view experimentation as constructing scientific "facts" instead of only revealing them, then arguments for methods are most appropriate in the Results section, where the validity of the data comes under careful scrutiny. (p. 116)

Discussion Section

The Discussion section is probably the most variable and complex part of a research article. This is why, as Day (1998) states, the Discussion section is usually the most difficult to write. There is a direct relationship between the Discussion and the Introduction section: "Whereas the introduction introduces the research question and reviews the state of knowledge in the field that motivated the question, the discussion explains how the question has been answered (at least in part) by the new research and shows how the field's knowledge is changed with the addition of this new knowledge" (Penrose & Katz, 1998, p. 57). Day (1998) specifies six essential features of a good Discussion: (a) Principles, relationships, and generalizations shown by the results should be presented, (b) anomalous data should be pointed out, (c) results and interpretations should be compared with previous research, (d) theoretical and practical implications should be discussed, (e) each conclusion should be stated as clearly as possible, and (f) the evidence for each conclusion should be presented in summarized form. Furthermore, the Discussion section is the place where the limitations of the study

Table 3.5. Moves in the Results section of research articles according to Thompson (1993), Nwogu (1997), Williams (1999), and Kanoksilapatham (2005).

Thompson (1993)	Nwogu (1997)	Williams (1999)	Kanoksilapatham (2005)
		Procedural information	Move 8: Stating procedures by Step1: Describing aims and purposes
			Step 2: Stating research questions
			Step 3: Making hypotheses
			Step 4: Listing procedures or methodological techniques
Methodological justifications			Move 9: Justifying procedures or methodology by Step 1: Citing established knowledge of the procedure
			Step 2: Referring to previous research
		Statement of finding: (a) Comparison of two of more subjects	
		(b) Description of change over a period of time	
		(c) Relation between variables	
		(d) Descriptions of quantitative data	
	Move 7: Indicating consistent observation by (1) Highlighting overall observation (2) Indicating specific observations	Substantiation of finding: (a) Findings which support or do not contradict other findings of the study	Move 10: Stating results by Step 1: Substantiating results
		(b) Complementary details of results which support a more general result	
	(3) Accounting for observations made		
	Move 8: non-consistent observations	Non-validation of finding	Step 2: Invalidating results
		Explanation of finding	Move 11: Stating comments on the results by Step 1: Explaining the results
Interpretations of experimental data			Step 2: Making generalizations or interpretations of the results

(Continued)

Table 3.5. Continued.

Thompson (1993)	Nwogu (1997)	Williams (1999)	Kanoksilapatham (2005)
Agreement with pre-established studies		Comparison of finding with literature: (a) Findings are the same	
		(b) Findings are neither the same nor different	
Comments on discrepancies		(c) Findings are different	
		Evaluation of finding re hypothesis: (a) Finding is in line with the hypothesis	Step 3: Evaluating the current findings
		(b) Finding is not in line with the hypothesis	
Calls for further research		Implications of finding	
			Step 4: Stating limitations
			Step 5: Summarizing
Evaluations about the experimental data's accuracy			
Interpretative perplexities			

are discussed. These are remarks about "problems with errors, methods, and validity" (Ioannidis, 2007, p. 324) and may also be described as weaknesses, caveats, or shortcomings of the research.[4]

The Discussion section is the place where the main knowledge claim can be found: "the assertion for which the authors hope to be cited – and credited – in future articles" (Myers, 1992, p. 296). This main knowledge claim can also be described as the main conclusion or thesis: the claim that the whole research article supports (Booth, Colomb, & Williams, 2003). When expressing claims, authors generally use hedges.

In Table 3.6, the Discussion section's moves that are proposed by Dudley-Evans (1994), Nwogu (1997), Peacock (2002), and Kanoksilapathan (2005) are put next to each other. Dudley-Evans's (1994) nine moves occur often in recurring move cycles. For example, a statement of result or finding is regularly followed by a reference to previous research. Peacock (2002) used the subdivision of Dudley-Evans for his analysis of the Dis-

[4] Interestingly, few journals encourage authors to include limitations. As a result, a minority of journal articles contains words that denote limitations (Ioannidis, 2007).

Table 3.6. Moves in the Discussion section of research articles according to Dudley-Evans (1994), Nwogu (1997), Peacock (2002), and Kanoksilapatham (2005).

Dudley-Evans (1994)	Nwogu (1997)	Peacock (2002)	Kanoksilapatham (2005)
Move 1: Information move (background information about theory, research aim, methodology, or previous research)		Move 1: Information move	Move 12: Contextualizing the study by Step 1: Describing established knowledge
			Step 2: Presenting generalizations, claims, deductions, or research gaps
			Move 13: Consolidating results by Step 1: Restating methodology (purposes, research questions, hypotheses restated, and procedures)
	Move 9: Highlighting overall research outcome		
Move 2: Statement of result (presents a numerical value or refers to a graph or table)	Move 10: Explaining specific research outcomes by (1) Stating a specific outcome	Move 2: Finding (with or without a reference to a graph or table)	Step 2: Stating selected findings
Move 3: Finding (observation arising from research; contains no reference to a graph or table)			
	(2) Interpreting the outcome		
	(3) Indicating the significance of outcome		
Move 4: (Un)expected outcome (comment on an expected or unexpected/surprising result)		Move 3: (Un)expected outcome	Step 3: Referring to previous literature
Move 5: Reference to previous research (used to compare results or as support for claim)	(4) Contrasting present and previous outcomes	Move 4: Reference to previous research	
Move 6: Explanation (reasons for an unexpected result)		Move 5: Explanation (reasons for expected or unexpected results)	Step 4: Explaining differences in findings

(Continued)

Table 3.6. Continued.

Dudley-Evans (1994)	Nwogu (1997)	Peacock (2002)	Kanoksilapatham (2005)
Move 7: Claim (a generalization arising from the results)		Move 6: Claim (contribution to research; sometimes with recommendations for action)	Step 5: Making overt claims or generalizations
			Step 6: Exemplifying
Move 8: Limitation (caveats about the findings, methodology or claims)	(5) Indicating limitations of outcomes	Move 7: Limitation	Move 14: Stating limitations of the study by Step 1: Limitations about the findings
			Step 2: Limitations about the methodology
			Step 3: Limitations about the claims made
	Move 11: Stating research conclusions by (1) Indicating research implications		
Move 9: Recommendation (suggestions for future lines of research or methodology)	(2) Promoting further research	Move 8: Recommendation	Move 15: Suggesting further research (optional)

cussion section from 252 research articles across seven disciplines. Based on his results, he made some small alterations of Dudley-Evans's model. For example, Peacock suggests that Moves 2 and 3 should be combined, "as in both moves authors refer back to a finding, either with or without a reference to a graph or table" (p. 492). Kanoksilapatham (2005) came up with a model that is not so much dissimilar, but more detailed than the models of Peacock and Dudley-Evans. Just like Peacock, Kanoksilapatham combines statement of result and finding into "stating selected findings." Furthermore, Kanoksilapatham added a new step: exemplifying. In this step, authors give an example that justifies a certain statement.

Aspects not explicitly mentioned by genre analysts are alternative explanations or alternative interpretations of results. These doubts, as Suppe (1998) calls them, are included in a selective way: "…confining attention only to those specific doubts the discipline recognizes as legitimate counterpossibilities" (Suppe, 1998, p. 384). According to Suppe, doubts are often coupled with rejoinders, which impeach these alternatives as much as possible (see Section 3.2). The explicitation of these doubts and rejoinders plays an important part in the persuasive process:

> There will always be more than one plausible interpretation of any piece of data and more than one way of looking at any problem. So while all claims require ratification to become knowledge, readers always have the option of refuting them. At the heart of academic persuasion then, is the writer's attempts to anticipate possible negative reactions to his or her claims. (Hyland, 1999, p. 78)

Genre analysts do mention recommendations (for further research), but remarks that summarize the potential significance of the findings are often not included in their descriptions of the Discussion section (Alexandrov, 2004). Nwogu (1997) is one of the few who distinguishes these research implications: a summing up of "the writer's views on the contributions which the study has made to the field" (p. 133).

3.4 The Comprehensive Argumentative Structure and the Scientific Argumentation Model

In this section, we will construct a model that describes a research article's intrinsic line of reasoning, stretching from the very reason to undertake the research, through data collection and interpretation, to the implications of the outcome. This model will be based on the abovementioned rhetorical moves and ideas from argumentation theory. We call this model, depicted in Figure 3.4, the Comprehensive Argumentative Structure (CAS). This model contains those rhetorical moves that, separated from the text and reassembled into a scheme, coherently show the line of reasoning of the research article as a whole. Next, we will further adapt this structure into a more concise version that is fitted for novice readers: the Scientific Argumentation Model (Figure 3.5). We use the words argumentative and argumentation because the encompassing feature of these models is the use of evidence to support claims and propositions (Andrews, 2010): the motive supports the objective, supports support each other and the main conclusion, and so forth.

The Comprehensive Argumentative Structure

Introduction section. As stated above, Swales (1990) identified three rhetorical moves in the Introduction section: establishing a territory, establishing a niche, and occupying the niche. This first move is quite general, because the significance of a research field is not unique to a specific research article. Other research articles from the same field will often describe the same significance. So, this move was not included in CAS. In contrast, establishing a niche and occupying the niche are very specific for a study. Because the concept of niches is probably difficult to grasp for students, we called the establishment of a niche the *motive* of the study (why was the study done) and the occupation of a niche the *objective* of the study (what did the authors want do know?). The objective may be formulated as a research question, a research aim, or a hypothesis that needs to be tested. The motive and objective are related to each other, as the objective emerges from the motive.

Method section. Nwogu (1997) and Kanoksilapatham (2005) both show that the Method section contains moves that describe experimental procedures and moves that describe data analysis/statistical procedures (or, as we called this move: data processing procedures). Additionally, Nwogu (1997) mentions a move describing data collection procedures. Kanoksilapatham (2005) also mentions a move describing the materials and a move detailing the equipment. For the sake of simplicity, we share these methodological aspects under experimental procedures. So, this leaves us with two moves: (1) description of *experimental procedures* and (2) description of *data processing procedures*. The former is directly connected to the objective, because the objective has influence on how the experimental procedures will be designed. In turn, the experimental procedures will influence the data processing procedures.

Results section. Statement of finding (after Williams, 1999) is arguably one of the most essential moves in the Results section. These are statements that provide relatively simple interpretations of the *inscriptions* (usually figures or tables). Statements of findings can lead to a *preliminary conclusion*. This is a generalization/interpretation of a statement of finding (Kanoksilapatham, 2005). Preliminary (or sub-) conclusions can be located in the Results or Discussion section. A statement of finding could be: “Rats which were given drug X had a lower blood pressure than rats in the control group.” A preliminary conclusion could be: “Drug X lowers the blood pressure in rats.” Of course, the difference between statements of findings and preliminary conclusions will not always be clear-cut. It is possible that the authors only present statements of findings in the Results section and save their preliminary conclusions for the Discussion section. Inscriptions, statements of findings, and preliminary conclusions are comparable to Kelly and Takao’s (2002) epistemic levels (see Section 3.2). Furthermore, the lines of reasoning between these epistemic levels may be described as a chain of interpretative steps or “externalisations” (Pinch, 1985).

Genre analysts who studied biochemistry articles have identified a number of other moves in the Results section (e.g. statements about why and how the data was collected), but these seem to be rather rare in biomedical research articles (Williams, 1999). This is why we left these moves out in the section of CAS that represents the Results section.

Discussion section. One of the most important elements in the Discussion section is the claim: generalizations arising from the results (Dudley-Evans, 1994). In our model, there are two types of claims (which we call conclusions, following Day, 1998): preliminary conclusions and the *main conclusion*.[5] The main conclusion forms the authors' thesis. This main conclusion is ultimately supported by inscriptions and a wide variety of moves that are located in the Results and Discussion sections: statements of findings (often based on inscriptions), preliminary conclusions and references to previous research. Together, these moves (and inscriptions) form the *supports* of the main conclusion. Our main conclusion may also include a hedge/qualifier. The main conclusion leads to remarks that present the potential significance of the findings and the possible influence on society, research, or clinical practice (Alexandrov, 2004) and recommendations. We call these remarks the *implications*.

Furthermore, we grouped comments about an unexpected outcome and limitations as *counterarguments*, because they shed doubt on the validity or generalizability of the main conclusion. Suppe's (1998) doubts also fall into this category. A counterargument can be "weakened" by *refutations* (or rejoinders, as Suppe calls them). An explanation (reasons for an unexpected result) can serve as a refutation in the case of an unexpected outcome.

We did not incorporate the information move of Dudley-Evans (1994) in our model, because this move (which describes background information about theory, research aim, methodology, or previous research) is a recapitulation of the motive and/or objective. We also left out the exemplifying move of Kanoksilapatham (2005). These exemplifications support the arguments in a research article by contributing "to the creation of coherent, reader-friendly prose while conveying the writer's audience-sensitivity and relationship to the message" (Hyland, 2007, p. 266). However, they are not central to the article's argumentation and that is why we did not include them in our model.

In summary, we think that the following moves describe the most essential part of the rhetorical structure of a research article: motive, objective, experimental procedures, data processing procedures, statements of findings,

[5] Preliminary conclusions are different from additional conclusions. In contrast to preliminary conclusions, additional conclusions do not support other conclusions.

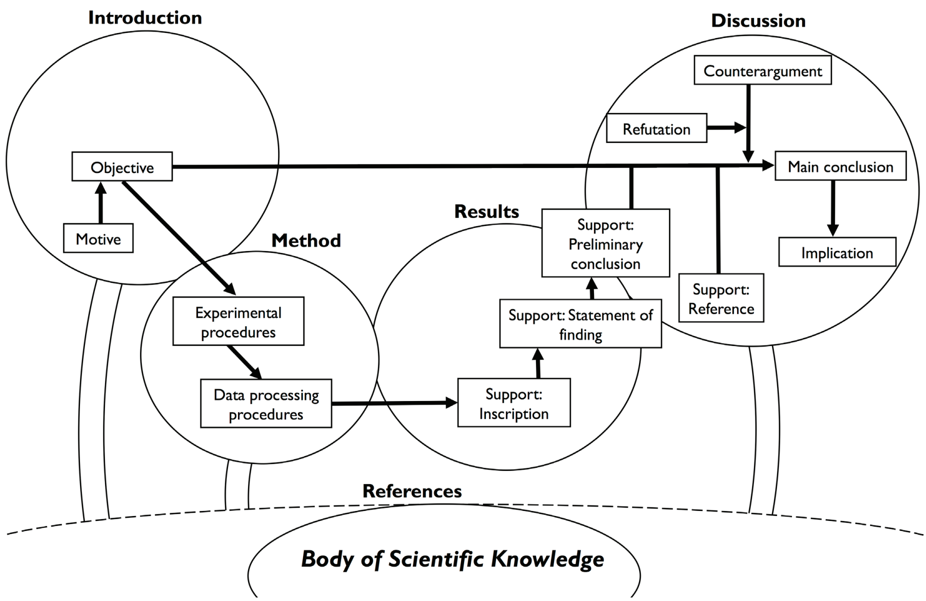


Figure 3.4. The Comprehensive Argumentative Structure (CAS), depicting the rhetorical moves of a research article and their relations.

preliminary conclusions, main conclusion, implication, counterarguments, and refutations. Inscriptions, to which statements of findings often refer, also play an important role in a research article's argumentation but are not rhetorical moves due to their extra-textual nature. However, because of their importance, we will subsequently group them together with rhetorical moves. Inscriptions, statements of findings, preliminary conclusions, and references to previous research are called supports, as they all support the main conclusion. These moves and their relations are shown in CAS (Figure 3.4). The four different sections of a typical research article (Introduction, Method, Results, and Discussion) are represented by circles. In these circles, the different moves are represented. The central arrow of the model is the connection between the objective and the main conclusion.

Below the central arrow[6] the supports – which form the pillars that justify the main conclusion – are placed. Supports in these pillars are connected to each other with arrows. Above the central arrow, all the counterarguments are placed. The counterarguments "cut" with an arrow through the connection between objective and main conclusion.[7] In turn,

[6] CAS contains three types of arrows/lines: (1) the central arrow, (2) the supportive arrows/lines (located below the central arrow), and (3) arrows that "cut" (located above the central arrow).

[7] The arrow symbolizes the potential nature of a counterargument. When a counterargument is completely refuted, it becomes a neutral element in the argumentation

a counterargument may be weakened or refuted by refutations that cut its connection with the central arrow. These refutations are not always present; authors do not always refute a counterargument. We chose to place counterarguments above the central arrow to emphasize that their argumentative function is the exact opposite from supports. Finally, the stems, connecting the circles to the "body of scientific knowledge," symbolize how the references in the Introduction, Method, and Discussion sections connect the article with the body of scientific knowledge.

The Scientific Argumentation Model

For educational purposes we stripped the comprehensive model CAS down to such a level that on one hand it contains only those text elements that are understandable for non-expert readers and on the other hand still shows a research article's entire line of reasoning. We omitted all the elements from the Method section, because understanding this section requires a great amount of prior knowledge due to its technical details. We kept the chains of supports, but decided to group all different supports (inscriptions, statements of findings, preliminary conclusions, and references) together in one category. This makes the identification of supports less complicated. We will call the remaining framework the Scientific Argumentation Model (SAM). SAM consists of a description of seven rhetorical moves (see below) and an argumentation scheme showing how these seven moves are related to each other. This visual representation is called the SAM scheme and is depicted in Figure 3.5.

The graphic properties of the SAM scheme are derived from the CAS model. Although the scheme lacks the lowermost connection from the objective to the supports through elements from the Method section, it still features a complete line of reasoning because of the direct link between objective and main conclusion. The objective and main conclusion relate to each other as question and answer or as expectation and evaluation. They often share the same wording, which signals readers that they are connected. The elements of the SAM model are described as follows:

- Motive: Why was the research done (e.g. a gap in knowledge, contradictory results)? The motive leads to the objective.
- Objective: What do 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.
- Main conclusion: What is the main outcome of the research? The main conclusion is closely connected to the objective. It answers the

scheme; in these cases there is no "cut."

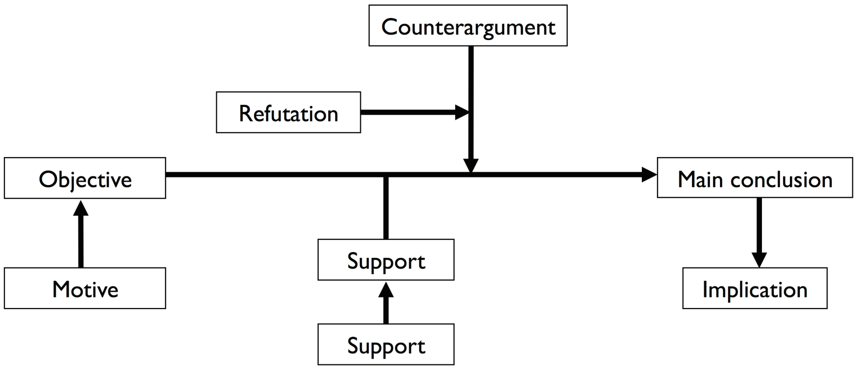


Figure 3.5. The Scientific Argumentation Model (SAM) scheme, showing the abridged argumentative structure of a research article.

research question, it says if the research aim was reached, or it states if the hypothesis was true or false. The main conclusion will lead to an implication.

- Implication: What are 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.
- Support: All the elements the authors use to justify their main conclusion. These elements can be based on own data (or their interpretation) or are statements from 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).
- Counterargument: Statement that weakens or discredits the main conclusion. For example, possible methodological flaws, anomalous data, results that contradicts previous studies, or alternative explanations. Counterarguments are sometimes presented as limitations.
- Refutation: Statement that weakens or refutes a counterargument.

At the beginning of this chapter, we formulated five criteria to which our heuristic should adhere. We think that SAM adheres to all of these criteria: it describes seven essential moves and their relations, the move descriptions are relatively straight-forward, it is visual, and it is generic due to its broad definitions of the moves (e.g. it takes into account that some authors only mention a research question, some only a hypothesis, et cetera).

We expect that SAM can help students with three activities: reading research articles, critically evaluating research articles (as it gives insight into a research article's argumentation), and writing research articles. In Chapters 4, 5, and 6 we will study the effects of SAM on the first of these three activities. In Chapter 7, students will use SAM to critically evaluate research articles. Writing research articles with the support of SAM will not be studied in this thesis. However, in Chapter 8 we will present some ideas about a reading and writing curriculum for undergraduate science students.

In the next chapter we will use the SAM model in a teaching strategy that aims to help students with reading research articles. For this purpose, we constructed so-called information sheets in which we describe the SAM model's seven moves in a similar way as we did above (see Appendix 1). These descriptions were coupled with examples taken from authentic research articles.

Chapter 4
A Teaching Strategy to Improve Students' Rhetorical Consciousness of Research Articles

4.1 Introduction

In a preliminary investigation – described in Chapter 2 – we described several difficulties that undergraduate students have with reading research articles. Students identified a wide range of sentences as conclusions that were not seen as conclusions by experts. Our results also indicated that students and experts – to which we compared students' performances – used different criteria to identify the two studied rhetorical moves (conclusions and supports).

Based on these results, we decided to develop a heuristic, the Scientific Argumentation Model (SAM), which could potentially help students with reading research articles. SAM was based on ideas from genre analysis and argumentation theory (see Chapter 3). SAM consists of a set of seven rhetorical moves (motive, objective, main conclusion, implication, supports, counterarguments, and refutations) and an argumentation scheme (the so-called SAM scheme, see Figure 3.5), which depicts the relations between these moves.

In this chapter, we will describe our teaching strategy (in which SAM played a crucial role) and then we will evaluate to what extent students are able to identify the before-mentioned rhetorical moves. We will not evaluate students' ability to construct a SAM scheme. This will be done in Chapter 6.

4.2 Teaching Strategy

The first and most central design principle (see Section 1.8) of our teaching strategy is the focus on the rhetorical structure of research articles. This approach – based on the work done in the field of genre analysis and argumentation theory – means that we teach students how to read research articles by instructing them which rhetorical moves occur in research articles and how they can identify these. As research articles are persuasive in nature, we focus on the rhetorical moves that play an important role in the authors' argumentation.

In this way, we wanted to improve students' rhetorical consciousness: the knowledge about the rhetorical structure (i.e. the arrangement of rhetorical moves) in a genre-specific text (Swales, 1990). In van den Akker's (1999) conceptualization of design experiments (see Section 1.8) this central design principle concerns a substantive emphasis.

Thus, the focus on rhetorical structure formed the backbone of our learning ecology (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003). The second design principle, which complemented our learning ecology, is cognitive apprenticeship. Cognitive apprenticeship is – following the terminology by van den Akker (1999) – a procedural emphasis.[1] We chose this approach because cognitive apprenticeship aims to enable students to acquire, develop, and use cognitive tools in an authentic domain activity. In this way, students enter a culture of practice – in our case the scientific community (Lave & Wenger, 1991) and move closer into its activity system (see Section 1.7). In cognitive apprenticeship the centrality of activity in learning and knowledge is emphasized while the context-dependent nature of learning is highlighted (Brown, Collins, & Duguid, 1989).

According to Collins, Brown, and Holum (1991), cognitive apprenticeship involves three characteristics: Firstly, the processes of the task should be identified and made visible for students. Secondly, abstract tasks should be situated in authentic contexts (so that students will understand the relevance of the task). Thirdly, the diversity of situations should be varied and common aspects should be articulated (so that students may transfer what they learn).

Collins, Brown, and Holum (1991) describe six teaching methods of which the first three form the core component of cognitive apprenticeship:

- Modeling: An expert performs a task while students observe him or her. In this way, students may build a conceptual model of the processes that are necessary to accomplish the task.
- Coaching: Students are offered hints, feedback, modeling, reminders, scaffolding, and new tasks while they carry out their task.
- Scaffolding: Students are given supports that help them to carry out the task. Eventually, fading takes place: the gradual removal of supports until students are able to accomplish a task on their own.
- Articulation: Students articulate their knowledge, reasoning, or problem-solving processes.

[1] This design principle was also used during our preliminary investigation.

- Reflection: Students are enabled to compare their own problem-solving processes with another student, expert, or – ultimately – an internal cognitive model of expertise.
- Exploration: When all supports are removed, students are pushed to explore problems or questions on their own.

This second design principle led to two derived design principles. The first derived design principle is *authenticity*. We used authentic research articles that were not edited, translated, or adapted. The second derived design principle is *interaction*. Students should be encouraged to discuss with peers and more experienced readers what they have read, which results in an environment where appropriation can take place (see Section 1.7). To achieve this, we used cross-year small-group tutoring (tutoring by students from other years). By guiding the students through research articles (via modeling and coaching), the tutors are functioning as experts who train the newcomers in a certain skill. Tutors used the heuristic described in the previous chapter as support (scaffolding). Furthermore, during the tutor group meetings students could articulate their knowledge, reasoning, or problem-solving processes and compare these with other students and the tutor (articulation and reflection). Topping (1996) lists a number of supposed advantages of peer tutoring:

> Pedagogical advantages for the tutee include more active, interactive and participative learning, immediate feedback, swift prompting, lowered anxiety with correspondingly higher self-disclosure, and greater student ownership of the learning process. The 'pupil/teacher' ratio is much reduced and engaged time on task increased. Opportunities to respond are high, and opportunities to make errors *and be corrected* similarly high. In addition to immediate cognitive gains, improved retention, greater meta-cognitive awareness and better application of knowledge and skills to new situations have been claimed. Motivational and attitudinal gains can include greater commitment, self-esteem, self-confidence and empathy with others. (p. 324–325)

Topping (1996) reviewed a number of studies on the effectiveness of cross-year small-group tutoring and concludes that "much of the research is not of the highest quality, but good quality studies (…) do clearly demonstrate improved academic achievement" (p. 327).

It has to be noted that peer tutoring could have some disadvantages. For example, the tutor's mastery of content is probably less than that of a faculty member, or it is possible that tutors lack teaching skills to guide the learning process of the students as intended.

4.3 Research Questions

The teaching strategy was executed in 2009 as the module Reading Research Articles. This module was part of the course Biomedical Research (see Section 1.9 for a description). In this chapter, we will confine ourselves by testing our first design principle (the focus on rhetorical structure). By determining how well students are able to identify rhetorical moves (i.e. the moves that are described in SAM) in a research article, we want to measure the level of their rhetorical consciousness. Our first two research questions were:

1. How does the rhetorical consciousness of undergraduate life science students improve after following our course module?
2. What is students' own perception on their improvement of their rhetorical consciousness?

We answered the first research question by determining students' progression in the identification of rhetorical moves. For this purpose, we tested the students at the beginning and at the end of the module. We answered the second research question – about students' own perception on their improvement of their rhetorical consciousness – by analyzing questionnaire data collected at the beginning and end of the module. Our third research question was:

3. How do students' reading behaviors change during the course module?

To answer this question, we monitored their improvement by studying students' self-reported change in reading behaviors. Bazerman (1988) already established that experts read selectively (skipping parts of the text; only reading the newsworthy parts) and non-sequentially (not reading the parts of an article in order). That is why we determined if more students read selectively and non-sequentially at the end of the module. Furthermore, we let the tutors fill in questionnaires to check how our teaching strategy was implemented.

4.4 Method

Participants

The total number of first-year undergraduate life science students that entered the course was 125. They were approximately 18–20 years old. Fifty-five students were male and 70 students were female. At the beginning of the course, the students filled out a questionnaire in which they had to indicate how many research articles they read before the course. The 125 students were randomly divided into 14 tutor groups. The tutors were between 20 and 23 years old, studied life sciences (n = 8) or medicine (n = 6), and were third-year bachelor students or master students.

Teaching Strategy

In Week 1 and Week 8 of the module students handed in their answers to a pre-test and post-test to measure their rhetorical consciousness. From Week 2 up to Week 7 the teaching strategy was implemented (i.e. the rhetorical moves were introduced to students). Our module consisted of tutor group meetings and homework assignments. Each week, at the end of the meeting, students received a new research article (the "article of the week"), an assignment, and instructions from the tutor (Figure 4.1). During preceding lectures students were made familiar with the concepts discussed in the article. Additionally, students received information sheets – based on SAM – in which we listed the different moves of a research article, together with definitions and examples taken from authentic research articles (see Appendix 1). To the tutors, we emphasized that they should demonstrate how they read a research article and identified rhetorical moves (see Chapter 5 for more details about the instructions we gave to the tutors). The assignment consisted of reading the research article and answering questions on paper. Students made the assignment as homework. During the next tutor group meeting, students discussed the research article and the answers to the assignment. For this purpose, tutors could use so-called discussion prompts that were developed and provided by us and concerned the methodology, meaning of certain technical terms, interpretation of results, and connections with other studies. Additionally, students received written feedback from the tutor on their answers. The

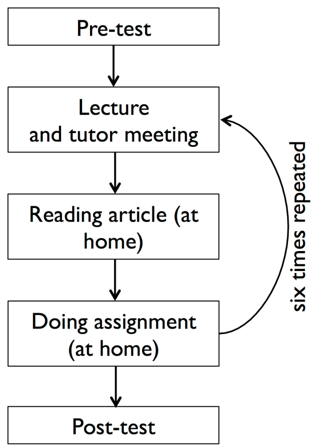


Figure 4.1. Schematic representation of our teaching strategy.

tutor meetings were held in small meeting rooms at the university. Meetings usually lasted 2–2.5 hours and were mandatory.

Move identification assignments. During the module students received six research articles and six homework assignments. Central part of these assignments was the identification of the seven rhetorical moves that are central to SAM: motive, objective, main conclusion, implication, supports, counterarguments, and refutations. We did not want to overload students at the beginning of the module by letting them identify all rhetorical moves at once. That is why we followed a cumulative approach: in Week 3 they identified only the motive and objective, in Week 4 the motive, objective, main conclusion, and implication, et cetera. Once we introduced a move, the students had to identify this move in all subsequent assignments. We hoped that by repeating the identification process, students would rely less on the information sheets (our scaffolding method) as the module progressed. In this way, fading could occur. Each week, a different research article was used. The assignments students made before the tutor group meeting were:

- Week 2: Students did not identify rhetorical moves; instead they focused on the general structure of research articles.
- Week 3: The Introduction section was closely examined. Students identified the motive and objective.
- Week 4: The Discussion section was closely examined. Students identified the motive, objective, main conclusion, and implication.
- Week 5: The Results section was closely examined. Students identified all rhetorical moves.
- Week 6: The Method section was closely examined. Students identified all rhetorical moves.
- Week 7: Students identified all rhetorical moves and constructed a SAM scheme.

See Chapter 5 (in particular Table 5.1) for more information about these assignments.

Additional assignments. Students did not only identify rhetorical moves for the homework assignments. Each week, they also answered additional questions that were partly based on assignments published by Yudkin (2006) and Gillen (2007). A summary of the additional questions students had to answer before the tutor meeting in each week is described below.

For Week 2, students answered questions about the general structure of research articles. We asked them to formulate the function of the different sections (Abstract, Introduction, Method, Results, and Discussion) and to summarize these sections of the article of the week. They formulated criteria for a good title and determined if the article of the week's title followed these criteria. Furthermore, they answered questions about the article of the week's references and self-references, the time between acceptance and publication, and the funding of the research.

For Week 3, students chose five important concepts mentioned in the Introduction section and explained their meaning. We asked them to describe the field of research and why authors refer to previous research in their Introduction section. They summarized this previous research.

For Week 4, students were asked if the main conclusion and implication were related to each other and how certain the authors were about their main conclusion.

For Week 5, students identified the most essential figure or table and motivated their choice. Students evaluated the quality of the refutations. Finally, the students devised a new counterargument (not mentioned by the authors) and refutation.

For Week 6, students described the article's experimental and control groups and summarized four experiments.

For Week 7, no additional questions were asked.

Pre-test and Post-test

To measure the effectiveness of our teaching strategy, students made a pre- and post-test as homework in Week 1 and Week 8. The pre-test article and assignment were given at an introductory meeting a week before the module started. The post-test article and assignment were given in Week 7. The assignments consisted of the following questions:

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

When we formulated the questions, we deliberately did not use the term counterargument. At the time of the pre-test, students had not received instructions about the meaning of this term and we wanted to make sure that students would not become confused. We assumed that students would be able to understand the other terms during the pre-test. We implemented the pre-test and post-test via a parallel test. The tutor groups were divided into Group A (n = 72 students) and B (n = 53 students). The groups were not of the same size because of logistic reasons. During the pre-test, Group A received Article 1 ("The Protective Effect of Fish n-3 Fatty Acids on Cerebral Ischemia in Rat Hippocampus" by Bas et al., 2007) and Group B received Article 2 ("The Protective Effect of Fish n-3 Fatty Acids on Cerebral Ischemia in Rat Prefrontal Cortex" by Ozen et al., 2008). During the post-test, Group A received Article 2 and Group B Article 1. By switching the articles, we could eliminate the possibility that the measured progression was the result of a post-test article that was easier to read and understand than the pre-test article.

Articles 1 and 2 described the effects of a fish oil (n-3 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 number of apoptotic neurons was counted and the levels of several biomarkers (NO, CAT, SOD, and MDA) were measured to determine the amount of damage to the rats' brains. In contrast to the other articles used in the module, the concepts mentioned in Articles 1 and 2 were not explicitly discussed by the lecturers of the course.

The main body of Article 1 contained approximately 4,000 words, one figure, and two tables. Article 2 contained approximately 3,300 words, two figures, and one table. Readability of the articles was measured using the Flesch Reading Ease Score (see Section 2.3). Article 1 had a Flesch Reading Ease score of 50. Article 2 had a score of 46. This means that Articles 1 and 2 were, respectively, *fairly difficult* and *difficult to read*. Although the articles fall in different scoring categories, we will assume that the articles are more or less equivalent in readability, as the difference between the two scores is very small.

Students' answers (paraphrases or quoted verbatim text fragments) were collected via e-mail. Because 17 students did not hand in the pre-test and/or post-test, we analyzed the answers of 66 (Group A) and 42 (Group B) students. After each pre-test or post-test, we asked the students via an on-line questionnaire how much time they spent on the test (including reading the article) and how they rated the understandability of the article on a 4-point scale: *Hard to understand*, *reasonably difficult to understand*, *easy to understand*, or *very easy to understand*. Furthermore, to determine their reading behaviors, students had to indicate on a 4-point scale how well they read the different sections of the research article: *not*, *casually*, *good*, or *very good*. Their answers were used to assess to what extent they read selectively. Students also indicated if they read the article in a sequential or non-sequential way.

Questionnaires

Students' self-assessment. After making the pre-test and post-test, students completed an on-line self-assessment in which they indicated via a 5-point rating scale (from *strongly disagree* to *strongly agree*) how much they agreed to five different statements. The statements dealt with students' ability to identify certain rhetorical moves in research articles:

- I am able to read a scientific research article in a structured way.
- I am able to identify the research question.
- I am able to identify the results.
- I am able to identify the conclusion.
- I am able to identify the supports used in the discussion.

Tutors' views on the implementation of the design. To check if the tutors (n = 14) followed the instructions for the meetings, they filled out each week (from Week 2 to Week 7 – only the weeks when they gave instructions about rhetorical moves) an evaluation form about how their tutor meetings proceeded. We asked them the following open questions:

- What was the structure of the meeting? Did you use the given instructions? If not, why?
- If you used the instructions: did you deviate from them? If so, what did you do different and why?
- How did the meeting proceed? What went well and what could have gone better?
- Did the students encounter problems? If so, which problems?
- Do you have any further comments about the assignment or the instructions?

The questionnaires were collected via e-mail. On average, about 11 tutors sent in a completed questionnaire each week.

Data Analysis

Pre-test, post-test, reading behaviors, and self-assessment. For the analysis of the answers to the pre-tests and post-tests, I blindly rated students' answers (i.e. I did not know which answer came from the pre-test and which answer came from the post-test). For this analysis we devised a scoring model based on the definitions of the different rhetorical moves (Table 4.1). For each move, we determined a number of elements that should be present in the student's answer. For each element, we gave one point. Then, we calculated percentage scores for each answer. If, for example, the answer of a student contained the maximum amount of elements, his or her score was 100%. We did this for the following moves: motive (the answer to Question 1 of the abovementioned assignment), objective (Question 2), main conclusion (Question 5), implication (Question 6), and counterargument (Question 7). For the analysis of the supports, we made a selection of the most important supports (the answer to Question 4) for the main conclusion and counted how many students mentioned these (but only for the students who mentioned the correct main conclusion in their answer to Question 3). The results of this analysis are shown in Tables 4.2 and 4.3.

To check the reliability of our method, a second researcher also blindly rated 120 randomly chosen items. Krippendorff's alpha (Hayes & Krippendorff, 2007) was 0.98, which indicates that there was a high inter-rater agreement. We used the first author's ratings to calculate the average score per group per move.

For determining the significance of the differences between the pre-test and post-test scores of each article we used a Mann-Whitney *U* test.

We also used the Mann-Whitney *U* test for determining the significance of the differences between the perceived understandability of the articles used in the pre-test and post-test.

The significance of the differences in time spent on reading the article and making the assignment between the pre-test and post-test was determined by an independent samples student *t*-test.

A Wilcoxon signed-rank test was used to analyze students' answers to our question about how well they read the different sections of the research articles.

Regarding the self-assessment, significant differences between the beginning and end of the module were determined via a Wilcoxon signed-rank test.

To determine if students' perception of their improvement differed from their actual improvement we calculated for each student the difference between the pre-test and post-test percentage scores (see above). Students were divided into two groups: students who showed progression (having a post-test score which was higher than the pre-test score) and students who did not show progression (having a post-test score which was equal to or lower than the post-test score). For both groups, the average scores

Table 4.1. Scoring model used for the analysis of the pre-test and post-test. For each element placed between brackets, the student was awarded with one point.

	Article 1	Article 2
Motive	However, to our knowledge [there is no report] on the [number of apoptotic cell] as [a marker of neuron damage] and [antioxidative functions] of [n-3 EFA treatment] in [ischemic hippocampal formation.]	However, we think that [fish n-3 fatty acids] [may show] [different effects in different parts and/or sections] of the [brain] because of their functional and structural specialties.
Objective	The aim of this study was [to investigate] the [antioxidant] and [neuroprotective effects] of [fish n-3 EFA] on [cerebral I/R injury] in Sprague–Dawley [rats'] [hippocampal formation]	In this study, we [investigated] the [antioxidant] and [neuroprotective effects] of [fish n-3 EFA] on [global cerebral I/R injury] in [rats'] [prefrontal cortex.]
Main conclusion	Also [fish n-3 EFA] [ameliorates] the [oxidative status] and [apoptotic changes] in [rat] [hippocampus] following [I/R.]	Additionally, [fish n-3 EFA] [ameliorates] the [oxidative stress] and [apoptotic changes] in [rats'] [prefrontal cortex] following [I/R injury.]
Implication	We conclude that [dietary supplementation] of [fish n-3 EFA] [may be] [beneficial to preserve or ameliorate] on [ischemic cerebral vascular disease.]	This preliminary study suggest that [dietary supplementation] of [fish n-3 EFA] [may be] [beneficial to preserve or ameliorate] the [ischemic cerebral vascular disease.]
Counterargument	[It is questionable] whether [MDA] is really [a reliable marker] for [oxidative stress]	Although [the biochemical tissue results] of [this study] have [some different parameters] with [our hippocampal study]…

on the two self-assessments were calculated (for this purpose, *strongly disagree* was given 1 point, *disagree* 2 points, et cetera).

Tutors' views on the implementation of the design. For the analysis of the tutor questionnaires, two researchers analyzed the answers and marked fragments that gave an indication to what extent the tutors followed our instructions. We looked especially at possible problems that tutors had with the implementation of our teaching strategy.

4.5 Results

Students' Reading Experiences and Academic Performance

All 125 students were novice readers of primary literature and their native language was Dutch. At the beginning of the module, these students (minus 15) filled out the questionnaire about their reading experiences regarding primary literature. Five students had not read any research article at all before the beginning of the module. Eighty-three students had read 1–6 articles, 19 students had read 7–12 articles, and 3 students had read more than 12 articles. Groups A and B were comparable, since we determined that there was no significant difference between the two groups regarding (1) the number of articles read before the module and (2) their grades on the course's final examination (data not shown).

Identification of Rhetorical Moves

We determined students' ability to identify rhetorical moves (motive, objective, main conclusion, implication, and counterargument) by attributing points to their answers based on our scoring model (Table 4.1). The distributions of points given for the different answers of students are depicted in Figures 4.2 and 4.3. As can be seen, students scored more often the maximum number of points for their answers in the post-test than in the pre-test (with the exception of the counterargument of Article 1). Furthermore, the number of semi-correct answers (i.e. answers which were awarded with 1 point or more but not the maximum amount of points) decreased. This means that students' progression is due to the fact that more students identify a certain move, and not because students give more complete answers.

The average percentage scores for Articles 1 and 2 are depicted in Figures 4.4 and 4.5 respectively. With the exception of the counterargument of Article 1, the average scores on all moves in the two articles were higher for the post-test than the pre-test. All these increases in average scores were significant, except for the objective of Article 1 and the counterargument of Article 2. The decrease in average score on the counterargument of Article 1 was not significant. In general, the students scored higher on the motive and objective than on the main conclusion, implication and counterargument.

We distinguished respectively five (Table 4.2) and four (Table 4.3) essential supports for the main conclusions of Articles 1 and 2.[2] In Article 1, three supports were more often identified in the post-test than in the pre-

[2] Interestingly, Article 1 shows that the levels of MDA and SOD are increased in animals that are treated with fish oil. In contrast, levels of CAT are decreased. In Article 2 these effects are reversed. However, it seems that in both articles the biomarker data are used as supports for the main conclusion that a fish oil diet is effective.

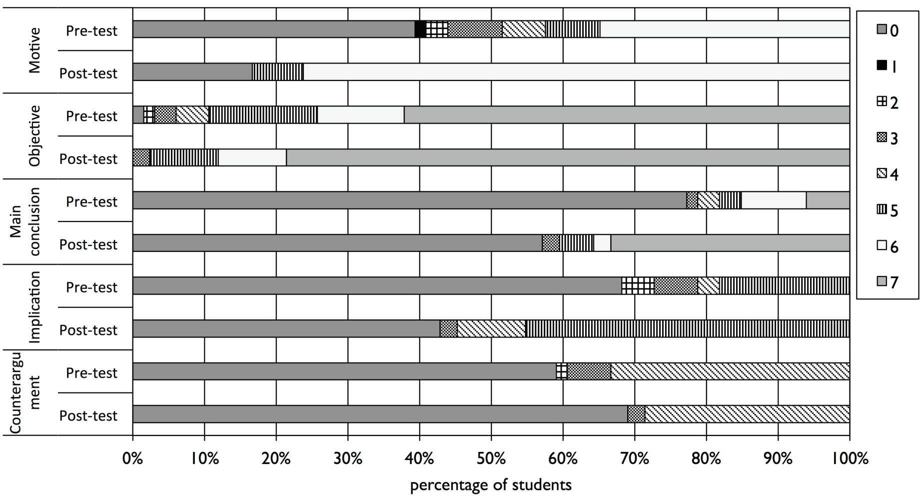


Figure 4.2. Distribution of points given for the sentences identified as moves by students in Article 1. The maximum number of points which could be scored for the motive, objective, main conclusion, implication, and counterargument is respectively 6, 7, 7, 5, and 4.

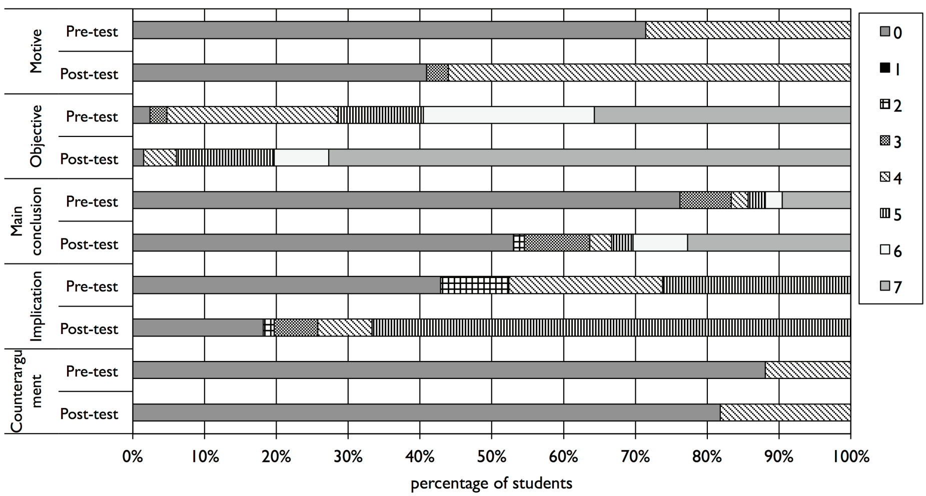


Figure 4.3. Distribution of points given for the sentences identified as moves by students in Article 2. The maximum number of points which could be scored for the motive, objective, main conclusion, implication, and counterargument is respectively 4, 7, 7, 5, and 4.

test. The other two supports showed a small decrease. Regarding Article 2, we saw with two supports an increase, and with two supports a small decrease. It has to be noted that only nine students were able to find the main conclusion in Article 2 during the pre-test, so the sample size is very small.

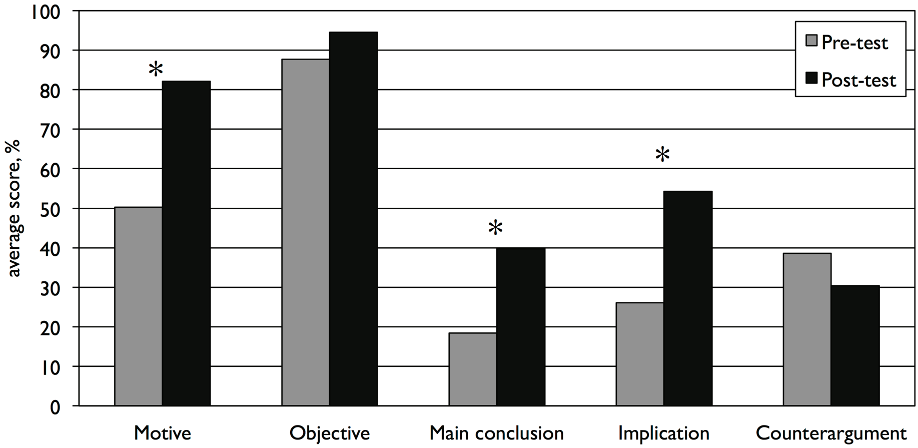


Figure 4.4. Average student scores when identifying rhetorical moves in Article 1 at the beginning (pre-test; $n = 66$) and end (post-test; $n = 42$) of the module. Significancies of the differences between pre-test and post-test (Mann-Whitney U test): motive ($U = 815.5$, $z = -3.9$, $p < 0.001$), objective ($U = 1135.5$, $z = -1.9$, $p = 0.055$), main conclusion ($U = 1043.5$, $z = -2.7$, $p = 0.008$), implication ($U = 965$, $z = -3.0$, $p = 0.003$), counterargument ($U = 1267.5$, $z = -0.9$, $p = 0.379$). An asterisk indicates $p < 0.05$.

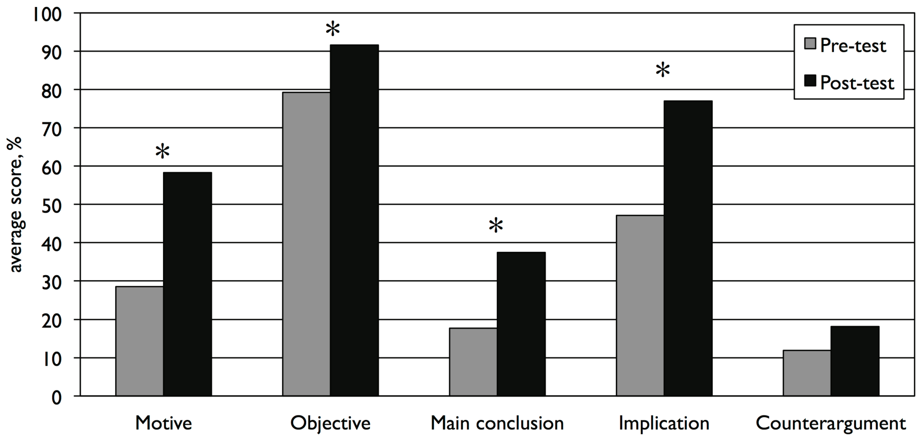


Figure 4.5. Average student scores when identifying rhetorical moves in Article 2 at the beginning (pre-test; $n = 42$) and end (post-test; $n = 66$) of the module. Significancies of the differences between pre-test and post-test (Mann-Whitney U test): motive ($U = 975$, $z = -3.0$, $p = 0.003$), objective ($U = 854$, $z = -3.8$, $p < 0.001$), main conclusion ($U = 1047.5$, $z = -2.5$, $p = 0.014$), implication ($U = 817.5$, $z = -3.9$, $p < 0.001$), counterargument ($U = 1299$, $z = -0.9$, $p = 0.385$). An asterisk indicates $p < 0.05$.

In their written answers, students gave a number of different reasons to explain why they identified a certain statement as main conclusion. In their answers to the pre-test, 19 of the 108 students motivated their choice by stating that their conclusion answers the research question. However, some students gave a reason that was based on the place of the sentence in the text. For example, during the pre-test, one student wrote: “The first conclusion [is the main conclusion]. They go on about this one the most and it is first mentioned.” Only 2 students mention that the main conclusion is a summary. In the answers to the post-test, students did not explicate much. Regarding the identification of the Article 2’s main conclusion, 31 (pre-

Table 4.2. Percentage of students mentioning certain supports for the main conclusion of Article 1 (n = 29 for the pre-test and n = 31 for the post-test). The treated group are the rats with cerebral injury which have been given fish oil. The non-treated group are the rats with cerebral injury which have been given a normal diet.

Support	Pre-test	Post-test
MDA higher in treated group than in non-treated group	14	23
SOD higher in treated group than in non-treated group	24	32
NO lower in treated group than in non-treated group	34	29
CAT lower in treated group than in non-treated group	17	32
In treated group less apoptotic neurons than in non-treated group.	45	42

Table 4.3. Percentage of students mentioning certain supports for the main conclusion of Article 2 (n = 9 for the pre-test and n = 29 for the post-test). The treated group are the rats with cerebral injury which have been given fish oil. The non-treated group are the rats with cerebral injury which have been given a normal diet.

Support	Pre-test	Post-test
MDA lower in treated group than in non-treated group	33	41
SOD lower in treated group than in non-treated group	22	41
CAT higher in treated group than in non-treated group	44	38
In treated group less apoptotic neurons than in non-treated group	56	52

test) and 10 (post-test) students identified the following sentence: "These results demonstrate that n-3 EFA has neuroprotective effect on global cerebral I/R injury with biochemical and histopathological parameters" (p. 151). Although this answer is not completely wrong, we think that our main conclusion relates in a more precise way to the objective of the study. Also, some students wrote down a sentence that was more an implication than a conclusion. A good example is the implication of Article 2, which was often identified as main conclusion by students: "This preliminary study suggest [*sic*] that dietary supplementation of fish n-3 EFA may be beneficial to preserve or ameliorate the ischemic cerebral vascular disease" (p. 152).

Understandability and Time Spent on Reading

We asked the students to rate the understandability of the two articles. During the pre-test, 68% of the students said they found the Article 1 *hard* or *reasonably difficult to understand* and 32% found the article *easy* or *very easy to understand*. During the post-test, these percentages were 10% and 90% respectively. This difference was significant (Mann-Whitney U test, $U = 468.5$, $z = -5.8$, $p < 0.001$). We saw a similar pattern for Article 2. During the pre-test 36% of the students said they found the article hard or reasonable easy to understand and 64% found the article *easy* or *very easy to understand*. During the post-test, these percentages were 14% and 86% respectively. This difference was also significant (Mann-Whitney U test, $U = 905$, $z = -2.0$, $p = 0.041$). So, during the post-test students found both articles easier to understand compared to during the pre-test.

Students stated that they spent significantly less time on the post-test compared to the pre-test. For Article 1, the average time spent on the pre-test (reading the article and making the assignment) was 163 minutes while the average time spent on the post-test was 107 minutes. This difference was significant (independent-samples t-test, $t(99) = 4.007$, $p < 0.001$). For Article 2, the average time spent on the pre-test was 142 minutes while the average time spent on the post-test was 105 minutes. This difference was also significant (independent-samples t-test, $t(97) = 3.215$, $p = 0.002$).

Students' Self-assessment

After the abovementioned pre- and post-test students assessed their abilities regarding reading and understanding research articles. Comparing the answers to the pre- and post-self-assessment questionnaires, we saw for all five statements a significant increase in the degree students agreed with them: I am able to read a scientific research article in a structured way (Wilcoxon signed-rank test, $z = -3.6$, $p < 0.001$); I am able to identify the research question (z = -4.8, $p < 0.001$); I am able to identify the results (z = -2.7, p = 0.008); I am able to identify the conclusion (z = -3.0, p = 0.002); I am able to identify the supports used in the discussion (z = -3.3, p = 0.001). This indicated that students thought that they had become better readers.

To determine the relationship between the scores on the pre-test, post-test, and self-assessments, we made a distinction between students who showed no progression (n = 28) on the post-test compared with the pre-test, and students who did show progression (n = 80). For the first group, the average pre-self-assessment score was 3.58 and the average post-self-assessment score was 4.03. For the second group, these numbers were respectively 3.67 and 4.03. So, the students who showed no progression between their pre-test and post-test scores do not differ greatly in self-assessment scores with students who showed a progression.

Reading Behavior

After completion of the pre- and post-test, students filled in a questionnaire about how they read the articles. Ninety-four students said they read the article during the pre-test in a sequential way and 16 students said they read the article in a non-sequential way. After the post-test, 80 students said they read the article in a sequential way and 38 students said they read the article in a non-sequential way.[3]

We asked students to indicate on a 4-point scale how much attention they paid to the different sections of the research article during the pre- and post-test. When we compared the post-test with the pre-test, students paid significantly less attention to the Abstract (z = -2.7, p = 0.007; Wilcoxon signed-rank test, two-tailed), Method (z = -5.3, $p < 0.001$), and Results (z = -3.7, $p < 0.001$) sections (Table 4.4). The attention paid to the Introduction section, Discussion section, and the figures and tables stayed the same. So, it seems that less students tended to read research articles in a sequential way and they read more selectively at the end of the module.

Tutors' Views on the Implementation of the Design

To determine if the tutors were following our guidelines, we asked them from Week 2 until Week 7 to fill out questionnaires with open-ended questions after each meeting. Tutors generally followed our guidelines. However, it seems that the tutors found it rather difficult to demonstrate how they identify certain rhetorical moves. As one tutor said: "I find it difficult to demonstrate; of course we discussed it." Some other tutors

[3] For this analysis, we also included the students who were left out in the analysis of the pre-test and post-test.

Table 4.4. Frequency of students' answers (in percentages) when asked how well they read the different sections of the pre-test and post-test articles (n = 110 for the pre-test and n = 118 for the post-test).

	Pre-test		Post-test	
	not/casually	*good/very good*	*not/casually*	*good/very good*
Abstract	24	76	42	58
Introduction	26	74	22	78
Method	40	60	75	25
Results	12	88	19	81
Discussion	14	86	9	91
Figures and tables	40	60	34	66

Note. For this table we grouped the students who answered *not* or *casually* and the students who answered *good* or *very good*.

skipped this part: "A quick poll revealed that everybody knew [how to identify the motive and objective] already, so it didn't seem necessary to examine the matter further." Other tutors relied heavily on the information sheets when they instructed the students: "The information sheets were so clear that I distributed them, let the students read them and gave them the opportunity to directly ask questions (…) Sometimes the assignment showed that they switched the motive/objective (or other elements), but this was discussed during the group meeting and in the feedback I sent them."

Summary of the Results

In summary, with the exception of counterarguments and certain supports, students' average move identification scores were higher on the post-test (compared to the pre-test). Regarding the identification of supports, there was, in general, not much difference between the pre-test and post-test. According to the self-assessment, students who scored higher on the post-test than on the pre-test, did not rate their progression in skills higher than students who did not score higher on the post-test. Both groups scored higher on the self-assessment given at the end of the module compared to the one administered at the beginning.

Furthermore, at the end of the module students found the articles easier to understand and they reported spending less time on reading the article and making the assignment. Students also reported that their reading behaviors changed: more students read in a non-sequential manner and they read more selectively.

Tutors generally followed our guidelines, although they had some difficulty with modeling (i.e. demonstrating how they identify certain moves). That is why some tutors choose to discuss rather than demonstrate, skipped this part, or relied heavily on the information sheets (scaffolding).

4.6 Discussion

Our study shows that our teaching strategy improves undergraduate students' rhetorical consciousness of research articles. Regarding the identification of the motive, objective (only of Article 2), main conclusion, and implication, the students showed significant improvements in their average scores. According to their own estimates, students made the assignment in less time. Given these results, it seems likely that our students have become better readers – i.e. they have a better grasp of rhetorical moves. Because the tutors did generally follow our guidelines, we conclude that the focus on rhetorical structure – our first design principle – may be a powerful tool for introducing undergraduate students to primary literature.

However, our study also showed that a majority of the students still have difficulty with identifying the main conclusion. This was quite puzzling, because in our previous study (see Chapter 2), we found that almost all students found the main conclusion in a research article. However, in this previous study we also showed that students relied on lexical features such as reporting verbs (e.g. *suggests*, *found*, *show*) and transition words/phrases (e.g. *overall*, *so*, *in summary*) for identifying conclusions. The main conclusions of Articles 1 and 2 that we identified did not contain these types of lexical features. This could explain why a majority of students did not find them. This is supported by the fact that a number of students identified a sentence containing the word "demonstrate" as the main conclusion of Article 2.

Looking in more detail at the texts of the two articles provides us with clues why some moves were easier to find than others. For example, the motive of Article 1 was identified correctly by more students than the motive of Article 2. The reason for this could be that the motive of Article 1 referred to a clear gap in knowledge. In contrast, the motive of Article 2 is more a speculation. Students' average score for the implication of Article 1 was lower compared to Article 2. This could be explained by the use of "we conclude" in the implication of Article 1. These words could have led some students to believe this statement was a conclusion.

We have observed that students had difficulties with finding counterarguments. This accords with Kuhn (1991), who found that her subjects had difficulties with recognizing the critical status of the counterinstances they encountered. A 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). Also, the counterarguments in the articles contained no distinctive lexical features. Another possible explanation is that students are used to reading textbooks. Textbooks tend to present knowledge claims without explaining how these claims came to be (e.g. Goldman & Bisanz, 2002). As a result, these texts seldom contain counterarguments. So, students tend to have limited experience with identifying counterarguments.

We observed that three of the five (in the case of Article 1) and two of the four (in the case of Article 2) essential supports were identified by more students at the end of the module compared to the beginning. However, for the other supports no increase was observed. Also, the support that was identified the most was identified by only 56% of the students. So, it seems that after our module students have problems with finding the evidence for a claim (i.e. the coordination of evidence). The same phenomenon was observed in Chapter 2. We can only speculate about possible reasons for this result. One reason may be that supports are quite numerous; this makes it difficult for students to determine which supports are important evidence for the main conclusion.

The observations listed above point at an interesting feature, namely that specific features of scientific texts are essential for students when identifying rhetorical moves. The occurrence of lexical features seemed to influence students' rhetorical consciousness. This will be further explored in Chapter 5.

For our analysis of the self-assessment, we separated students who showed progression on the pre-test/post-test from the students who did not show progression. Both groups scored higher on the self-assessment given at the end of the module compared to the one administered at the beginning. This shows that students thought that they have become better readers. The self-assessment was clearly not sensitive enough to differentiate between readers who showed progression and readers who showed no progression. This emphasizes the need for researchers to be very careful when drawing conclusions from data collected with a self-assessment instrument that is not carefully validated. As was shown in Section 1.6, the use of self-assessments is common in studies in which primary literature courses are evaluated. Our study suggests that self-assessments are limited in what they can tell about the progression of students.

Our study indicates that students adopt different reading behaviors during the module. There is an increase in the number of students who stated that they read their article non-sequentially. Compared to the beginning of the module, students say they read the Abstract, Method, and Results sections less intensively, which indicates that they have become more selective in their reading. These two findings suggest that students' reading has become more goal-directed; thus the students have adopted more expert-like reading behaviors. Previous studies (Bazerman, 1988; Berkenkotter & Huckin, 1995) have shown that experts' reading behaviors are strongly influenced by the goal with which they read a research article.

It could be argued that students performed better on the post-test because their knowledge of the concepts discussed in the articles has increased. This is unlikely, since the subject of the pre-test and post-test 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, since there were six weeks between the pre-test and post-test.

Our two derived design principles (authenticity and interaction) were based on our second design principle: cognitive apprenticeship. Following this approach, we used modeling, coaching, scaffolding, articulation, and reflection as teaching methods in our module. We did not use exploration (pushing students to explore problems or questions on their own) in our teaching strategy, because we think that exploration should be implemented when students are further in their academic development. The tutors indicated that they had difficulty with modeling. This will be further explored in the next chapter, where we will analyze recordings of the tutor group meetings. Although the research articles used in the module were authentic (i.e. they were not translated or edited), the goal with which students read the articles was not authentic. Students' reading is assignment-driven, while experts read because they want to incorporate the information into their research. In the setting of an undergraduate course, it is difficult to bridge this difference.

Our students were non-native speakers of English. As explained in Chapter 2, we have no reason to suspect that this influenced our results in a major way.

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

This study demonstrates how ideas from the field of genre analysis can be used to improve undergraduates' reading strategies. Our central design principle, the focus on rhetorical structure, may be an effective method for introducing novice readers to primary literature. Ultimately, we expect that the strategy that students developed during our teaching strategy – reading by focusing on the different rhetorical moves – will become automated and more like a skill when they gain more reading experience (see Section 1.7).

Chapter 5
Features Used by Students to Identify Rhetorical Moves

5.1 Introduction

To help undergraduate life science students with reading research articles, we developed a teaching strategy for this purpose. This teaching strategy was based on the Scientific Argumentation Model (SAM), a heuristic which development we described in Chapter 3. Students were expected to follow a strategy that focused on the identification of rhetorical moves. In a previous study (Chapter 4), we implemented this teaching strategy in the module Reading Research Articles and determined with a pre-test and post-test that students' ability to identify rhetorical moves had been improved. However, students had still difficulty with identifying main conclusions, supports and counterarguments (we did not study their ability to identify refutations). Students also stated that they finished the task (reading the article and making the test) in less time and read in a more selective and non-sequential way, which suggested to us that they adopted more expert-like reading behaviors.

In this chapter, we want to investigate how certain features of sentences influence their identification as rhetorical moves by undergraduate life science students. Some studies describe the identification of moves in research articles (i.e. determining the boundaries between moves), although not by novice readers. These are (a) genre analysis studies and (b) information retrieval studies. In genre analysis studies, moves are identified in order to characterize texts from a certain corpus (e.g. a selection of biochemistry articles). This is usually done by coding text fragments in a qualitative way (Holmes, 1997). As Paltridge (1994) wrote, boundaries between moves are usually determined by cognitive criteria (based on content) rather than linguistic ones. This is sometimes difficult: Hunston (1989) remarked that boundaries between moves can be fuzzy edged because of poor writing or because the writer wants to make the prose more flowing. In information retrieval studies, researchers try to develop software that can automatically retrieve information from a certain text (e.g. for automatic summarization). In these studies, moves are usually identified by linguistic means, for example by analyzing lexical cues (e.g. Teufel, 1999). In contrast to genre analysis studies, the content of texts is not qualitatively analyzed.

In our previous research (Chapters 2 and 4) we made the assumption that students' ability to identify a move relies on certain features of this move. We distinguished three types of move features: content-based, organizational, and lexical features. Content-based features are (after Paltridge, 1994) cognitive: these features relate to the function of a move and may be similar to the descriptions of the seven moves mentioned in Section 3.4. Organizational features describe the location of the move in the research article. For example, some moves (e.g. the objective) can always be found in the Introduction. Lexical features are words/phrases that may trigger the reader to identify a certain statement as move. These words/phrases are often a form of metadiscourse: devices that help the reader to "form a convincing and coherent text" (Hyland, 1998a, p. 442). Examples of metadiscourse are words like *in addition, finally, noted above, according to,* and *namely*. Another type of lexical features are reporting verbs like *suggest* or *show* that may signal a conclusion (Bloch, 2010). Lexical features may also consist of words that are specifically associated with a certain move. For example, *it is still unknown* is a phrase indicating a motive.

5.2 Research Question

In our study, undergraduate life science students who followed our teaching strategy were assigned to identify rhetorical moves in research articles. Students received feedback on their answers to the assignments during tutor group meetings. In contrast to our preliminary investigation (Chapter 2), we gave students guidelines that helped them with the identification of rhetorical moves (see Appendix 1). By analyzing students' answers and analyzing the tutor group meetings, we want to answer the following research question:

- Which features (content-based, organizational, lexical) do undergraduate life science students use when identifying rhetorical moves in a research article?

With this information, we want to make suggestions about how to further improve our teaching strategy. For example, we found in Chapter 4 that students had difficulty with identifying main conclusions, supports, and counterarguments. Knowing how students identify rhetorical moves may help us understand why this is the case and provide us with clues for improving our teaching strategy.

5.3 Method

Educational Setting

The research took place in 2009 during the implementation of the module Reading Research Articles, which lasted eight weeks and was part of a course called Biomedical Research. During this module, students made homework assignments in which they – among other things – identified rhetorical moves in research articles. In subsequent tutor group meetings they discussed their answers.

Each week students received instructions from the tutor, and at the end of the meeting a new research article (the "article of the week") and an accompanying assignment the students had to complete at home. In total, students received six research articles and six assignments during the module (excluding the pre-test and post-test in respectively Week 1 and Week 8, which results are described in Chapter 4).

Central part of the assignments was the identification of the seven rhetorical moves that were part of our heuristic, SAM (see Chapter 3): motive, objective, main conclusion, implication, supports, counterarguments, and refutations. Before the assignment, students were instructed by the tutor how to identify these moves. For example, students were instructed in Week 2 how to identify the motive and objective. Right afterwards, the students received the assignment and a research article. In the next week (Week 3), during a tutor group meeting, the article and the assignment were discussed. Students sent their answers by e-mail to the tutor a few days before the meeting. After the meeting, students received written feedback from the tutor on their answers.

In Table 5.1, the tutor group meetings' instructions, assignments, and research articles of Week 2 until Week 6 are shown. The number of words, tables, and figures and the Flesch Reading Ease Score (see Section 2.3) of each article are also included in this table. The assignment and research article of Week 7 will be discussed in Chapter 6. We left out the assignment and article that were discussed in Week 2. For this assignment, students only studied the general structure of research articles and not specific rhetorical moves. Further details about the module are given in Section 4.4.

We asked the tutors to instruct the students by:

- Discussing the information sheets (Appendix 1) in which we listed the seven rhetorical moves of a research article, together with definitions and examples taken from authentic research articles. In these information sheets, we emphasized especially content-based features of these seven rhetorical moves. Lexical features can be rather variable, so we only explicitly mentioned these features for the main conclusion and counterarguments. Organizational features were limited to a specification in which section or sections a rhetorical move can be found;

- Showing as much as possible how they read a research article themselves when giving instructions. For this purpose, tutors were instructed to use each week the sample article: "Effects of Rat Sinoaortic Denervation on the Vagal Responsiveness and Expression of Muscarinic Acetylcholine Receptors" by da Silva Soares et al. (2006). For example, when a tutor explained how to identify the motive, he/she demonstrated how he/she identifies this particular move in the sample article;
- Giving feedback to students. Students received oral feedback when discussing the assignments during the tutor group meetings. Tutors were free in deciding how to give this feedback. After each meeting, students received also written feedback from the tutor on their answers via e-mail or on paper. The feedback tutors gave was based on answer models we developed.

Participants

For this chapter, we chose to closely observe one randomly selected tutor group. This tutor group consisted of ten undergraduate life science students: Amber, Chris, David, Nate, Olivia, Claire, Ruth, Brenda, Lisa, and Maggie (pseudonyms). The students were approximately 18–20 years old. Before the module, Brenda and Maggie stated that they had never read a research article. Claire, Amber, Olivia, and Ruth stated that they had read 1–3 research articles. Chris and Nate read 4–6 research articles, David read 10–12 research articles, and Lisa (who also studied psychology) read more than 15 research articles.

The tutor was a female third-year undergraduate life science student who followed the course herself two years before. This was the second time she acted as tutor in the course. Before the module there was a meeting in which she and the other tutors received information from me about the module's teaching strategy.

Data Collection and Analysis

In this chapter we analyzed two data sources: students' written answers to the weekly module assignments (Weeks 3, 4, 5, and 6) and the audio recordings of the tutor group meetings that took place in Weeks 2, 3, 4, 5, and 6. For these assignments, students identified rhetorical moves in a set of different research articles (see Table 5.1). We answered our research question by looking at how the differences in rhetorical move features influence their identification. The tutor group meetings provided us with additional information about students' choices when identifying moves.

Table 5.1. Tutor group meetings' instructions, assignments and information about the research articles of weeks 2 until 6.

Week	Instructions given about the identification of a…	Students handed in assignment about the identification of a…	Article used for assignment:	Content of article	Flesch Reading Ease Score
2	Motive, objective	-	-	-	-
3	Main conclusion, implication	Motive, objective	Article 1: "Time Course of Echocardiographic and Electrocardiographic Parameters in Myocardial Infarct in Rats" by Miranda et al. (2007).	4,000 words, one table, two figures	62 (*standard readability*)
4	Support, counterargument, and refutation	Motive, objective, main conclusion, implication	Article 2: "Decreased Heart Rate Variability and Its Association with Increased Mortality After Acute Myocardial Infarction" by Kleiger, Miller, Bigger, and Moss (1987).	5,000 words, six tables, five figures	27 (*very difficult to read*)
5	-	Motive, objective, main conclusion, implication, supports, counterarguments and refutations	Article 3: "Delayed Erythropoietin Therapy Reduces Post-MI Cardiac Remodeling Only at a Dose That Mobilizes Endothelial Progenitor Cells" by Prunier et al. (2007).	6,000 words, three tables, six figures	47 (*difficult to read*)
6	-	Motive, objective, main conclusion, implication, supports, counterarguments and refutations	Article 4: "Angiotensin II–mediated Hypertension in the Rat Increases Vascular Superoxide Production via Membrane NADH/NADPH Oxidase Activation" by Rajagopalan et al. (1996).	4,000 words, four tables, seven figures	35 (*difficult to read*)

Assignments – data collection. Each week, students sent their written answers to the assignments to the tutor via e-mail. She forwarded these answers to us. Because we did not collect all answers of Amber, David, Ruth, Lisa, and Maggie, we analyzed only the written answers of the other five students (Chris, Nate, Olivia, Claire, and Brenda). The assignments were formulated in this way: “What is/are in the article the [motive(s)/objective(s)/main conclusion(s)/implication(s)/counterargument(s)/refutation(s)]?” In the case of the supports of Article 3, the question was formulated as follows: “Which supports from the Results and Discussions sections are used to justify the main conclusion(s)? Try to depict the supports in at least three support chains[1] (an example of a support chain can be found in the information sheets[2]).” In the case of the supports of Article 4, the question was formulated in this way: “Which supports from the Results and Discussions sections are used to justify the main conclusion(s)? Try to depict the supports in at least two support chains (an example of a support chain can be found in the information sheets).” Students identified moves explicitly (by quoting a certain text fragment) or implicitly (by paraphrasing a certain text fragment).

Assignments – data analysis. I identified the different rhetorical moves (motive, objective, main conclusion, implication, supports, counterarguments, and refutations) in each research article. I was mostly led by content-based features when identifying moves (following Paltridge, 1994). A second researcher checked my analysis and after a discussion we agreed upon which text fragments represented certain moves in each article. After that, I determined the lexical, organizational, and content-based features of these text fragments. The second researcher also checked these features. Then, I compared students' answers with the moves we identified and used the move features I determined to explain similarities and differences between my analysis and the students' answers. In this way I could determine if students used the presence of a certain feature in a text fragment to identify a move.

Tutor group meetings. I attended all tutor meetings of this group and audio recorded all sessions. I was not a passive observer during the meetings. For example, if the tutor did not know the meaning of a certain technical term, I tried to provide an answer. However, I did not directly influence the topics that were discussed or the pedagogical approach of the tutor.

For the analysis of the audio recordings, I marked episodes in which the students and/or tutor discussed:

- Instructions of the tutor regarding the identification of rhetorical moves;

[1] These support chains are an important feature of SAM (see Chapter 3). A thorough analysis of students' supports chains is presented in Chapter 6.

[2] The information sheets can be found in Appendix 1.

- Remarks of students regarding the identification of rhetorical moves;
- The perceived difficulty of the articles and the assignments.

A second researcher checked if all relevant episodes were marked. Then, verbatim transcriptions were made of the marked episodes. The episodes we describe in the Results section will contain utterances of all ten students.

5.4 Results

We will begin by presenting remarks made by students during the tutor group meetings about the perceived difficulty of the four different research articles. Then, we will show for each move the tutor's instruction, a description of the moves' features in the four articles, students' written answers, and selected episodes from the tutor group meetings.

Perceived Difficulty of Articles

In Week 3, when Article 1 was discussed during the tutor group meeting, Olivia said that she found the Results section very long. Nate added that it was not very clear from the Method section what exactly the experimental groups were. He only found this out when he read the Results section.

In Week 4, David stated that he found Article 2 "relatively superficial." Ruth said that the Discussion section was difficult. In contrast, Olivia found the Introduction section difficult, because of its shortness. Later on, Nate remarked that he liked the shortness of the Introduction section, because usually only the last two paragraphs are important. In Week 5, Nate stated about Article 3 that "some things were difficult" in the article, but he was not able to remember which things. Amber added: "I found the results not really (…) clarifying. [It was not easy] to identify the main points." Article 4, which was discussed in Week 6, gave rise to a number of negative comments from students. They said that they found the article rather difficult. Nate qualified the article as "horrible" and said that he spent three and a half hours on the assignment. Olivia complained that the Method and Discussion sections were too lengthy.

At the end of the module, students had to fill out a questionnaire in which we asked them if they had any remarks about the choice of research articles. Lisa stated that she found Article 4 too difficult and Brenda found Article 3 too long and too difficult. The other students had no remarks about the choice of articles.

So, students stated difficulties with all four articles – and especially their Results and Discussion sections – during the tutor group meetings. However, at the end of the module, only two students were negative about the choice of articles. This suggests to us that the articles were not too difficult for students, so their content probably did not influence students' performances in a major way.

Identification of Motives

Instruction by tutor. In Week 2, the tutor explained to the students how to identify the motive. For this purpose, she let them read the information sheets (see Appendix 1) and then asked them to identify the motive in the sample article. David answered this question by paraphrasing the following sentence from the sample article (located in the last paragraph of the Introduction section):

> Although many studies have analyzed the role of the sympathetic system on the hemodynamic changes after SAD, little is known about the parasympathetic efferent effect on the heart in this model. (p. 332)[3]

Then, Nate chimed in by noting that the word "although" was also mentioned in one of the information sheets' examples of motives (a lexical feature). The tutor acknowledged this and stated that David's answer was correct. She emphasized that in this case the motive indicates that there is something unknown (a content-based feature):

> *Tutor:* *Indeed. The motive is: we know much about the [sympathetic] part of the regulation of the heart (…) but there is not much known about the parasympathetic part, so that is what we want to investigate.*

Analysis of the article's motives. The motives of the four articles are shown in Table 5.2. The motives of Articles 1, 3, and 4 contain pronounced lexical features such as "still a challenge" or "was not clear." Their organizational features are also pronounced, as all three motives are located in the Introduction section, right before the objective. Regarding their content, all three motives describe a problem or state that something is unknown. In contrast, the motive of Article 2 was only implicitly mentioned in the Discussion section (there are new methods available to measure HR variability, which could be used as a predictor of mortality), so there was no specific text fragment that could be

[3] Dialogue fragments from the tutor group meetings are indented and in italic. Text fragments from the article are indented and contain a page number. Students' answers are indented and end with the name of the student between brackets.

Table 5.2. Lexical, organizational, and content-based features of the motives in Articles 1, 2, 3, and 4.

	Text fragment	Lexical	Organizational	Content-based
Article 1	However, after 50 years and even with the more recent use of imaging techniques such as echocardiography (…), at present the evaluation of myocardial infarct size in vivo and its relation to the actual lesion found in *post-mortem* histopathology is still a challenge. (p. 640)	"However," "still a challenge"	Placed in the penultimate paragraph of the Introduction, right before the objective	Describes a problem
Article 2	-	-	-	-
Article 3	Cardioprotective effects have been shown with high doses of EPO injected during the first hour after ischemic injury, but the effect of chronic treatment beginning a few days after coronary artery occlusion, before the reperfused myocardium has healed, is unknown. (p. 522)	"is unknown"	Placed in the last paragraph of the Introduction, right before the objective	There is something unknown
Article 4	These previous studies were performed in cultured vascular smooth muscle cells after several passages. Such cells undergo transformation to a synthetic phenotype and exhibit signaling responses to angiotensin II which may differ from responses of the vascular smooth muscle in vivo. Because of these issues, it was not clear that a similar effect could be produced by modest levels of angiotensin II in vivo. (p. 1916)	"was not clear"	Placed in the last paragraph of the Introduction, right before the objective	There is something unknown

marked as motive. The content of this motive does not fall into one of our three categories (see Appendix 1).

Students' answers. Table 5.3 contains our assessments of students' written answers to the assignments. In Article 1, all five students except Claire identified the motive. Claire's motive was not ostentatiously wrong, but lacked specificity:

> There is already something known about cardiovascular diseases, but there is also a lot unknown. In this article there is more research done on this. (Claire)

The motive of Article 2 was not mentioned in the Introduction section. Still, all students wrote down a motive. Brenda, Olivia, and Nate wrote down a motive derived from the text. For example, Brenda answered that HR variability was measured in the past for only 60 seconds; now they want to predict mortality with a 24 hours measurement. Chris formulated the motive as an objective, because it describes what authors want to know – and not why they want to know it:

> Determining the best predictor of mortality. (Chris)

Claire's answer was more a description of a possible implication and is not in accordance with the article's content:

> If you know the relation between HR variability and mortality than it is maybe possible to ensure that people with a myocardial infarction who have a low HR variability will get a higher HR variability, so that they're chance of survival will improve. (Claire)

Table 5.3. Assessment of students' answers regarding the identification of the different moves in the four articles. Answers could be in accordance with our own analysis (✓), not in accordance with our own analysis (×), or incomplete/non-specific (±). The hyphens indicate that there was no motive mentioned in Article 2.

	Article 1					Article 2					Article 3					Article 4				
Student	Chris	Nate	Olivia	Claire	Brenda	Chris	Nate	Olivia	Claire	Brenda	Chris	Nate	Olivia	Claire	Brenda	Chris	Nate	Olivia	Claire	Brenda
Motive	✓	✓	✓	±	✓	-	-	-	-	-	✓	✓	✓	✓	✓	×	✓	✓	✓	✓
Objective	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	±
Main conclusion						✓	×	±	✓	±	✓	✓	✓	✓	✓	±	±	±	±	±
Implication						✓	✓	✓	×	✓	×	×	×	✓	✓	±	✓	✓	±	✓
Supports											±	±	✓	✓	±	✓	±	✓	✓	±
Counter-arguments and refutations											✓	±	✓	±	✓	✓	×	✓	✓	×

The motive of Article 3 was found by all five students. The motive of Article 4 was only incorrectly identified by Chris, who paraphrased a text fragment in the Introduction section stating that

> [the pathways which lead to the increased $\cdot O_2^-$ production] remain undefined, but likely involve stimulation of phospholipase A_2 and the release of aracidonic [*sic*] acid. (p. 1916)

Although this sentence contains a lexical feature typical of motives ("remain undefined"), we did not identify it as a motive, as the authors did not study this pathway.

Tutor group meetings. During the tutor group meeting about Article 1, Amber stated that she had some difficulties with distinguishing the motive from the objective. David also found it difficult to find the motive. Later on, the tutor asked Amber to describe the problem that the authors want to solve (i.e. the motive):

> *Tutor:* *Why do they do this study?*
>
> *Amber:* *Yes. They want to determine the size of the myocardial infarction… that's better… [inaudible]*
>
> *Tutor:* *Yes, indeed. It is still difficult to determine with non-invasive methods the size of the infarction (…) that's the gap in knowledge.*

So, the tutor emphasized that the motive states that there is something unknown – a gap in knowledge (a content-based feature). Then, the tutor asked Brenda to identify the motive in the text. She identified the correct text fragment.

During the next tutor group meeting, about Article 2, the tutor established that some students had difficulty with identifying the motive. She asked the group to find the motive and then asked Nate to read out his answer. According to the tutor, Nate's answer – a motive derived from the text – was correct. Then, she emphasized the content-based features of the motive and objective:

> *Tutor:* *What is precisely the difference between the motive and objective? Could you explain this?*
>
> *Amber:* *Objective is what they want to do… Or not? And the motive is why they want to do the study.*
>
> *Tutor:* *Yes. Precisely. Indeed. So, the objective is in fact the*

way in which the study was done and the motive is indeed why.

Later on, she recapitulated her point about the content-based difference between motive and objective:

Tutor: *So, the motive is why you want to do the study: the gap in knowledge, what's missing. And the objective is: how you are going to do it and what they want to achieve with the study.*

During the next two tutor group meetings, no episodes occurred in which the motive was discussed, although the tutor did inventory students' answers.

Interpretation. The motives of Articles 1, 3, and 4 have all pronounced lexical, organizational, and content-based features. This is probably why students were quite successful in identifying the motives in these articles. Although Article 2 does not contain an explicit motive, three of the five students derived the authors' motive from the text. This suggests that students are able to use content-based features to distill an implicit rhetorical move from a text. The tutor described the objective as "how you are going to do it." This is different from the description given in the information sheets. Interestingly, the tutor emphasized the content-based features of the motive during the discussion of the information sheets, while Nate immediately focused on lexical features.

Identification of Objectives

Instruction by tutor. Besides the motive, the tutor also explained in Week 2 how to identify the objective. She let the students read the information sheets. One student had a question about a content-based feature of the objective:

Amber: *How you can apply it. This belongs to the objective? Or not?*

Tutor: *Well… no. (…) What you mean are the implications, which you can find often in the Discussion. (….) It is better to determine the implication afterwards. It is possible that something is mentioned about [possible implications] in the Introduction, but this is not always the case.*

Then, the tutor asked Olivia to identify the objective in the sample article. In her answer, Olivia referred to the last sentence of the last paragraph of the Introduction section (which almost immediately followed after the motive):

> Therefore, the objective of the present study was to analyze the effects of acute (1 day) and chronic (20 days) SAD on the vagal responsiveness and expression of mAChRs in the rat atrium. (p. 332)

The tutor said that it was correct. Then, the tutor summarized the main message of this exercise and emphasized the lexical and organizational features of the objective:

> *Tutor:* *You'll have [in the Introduction] a preamble and then, in the last paragraph, there is often a summary that contains their objective. Often it goes like: 'the aim of our investigation…' or 'therefore…' – so, with these kinds of words and sentences you can recognize quite well what the study's objective is.*

Analysis of articles' objectives. In all four cases the research articles' objectives had pronounced lexical features (Table 5.4). All objectives contained words such as "specific purpose" or "to gain insight." Furthermore, the objectives were all placed at the end of the Introduction section (organizational feature). Finally, all four objectives contained a research aim. In the case of Article 1, the objective consisted of a research aim *and* research questions, as the aim of the study was to answer three research questions.

Students' answers. As Table 5.3 shows, only Brenda did not identify all objectives. She identified the objectives of Articles 1, 2, and 3, but only wrote down three of the four objectives that were presented in Article 4. The other students were able to find all the objectives in the four research articles.

Tutor group meetings. During the tutor group meetings, no episode occurred in which the objective was discussed, except when the tutor inventoried students' answers and when she explained the difference between objective and motive (see above).

Interpretation. All students identified the objectives of the four articles, except for Brenda who gave an incomplete answer for Article 4. Students' suc-

Table 5.4. Lexical, organizational, and content-based features of the objectives in Articles 1, 2, 3, and 4.

	Text fragment	Lexical	Organizational	Content-based
Article 1	The specific purpose of the current study was to address the following questions: Can the routine electrocardiogram (ECG) and/or the echocardiogram (ECHO) predict adequately the extension of the infarct in the rat? How do these exams change with time after infarction? Is there a correlation between infarct size measured by histopathology and alterations in these exams? (p. 640)	"specific purpose," "following questions"	Placed in the last paragraph of the Introduction, right after the motive	Research aim/ Research question
Article 2	…to gain insight into the relations of these variables and HR variability to mortality. (p. 256)	"to gain insight"	Placed at the end of the Introduction (which consists of one paragraph)	Research aim
Article 3	Therefore, the aim of this study was to determine whether chronic treatment with clinically relevant doses of EPO beginning 1 wk after I/R reduces postinfarction cardiac remodeling. (p. 522)	"Therefore, the aim"	Placed in the last paragraph of the Introduction, right after the motive	Research aim
Article 4	In the present experiments, we examined the effect of angiotensin II–induced hypertension on vascular $\cdot O_2^-$ production [Objective 1] and attempted to characterize the oxidase activated [Objective 2]. Studies were also performed to determine if another form of pharmacologically induced hypertension, that due to the infusion of NE, had similar effects on vascular $\cdot O_2^-$ production [Objective 3]. Finally, we sought to determine if this alteration in vascular $\cdot O_2^-$ production had any impact on endothelial regulation of vasomotion [Objective 4]. (p. 1916)	"we examined," "attempted to characterize," "to determine," "to determine"	Placed in the last paragraph of the Introduction, right after the motive	Research aim

cessful identification of objectives is probably due to their pronounced lexical, organizational, and content-based features. The tutor did not discuss the criteria for the identification of the objective with the students. As a result, we cannot use this data to infer on which features students rely the most.

Identification of Main Conclusions

Instruction by tutor. In Week 3, at the end of the meeting, instructions were given about the identification of the main conclusion. First, the tutor asked the students to read the information sheets (Appendix 1). Later, the tutor discussed with them the organizational features of the main conclusion:

Tutor:	*If you're searching for the main conclusion: where would you look?*
Amber:	*In the Discussion section.*
Tutor:	*Is there a location in the Discussion section where you would look?*
David:	*It differs, I believe.*
Nate:	*In [the sample article] it has its own heading, but… I believe that in the previous one [Article 1]… in the beginning a small summary was given and at the end the main conclusion.*
Tutor:	*Yes, indeed. It is… Look, if you write a long article, you want to emphasize your most important message… so, you are closing the article with it. So, often… you can find the most important message in the last paragraph, because that's the information that has to be remembered.*

The students were assigned by the tutor to identify the main conclusion in the sample article individually. Nate was asked to give the main conclusion. He read out the following sentence:

> The present results provide evidence of the effects of SAD on mAChRs and parasympathetic function involved in regulating cardiac function. (p. 336)

The tutor said that this was correct. She explained to the students that this main conclusion could clearly be related to the objective of the study (a content-based feature). Then, Brenda asked a question about one of the content-based features of the main conclusion, which was mentioned in the information sheets (the authors decide if the hypothesis is correct or not, see Appendix 1):

Brenda:	*But here it's stated (…): they can also decide that the hypothesis was not correct. But this can never be found in a research article.*
Tutor:	*How do you mean?*
Brenda:	*Well, you (…) have a study and then you'll find out that the hypothesis is not correct. Then you'll do a new study. And you'll publish if you have a hypothesis that is correct.*

The tutor told Brenda that researchers sometimes publish a paper about a rejected hypothesis because they were not able to gather positive results and feel pressured to publish something. I added that publishing a rejected hypothesis may occur when the hypothesis is controversial.

Analysis of articles' main conclusions. In our analysis, we selected as the main conclusion the conclusion that related directly to the objective (content-based feature). The main conclusions of Articles 2, 3, and 4 are shown in Table 5.5. All three main conclusions – in contrast to what the tutor said during the instruction – are placed in the first paragraph of the Discussion section (organizational feature) and describe how the objectives are met (content-based feature). Furthermore, they contain words such as "we found" and "in the present study" (lexical features). Identifying the main conclusion was sometimes difficult, because all articles contain multiple conclusions.

Students' answers. Only Chris and Claire identified the main conclusion of Article 2 (Table 5.3). In contrast, Olivia and Brenda both paraphrased the last sentence of the Results section, in which the most important findings of the study are summarized:

> Decreased HR variability increases the risk of death irrespective of average HR (Fig. 2), variables reflecting left ventricular function (Fig. 3), those measuring ventricular ectopic activity (Fig. 4 and 5), clinical or demographic variables (Table V) or drug treatment (Table VI). (p. 260)

We did not designate this sentence as the main conclusion, as it is located in the Results section and is used by the authors as a support. However, the sentence does relate to the objective of the study (gaining insight into the relation between HR variability and mortality), so the answer

Table 5.5. Lexical, organizational, and content-based features of the main conclusions in Articles 2, 3, and 4.

	Text fragment	Lexical	Organizational	Content-based
Article 2	In the present study, HR variability (…) was significantly and strongly associated with subsequent mortality. (p. 260)	"In the present study"	Placed in the first paragraph of the Discussion	Description of how the aims were met
Article 3	These results show that chronic EPO treatment beginning 7 days after MI reperfusion in rats attenuates cardiac remodeling and improves cardiac function. (p. 526)	"These results show"	Placed in the first paragraph of the Discussion	Description of how the aims were met
Article 4	In these studies, we found that angiotensin II–induced hypertension is associated with increased vascular $\cdot O_2^-$ production [relates to Objective 1] and impaired vascular relaxations to acetylcholine, the calcium ionophore A23187, and nitroglycerin. In studies of vascular homogenates, the predominant source of this increased $\cdot O_2^-$ production seemed to be membrane-bound vascular NADH and NADPH oxidases [relates to Objective 2]. The alteration of vascular relaxations to endogenous and exogenous nitric oxide was likely, at least in part, due to the increase in vascular $\cdot O_2^-$ production, as it was partially corrected by augmenting vascular superoxide levels with liposome-entrapped superoxide dismutase [relates to Objective 4]. In contrast to the effect of angiotensin II infusion, NE infusion, which produced a similar degree of hypertension, did not increase vascular $\cdot O_2^-$ production and did not alter endothelium-dependent vascular relaxation [relates to Objective 3]. (p. 1920)	"In these studies, we found"	Placed in the first paragraph of the Discussion	Description of how the aims were met

was not completely wrong. Nate paraphrased the first sentence of the last paragraph of the Discussion section as main conclusion (which was titled "Therapeutic implications"). This was not the main conclusion, as it does not relate directly to the study's objective:

> The present study suggests that patients with decreased HR variability have decreased vagal tone or increased sympathetic tone and may have higher risk of ventricular fibrillation. (p. 261)

All students found the main conclusion of Article 3. The authors of Article 3 also presented an additional conclusion that provides an explanation for their results:

> These effects occurred only with an EPO dose that induced EPC mobilization in blood and increased capillary density in the MI border zone. The improvement in cardiac contractility was clearly related to EPC mobilization. (p. 526)

All students except Olivia also added the additional conclusion to their main conclusion. We did not mark answers in which this additional conclusion was included as incorrect, even though this conclusion is not directly related to the aim of study.

The main conclusion (or conclusions) of Article 4 was rather lengthy, because the study had four objectives. None of the five students did identify the complete main conclusion, despite the fact that – with the exception of Brenda – the students identified all objectives. Nate only incorporated the parts of this section that related to the first and second objectives. Olivia incorporated only the part that related to the first objective. Brenda incorporated in her answer only the part that related to the second objective and additionally paraphrased the following sentence – in which the authors compare their results with other studies – from the Discussion section:

> The present studies add to these recent observations by demonstrating that angiotensin II can exert this effect in vivo, and that this increase in $\cdot O_2^-$ production may contribute to alterations in endothelium-dependent vascular relaxation and responses to exogenous nitrovasodilators in the intact vessel. (p. 1920)

Chris also referred to this sentence. Claire's answer was also incomplete. She paraphrased the following sentence from the Abstract (probably because of the phrase "we conclude"):

> We conclude that forms of hypertension associated with elevated circulating levels of angiotensin II may have unique vascular effects not shared by other forms of hypertension because they increase vascular smooth muscle $\cdot O_2^-$ production via NADH/NADPH oxidase activation. (p. 1916)

This sentence does represent a conclusion, but because it was located in the abstract, we did not include it in Table 5.5.

Tutor group meetings. In Week 4, when Article 2 was discussed, much attention was paid to the main conclusion. The tutor asked the group to identify the main conclusion. Maggie answers this question:

Maggie:	*Well, I believe it's this one: their strongest association was the one between heart rate variability and ehm… yes, mortality.*
Tutor:	*Yes, indeed. Correct. (…) There were many things examined. And the correlation between heart rate variability and the mortality was the strongest of all things they measured. So, it was the best predictor. And there are a lot of other conclusions mentioned, but the main conclusion is in fact very short.*

So, the tutor made clear to the students that not all conclusions are main conclusions and that the main conclusion should contain information about the relation between heart rate variability and mortality, as this was the objective of the study (a content-based feature). Olivia asked if the sentence she identified as main conclusion (the last sentence of the Results section) was correct. The tutor indirectly answered this question by mentioning a different sentence as the correct main conclusion, which is located in the Discussion section and not in the Results section (organizational feature). Brenda raised the point that the article contained several text fragments conveying more or less the same message as the main conclusion:

Brenda:	*But actually it's a little bit the case… actually everything amounts to the same thing. Because it was almost in every conclusion mentioned that… well, yes… the HRV has simply the highest association.*
Tutor:	*Yes. Yes, correct. But if you, for example, wrote down as main conclusion: the heart rate variability has the biggest effect on mortality, independent of the ejection fraction… that's not a main conclusion, it's a sub-conclusion based on one graph.*

Brenda replied that her main conclusion (the last sentence of the Results section; she gave the same answer as Olivia did) summarized all the different graphs. This prompted the tutor to say that she would not mark her answer as incorrect. Amber stated that she did not think that the sentence read aloud by the tutor was the main conclusion, since this sentence was located in a paragraph containing a reference to a previous study. The tutor told her that the main conclusion begins with "In the present study" (a lexical feature). Amber realized her mistake and replied: "Oh, yes! Of course! I must have skipped this." Then, David stated that he thought that the first sentence of the last paragraph of the Discussion section was the main conclusion.

Amber: *Yes, that's what I've got here, too.*

Nate: *Yes. That one does not match the objective.*

David: *No, okay.*

Nate: *That's what I found out later on.*

Tutor: *Yes, that's indeed also a way of checking if you got the right conclusion: to determine if it matches the motive and objective. If it matches… well, yes… then you are in the right area.*

So, Nate and the tutor were reminding the rest of the group of a content-based feature of the conclusion. Lisa told the group that she has a different method for identifying the main conclusion:

Lisa: *What I've done is simply answering the objective by myself, after reading. I didn't look in the text where the main conclusion was located.*

Tutor: *No, but that's not obligatory. You can also simply… This is, let's say, a sentence that coincidentally mentions it. But you can also formulate it by yourself, of course.*

It seemed that she was the only student in the group who tried to interpret the article's data by herself, independently of the authors' interpretation. Then, Olivia had a question about the organizational features of the main conclusion:

Olivia: *I tried to find the [main] conclusion in the Results. (…) But that's not correct? It has to come from the Discussion?*

Tutor: *The Discussion is convenient because things will be interpreted for you, so to speak.*

Olivia: *Yes.*

Tutor: *But if you're very well able to interpret the conclu-*

	sions from the results, then it's better to do this first, before you read the Discussion. And then you're less prejudiced.
Olivia:	*Yes.*
Tutor:	*But I would certainly read the Discussion. Because that's where points of criticism are mentioned and where all the results and conclusions are summarized.*

The group discussed the sentence of the last paragraph of the Discussion section (which was identified as main conclusion by Nate). This discussion was instigated by Brenda, who doubted that this sentence described the authors' own results – despite the fact that it begins with "The present study suggests." The tutor was also puzzled by this. She remarked that the information that is presented could not be found in any of the graphs.

Ruth:	*It seems like they just come up with a reason why it could be the case.*
Tutor:	*Yes, indeed. Like this… If you read it in this way, you don't get the feeling that they directly studied this in their own research. But that they use previous studies as justification.*

The students looked for clues which could indicate if the information comes from the authors' own research. Lisa pointed out that the statement is followed by a sentence beginning with "these speculations." The tutor concluded: "This is an example of authors who present their own results in a flattering way. (…) It is stated 'the present study suggests,' but as far as we could see, they did not study this directly."

In contrast to Week 4, the main conclusion was not discussed in Week 5 (probably because all students identified the correct main conclusion). In Week 6, the tutor asked David to read out the text fragment he identified as main conclusion in Article 4. After David read out his answer, the tutor explained that the whole first paragraph of the Discussion section forms the main conclusion. "So much?" Olivia asked. "Yes," said the tutor, "but it was, of course, a large article with a lot of things [that the authors] did study."

Interpretation. Even though the main conclusions of Articles 2, 3, and 4 have pronounced lexical, organizational, and content-based features, students had in certain cases difficulty with identifying this move. Only two of the five students identified the main conclusion of Article 2. One student identified the implication as main conclusion and two students identified a sentence from the Results section. As Brenda stated, this sentence summarized all the different graphs, representing a preliminary conclusion and not the main conclusion. At that time, she did not make use of one of the organizational features of the main conclusion (it is located in the Discussion section) that was mentioned in the information sheets. Article 2 gave rise to an interesting discussion in which the students and tutor discovered that a lexical feature ("The present study suggests") in one sentence misrepresented the authors' research as the information in this sentence was not based on the authors' own results.

Students fared better with identifying the main conclusion of Article 3, possibly because the in-depth discussion about the main conclusion of Article 2. None of the students completely identified the complete main conclusion of Article 4, even though they all – with the exception of Brenda – identified the article's four objectives. It seems that students felt that a main conclusion has to be short; they ignored the content-based feature that the main conclusion relates to the objective.

Identification of Implications

Instruction by tutor. In Week 3, the tutor explained the students how to identify the implications in a research article. After students read the information sheets, the tutor asked Brenda to identify the implication in the sample article. She read out the following sentence:

> The functional data (…) may represent a new possibility of management in pathophysiological states associated with baroreflex sensitivity attenuation like heart failure and hypertension in which a clearly parasympathetic impairment may occur. (p. 336)

The tutor acknowledged that this was correct. Students had no questions or remarks about the identification of the implication.

Analysis of articles' implications. All three implications contained lexical features: "implications" (Article 2), "clinical setting" (Article 3), and "future studies" (Article 4). They were all three located in the last paragraph of the Discussion section (organizational feature) (Table 5.6). The implications of Articles 2 and 3 consisted of recommendations for improvements of therapies (content-based feature). The implication of Article 4 consisted of suggestions for future research (content-based feature).

Table 5.6. Lexical, organizational, and content-based features of the implications in Articles 2, 3, and 4.

	Text fragment	Lexical	Organizational	Content-based
Article 2	These speculations have therapeutic implications: agents that blunt sympathetic influence or agents that promote vagal influence may have therapeutic value after infarction… (p. 261)	"implications"	Placed in the last paragraph of the Discussion	Recommendation
Article 3	In the clinical setting, the acute antiapoptotic effects of EPO should add to the chronic effects seen in our experiments. (p. 528)	"clinical setting"	Placed in the last paragraph of the Discussion	Recommendation
Article 4	Future studies in which endothelium-dependent vascular relaxation is examined in humans should take into account the renin/angiotensin II profiles of the subjects examined. Finally, these studies may provide insights into why treatment with angiotensin II-converting enzyme inhibitors or angiotensin II-receptor antagonists may have beneficial effects not encountered with other drugs (37–40). (p. 1922)	"Future studies"	Placed in the last paragraph of the Discussion	Suggestions for future research

Students' answers. All students found the implication of Article 2, except Claire (Table 5.3). Claire paraphrased parts of the penultimate paragraph of the Discussion section. In this paragraph, the authors mention previous studies that discuss possible therapies against ventricular fibrillation. These therapies do not follow from the study's results, so we did not identify them as implications.

The implication of Article 3 was only found by Claire and Brenda. Chris' implication was rather non-specific: "The researchers are doing more research at this moment." Nate wrote that he could not find an implication and listed some ideas for possible new studies. Olivia gave two additional conclusions as implications.

The implication of Article 4, which consisted of two sentences, was found by Nate, Olivia, and Brenda. Chris' answer contained only the first sentence, while Claire's answer contained only the second sentence.

Tutor group meetings. In Week 4, students were still somewhat confused about the content-based features of the implication while discussing Article 2. That is why the tutor explicated a content-based feature of the implication:

Amber:	*The implication is just the use [of the study]?*
Tutor:	*Yes. Often this relates to drugs or therapies that may be developed for… well… humans or society.*
Lisa:	*Is a follow-up study an implication?*
Tutor:	*Yes. Actually everything that can happen with the research in the future.*

In Week 5, while discussing Article 3, students found identifying the implication still difficult:

Tutor:	*What was the study's implication?*
Nate:	*That one was difficult to find.*
Tutor:	*Yes?*
Nate:	*In my view.*
Tutor:	*Yes, it is indeed… it was less clearly laid down compared to the previous articles. That's right.*
David:	*[inaudible]*
Amber:	*I have… let's see… that it can have full effects. Chronic treatment of EPO combined with a standard dose. (…)*
Tutor:	*Yes, correct. That was very good, indeed. The important thing is that acute effects [of EPO] could be combined with chronic effects… for future treatments.*
David:	*[The authors] said in the last part that they are working on this at the moment.*
Tutor:	*Yes. Exactly. A follow-up study is always also an implication, of course.*

So, the tutor emphasized content-based features of the implication (suggestions for further treatment and research). While discussing Article 4 in Week 6, students did not state that they had any problems with identifying the implication:

Tutor:	*And what are subsequently the implications of this article?*
Olivia:	*That it may give insight into [the question] if it has positive effects. (…) That you gain more insight into*

	how [treatment] should happen.
Tutor:	*Indeed, because… can someone delve further into this matter? Because they are talking here about a combination of angiotensin II that leads to oxidative stress, and they say something about this in the implication, actually. (…)*
David:	*With the next study that they are going to do, renin/angiotensin profiles [should be] made of the subjects.*
Tutor:	*Yes, indeed.*

Interpretation. The implications in Articles 2, 3, and 4 have pronounced organizational and content-based features. However, their lexical features are sometimes rather subtle (e.g. "clinical setting" in Article 3). This might explain why students did not always identify the implication. During the tutor group meetings in Weeks 4 and 5, students were confused about the content-based features of the implication.

Identification of Supports

Instruction by tutor. The instructions for identifying the supports were given in Week 4. However, in Week 3 the tutor already made some remarks about the importance of inscriptions (see Section 2.3). During the tutor group meeting in that week she told the students how she reads an article herself:

Tutor:	*The first thing I do is not looking at the text, but at the graphs. (…) And then I read the text for myself. For me that's the best way (…) because you have a picture alongside [the text]. And… graphs show perfectly the important things.*

Nate asked the tutor if authors always mention important information in the figures and tables as well as in the text. In her reply, the tutor referred to pointers (a lexical feature): "Of course. This is always mentioned between brackets: see Figure 2 or see Table 1."

In Week 4, I (the tutor left the meeting early) explained with the use of the sample article how supports in the Results and Discussion sections form chains that lead to the main conclusion. Nate remarked that some supports (such as references to previous studies) may also follow after the main conclusion. Nate explained that this was clearly the case in Article 2. I replied that additional supports such as references to previous studies can follow after the main conclusion, but these are usually less important as justification for the main conclusion than the authors' own supports.

Analysis of articles' supports. Because of the large amount of supports, it was impossible for us to determine which features students used for identification: students' answers were too varied to identify patterns. This is why the supports of Articles 3 and 4 are not shown in a table. Instead, we will give a general description of the supports.

Article 3 contained six different data sources that served as supports: heart function, infarct size, capillary density, CD31, hematocrit, and apoptosis. These data were presented in six figures and three tables. The heart function data was probably the most important, as these demonstrate that rats treated with EPO showed more improvement after a myocardial infarction than rats not treated with EPO.

Article 4 contained six different data sources that served as supports: blood pressure, vascular superoxide production data, oxidase activity, vascular relaxation, effects of superoxide dismutase, and effects of low dose angiotensin II. These data were presented in seven figures and four tables.

Each of these data sources in both research articles can be incorporated in a support chain (The definition of a support chain is given in Section 3.4).

Students' answers. Our assessment of students' answers is shown in Table 5.3. In Week 5, Nate did not present Article 3's supports in chains. He only wrote down three different text fragments from the Results and Discussion sections. Brenda did the same: she wrote down five different (paraphrased) text fragments. In contrast, Chris, Olivia, and Claire presented their supports in chains. Chris wrote down four chains with text fragments. All his text fragments were taken from the Discussion section. His text fragments in the different chains were not always related to each other regarding their content. Olivia and Claire wrote down chains with text fragments (or, as in Claire's case, paraphrased text fragments) from the Results and Discussion sections. Their supports in the chains were also related to each other (which is why they were given a tick in Table 5.3[4]). Claire presented four different chains and Olivia three. With the exception of Chris, all students mentioned the improvement in heart function as a support.

In Week 6, Chris, Olivia and Claire wrote down respectively two, two and three support chains. Their supports in the chains were also related to each other (which is why they were given a tick in Table 5.3). Nate and Brenda just gave a list of supports. None of the students gave supports that

[4] The students' supports were not complete. However, we still gave them a tick as we did not explicitly instruct them to write down all supports.

related to all four main conclusions. For example, Nate only wrote down supports that related to Main conclusion 2 (although he also identified Main conclusion 1). Regarding the source of students' supports, Chris and Nate both referred to supports from the Results and Discussion sections. Olivia's supports exclusively came from the Discussion section. We were not able to say where Claire and Brenda got their supports from, because they paraphrased their answers in such a way that we could not identify the original text fragments.

Tutor group meetings. At the beginning of the meeting in Week 5, Olivia claimed that she found the identification of the supports in Article 3 difficult. Amber concurred with her. However, there was not much discussion during the tutor group meetings about the identification of supports and the construction of support chains. The tutor talked with the students in detail about individual supports (e.g. how to interpret certain graphs or tables), but their relation to each other and to the main conclusion was not discussed.

Something similar happened in Week 6. The tutor did not discuss the chains of supports. However, she said that she would give the students written feedback about this part of the assignment. When the students asked her about the relevancy of finding supports, she answers: "Look, with the chains it is simply the matter that you can link [text fragments from] the Results and Discussion with a figure. Or that you think: it is a figure and it will lead to a certain conclusion that will be taken to the Discussion. And you must be able to link these fragments together." Amber asked the tutor if these fragments are not always placed next to each other. The tutor: "No. It is possible that you have simply a figure and that (…) it will be discussed a little bit in the Results, and then [the authors] will follow it up in a conclusion [in the] Discussion."

Interpretation. Students did not identify all supports that were mentioned in the articles. At the most, they wrote down four different support chains (in Week 5). A reason for this might be that we did not explicitly ask them to identify all supports. Only three of the five students presented their identified supports in chains, even though they were instructed to do this. And if they presented their supports in chains, the supports did not always relate to each other in a correct way. During the tutor meetings, the tutor did discuss individual supports, but not the relations between them.

In the case of Article 3, students identified the most important support: four of the five students mentioned the improvement in heart function. In the case of Article 4, students failed to give supports that related to all four main conclusions. This is not surprising, as none of the students did identify all four main conclusions.

Because of the large number of supports in all research articles, students' answers were very varied. Therefore, we were not able to identify patterns, which could give us insight into students' use of features. Not all supports contain lexical features such as pointers or reporting verbs. This might explain why students have difficulty with identifying this particular move.

Identification of Counterarguments and Refutations

Instruction by tutor. In Week 4, I (the tutor left early) explained to the students how they could identify counterarguments and refutations. First they read the information sheets. David asked if authors will mention a counterargument if they are unable to come up with a refutation. I told him that – from an argumentative viewpoint – it is advisable to anticipate on readers' criticism, even if you cannot provide a refutation. David did not seem to be convinced by this. "But you will undermine your own research," he said. Then, I showed the students how I identified the counterarguments and refutations of the sample article. Amber had a question: is a lack of understanding of a certain mechanism also a counterargument? I answered that not understanding the mechanism behind a certain result does not necessarily imply that the main conclusion – based on this result – is incorrect.

Analysis of articles' counterarguments and refutations. In Article 3, the authors list two counterarguments under the heading "Limits": (A) Only CD31 staining was used to identify EPC's and (B) The beneficial effects of EPO are maybe unrelated to the increase in EPC's. The authors only provide a refutation for counterargument A. In Article 4, the authors mention one counterargument and a corresponding refutation in the Discussion section. In both articles, the counterarguments and refutations lacked pronounced organizational features: their location in the different Discussion sections was very variable. In Article 3 they were placed in the third and second to last paragraphs of the Discussion section and in Article 4 they were placed in the fifth paragraph of the Discussion section. Also, the lexical features of the counterarguments and refutations of Article 4 were quite subtle ("this is unlikely"). However, all counterarguments and refutations do contain rather pronounced content-based features (Table 5.7).

Students' answers. Chris, Olivia, and Brenda listed both counterarguments and the refutation of Article 3 (Table 5.3). Nate only found counterargument B. Claire only wrote down counterargument A (including the refutation). Additionally, all students (with the exception of Chris) identified text fragments as counterarguments and refutations in which the methodology is defended. For example, Nate, Olivia, and Brenda identified the following two sentences as respectively a counterargument and refutation:

> …however, the hematocrit reached 62% 3 wk after the first injection, and blood pressure also increased significantly. Because low EPO doses used to treat renal anemia have been shown to mobilize active EPCs in humans (5) we chose to test doses suitable for chronic treatment (10, 11). (p. 527)

In Article 4, Chris, Olivia and Claire identified the counterargument and refutation. Additionally, Olivia identified a statement in which the authors state that the mechanism behind the increased $\cdot O_2^-$ production remains “poorly defined” (p. 1920). We did not identify this statement as a counterargument, as the authors did not set out to unravel this mechanism. Not knowing the mechanism behind a certain phenomenon does not necessarily diminish the value of the results gathered about this phenomenon. Brenda’s answer consisted of statements (taken from the Discussion section) that described the findings of the study. So, they were not counterarguments but supports. Nate also identified a support as a counterargument. Furthermore, he identified a text fragment that states how the results are compatible with existing theories. Again, this was not a counterargument.

Tutor group meetings. During the discussion of Article 3 in Week 5, the tutor mentioned the counterarguments only briefly. She began by pointing out that the authors did not sufficiently prove their additional conclusion that the positive effects of EPO are due to the proliferation of EPC’s. Together with the students, she tried to think of experiments proving this. She told the students: “Maybe that you could also administer a surplus of EPC’s. Let’s see what the effect is of such action. These are the things you could think of when you want to really test this, because there are a lot of things suggested in this article, but there are no real proofs.” Later, she added: “Another counterargument is – and [the authors] point this out themselves – is that the improved function can also be caused by more oxygen. Because EPO increases the number of red blood cells, of course. And they do not discuss this further.”

During the tutor group meeting in Week 6, the tutor asked Chris to identify the counterargument and refutation in Article 4. The tutor said that his answer was correct. Then, Amber made a remark: “I had as refu-

Table 5.7. Lexical, organizational, and content-based features of the counterarguments and refutations in Articles 3 and 4.

	Text fragment	Lexical	Organizational	Content-based
Article 3	*Limits*. We used only CD31 staining to identify EPCs, whereas detailed phenotypic descriptions of circulating EPCs are based on coexpression of several other endothelial and hematopoietic antigens such as CD34, VE-cadherin, VEGF receptor 2, and Tei2. However, EPC mobilization by EPO has been clearly shown in species other than rats, including humans. (...) Although chronic EPO therapy initiated 7 days after reperfusion improved post-MI remodeling only at a dose capable of mobilizing EPCs and of increasing capillary density, we cannot exclude the possibility that the beneficial effects of EPO in this setting result, at least in part, from improved oxygen delivery to the hypoxic myocardium.	"Limits," "However," "Although"	Placed in the third and second to last paragraph of the Discussion	Methodological problem/Alternative explanation
Article 4	It is possible that other cell types such as monocyte/macrophages or neutrophils contributed to the increase in vascular superoxide production during angiotensin II infusion, but we believe this is unlikely. Previous studies have shown that the degree of inflammatory cell infiltration at these earlier time points of angiotensin II infusion is negligible (23). Further, the preferred substrate for oxidases in macrophages and neutrophils is NADPH rather than NADH. (p. 1920)	"this is unlikely"	Placed in the fifth paragraph of the Discussion	Alternative explanation

tation that they think it's unlikely." The tutor told her that she should add why the authors think that the counterargument is unlikely.

Interpretation. Three students identified the two counterarguments and the refutation of Article 3. The answers of the other two students were incomplete. Three students identified the counterargument and refutation of Article 4, while the answers of the other students' were incorrect. The counterarguments and refutations in the research articles have lexical and organizational features, but they are not very pronounced. So, students have to rely on content-based features for identification. Apparently, this was difficult for them. Additionally, the features of counterarguments and refutations were not discussed during the tutor group meetings.

5.5 Discussion

In this chapter, we described which lexical, organizational, and content-based features undergraduate life science students use for identifying rhetorical moves. Our results suggest that students rely heavily on lexical and organizational features, because students seemed to have the least difficulty with the identification of the motive and objective (see Table 5.3). These were the moves that had the most pronounced lexical and organizational features. They contain phrases such as "still unknown" (for the motive) and "the aim of this study" (for the objective). Furthermore, we saw that they were often located close to each other at the end of the Introduction section.

Students were slightly less successful in the identification of the main conclusions of Articles 2 and 4 (see Table 5.3) – even though their lexical, organizational, and content-based features were – in our opinion – relatively pronounced. It appears that students ignored certain organizational and content-based features, but we do not know for what reasons. For example, in Article 2 some students identified a sentence from the Results section as main conclusion because it summarized results. So, in this case they did not use the organizational feature that the main conclusion is located in the Discussion section. Furthermore, students did not identify the complete main conclusion of Article 4 because they felt that this move has to be short. It seems that they did not determine if the sentences they identified as conclusion contained the answers to the research questions mentioned in the Introduction section of Article 4.

Students' identification of implications was more problematic. Students stated during the tutor group meetings that they were confused about the content-based features of this move (e.g. they wondered if a follow-up study was an implication).

Students had also difficulty with the identification of supports, counterarguments, and refutations (see Table 5.3). The locations of these moves are rather variable and lexical features are rare. In these cases, it seems that students are more dependent on content-based features. Because research articles often contain multiple supports, counterarguments, and refutations, students do not know beforehand how many of these moves occur in an article. This could make it difficult for students to identify all these moves. Furthermore, the rhetorical value of counterarguments and refutations are concepts that are probably hard to grasp for students due to their lack of knowledge about effective rhetorical strategies. For example, David thought that scientists undermine their own research if they mention counterarguments that cannot be refuted.

Our results suggest that especially the identification of moves with less pronounced lexical and organizational features (implications, supports, counterarguments, and refutations) was problematic for students. So, it seems that students rely mainly on lexical and organizational features and less on content-based features when identifying moves. We hypothesize that students use lexical (and sometimes organizational) features to identify the move and then use content-based features to check if their assessment is correct. This could explain why students ignored, in some cases, the content-based features of moves that had no pronounced lexical or organizational features (such as counterarguments). In future studies, we could further test this hypothesis with task-based interviews and the think-aloud method, by which we can study in more detail the decisions individual students make when identifying moves (van Someren, Barnard, & Sandberg, 1994; Goldin, 2000).

We suspect that expert readers use – depending on the situation – organizational, lexical, and content-based features when they read a research article. Novice readers rely more on organizational and lexical features and less on content-based features. Because novice readers' knowledge of the research field and the use of rhetorical devices are limited, they have difficulty with decoding the content of each statement. Consequently, we suspect that they will pay more attention to features that are not based on content. However, we expect that as they gain more reading experience and more content knowledge, students will focus increasingly on the content when identifying rhetorical moves.

In our study, we have made a distinction between lexical and content-based features. However, these two features may overlap since textual content is always presented by lexical means. For example, the phrase *the aim of the study* is can be interpreted as a lexical feature because it signals an objective. However, the phrase is also part of the content of the sentence and text in which it is located.

A limitation of this study is that our data sources contained predominantly indirect evidence about students' use of features to identify rhetorical moves. Unfortunately, students were not triggered to discuss this use of features in great detail. A possible explanation for this is that they were not conscious of the features they used. It could also be that we did not provide them with enough opportunities to discuss these features. However, we see this study as the first step in gaining insight into students' rhetorical reading of research articles.

Another limitation is that we were not able to gain much insight into students' identification of supports. Students' answers were too varied to discern any patterns in them regarding the features they use to identify supports. For this reason, we will study in Chapter 6 students' identification of supports in more detail using a different set of data.

The tutor group meetings show that lexical, organizational, and content-based features of the motive, objective, main conclusion, and implication were discussed a number of times. In contrast, the tutor and the students rarely discussed the features of supports, counterarguments, and refutations. We cannot say with certainty why this is the case. It could be that the features of these moves are less tangible than the features of other moves.

We have three suggestions that could enhance students' ability to identify supports, counterarguments, and refutations:

1. The use of effective rhetorical strategies should be discussed more during the module, so that students will be more appreciative about authors' use of counterarguments and refutations.
2. Students should be provided with more opportunities to exchange their answers and their views on the features of rhetorical moves (e.g. by letting them discuss their answers to the assignment in pairs during the tutor group meetings). These kinds of exchanges between students were now relatively rare.
3. The instructions we gave to the tutors should be more explicit. In this study, they were too limited. As a result the tutors did not use the opportunities to discuss more extensively the identification of moves. A better preparation could improve their ability to discuss the features of rhetorical moves. This would help to bring suggestion 2 into practice and it would also ensure that the tutors stay close to the move descriptions given in the information sheets.

We asked each tutor to show the students how he/she reads research articles and identifies rhetorical moves. In this way, they would "model" the identification process: a phase of cognitive apprenticeship in which the expert performs a task so that students are able to observe and build a conceptual model of the processes that are necessary to accomplish it (Collins, Brown, & Holum, 1991). However, we observed that the tutor let students identify rhetorical moves during the tutor group meetings and then gave directly feedback on their answers. This corresponds with the remarks of some tutors in Chapter 4 that they found modeling difficult to do. If tutors would receive a better preparation, as suggested above, we could teach them to use effective modeling techniques.

Chapter 6
Students' Ability to Coordinate Rhetorical Moves with the Use of a Heuristic

6.1 Introduction

To help undergraduate life science students with adopting appropriate strategies when reading a research article, we developed a teaching strategy that was implemented in an eight-weeks module called Reading Research Articles. The central design principle of our teaching strategy was the focus on the rhetorical structure of a research article. For this purpose, students practiced the identification of seven rhetorical moves: the motive, objective, main conclusion, implication, supports, counterarguments, and refutations. Descriptions of these rhetorical moves are a part of the Scientific Argumentation Model (SAM) we developed (see Chapter 3). In Chapters 4 and 5 we described how we used SAM as a heuristic in our module.

An important component of SAM is an argumentation scheme (Figure 6.1). This so-called SAM scheme depicts the relations between the different rhetorical moves in a research article. In Chapters 4 and 5 we only described and analyzed assignments in which the students had to identify individual rhetorical moves – without putting them in the argumentation scheme. In this chapter, we will describe and analyze an assignment in which students were asked to complete a SAM scheme. We believe that completing such a scheme may help students with getting a clearer overview of the relations between rhetorical moves. This will lead to a better understanding of the text. That is why we want to determine to what extent students are able to construct a SAM scheme. Furthermore, in Chapter 5 we did not analyze students' identification of supports in great detail. Hence, in this chapter we will closely study students' identification of supports.

As explained in Chapter 3, the SAM scheme with its rhetorical moves represents the essential elements of a research article's argument. An important aspect of the SAM scheme is the justification of the main conclusion by means of supports. Below the central arrow, which connects the objective to the main conclusion, the article's supports are depicted. These

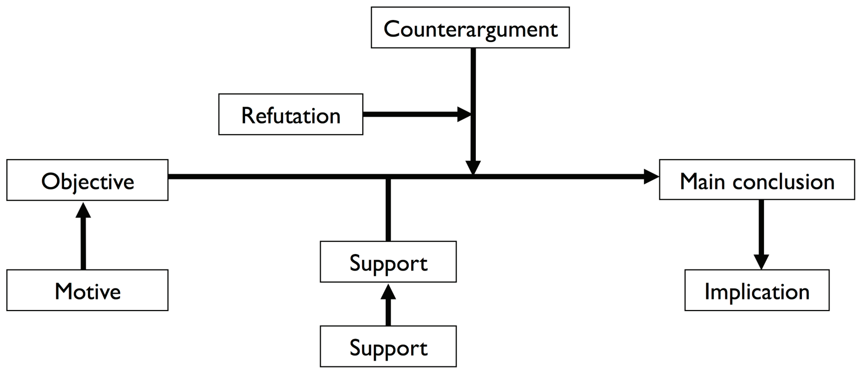


Figure 6.1. The SAM (Scientific Argumentation Model) scheme.

supports form "pillars" or "chains" that justify the main conclusion. We distinguish four types of supports: inscriptions[1], statements of findings, preliminary conclusions, and references to previous research.

Supports may be depicted in chains containing inscriptions and text fragments (statements of findings, preliminary conclusions, or references) that are related to each other. For example, the base of such a chain may contain an inscription. This inscription may lead to a statement of finding ("Figure 1 shows that the blood flow decreases"). Then, a preliminary conclusion may follow ("Drug X influences the blood flow"). This preliminary conclusion is connected to the central arrow, which points to the main conclusion. Understanding these support chains is important for interpreting and evaluating research articles. According to Kelly, Regev, and Prothero (2008), one of the criteria for determining the quality of (written) scientific argumentation is the coordination of evidence across epistemic levels (see Chapter 3) in a text: making explicit "how particular inscriptions or claims provide evidence for higher order, more generalized claims" (p. 133). These epistemic levels progress from specific propositions to more general and theoretical propositions. In Kelly et al. (2008) the epistemic levels consist of data inscriptions, low inference claims, claims supported by previous claims, theoretical claims, and thesis. It is difficult to compare these levels to our own description of supports, as Kelly et al. (2008) do not clearly define their levels; however, the lines of reasoning they envisaged between epistemic levels are comparable to our support chains.

[1] An inscription is not a rhetorical move, as it is an extra-textual element (see Section 2.3). However, we still included them as supports in SAM, since they are often at the base of a support chain.

6.2 Research Questions

In this chapter, we will determine students' ability to construct a SAM scheme. We will especially examine their ability to compile support chains. This leads to the first research question, consisting of three sub-questions:

1. To what extent are undergraduate life science students able to construct a SAM scheme?
   - 1a To what extent do undergraduate life science students identify motive, objective, main conclusion, implication, counterarguments, and refutations when constructing a SAM scheme?
   - 1b To what extent do undergraduate life science students coordinate motive, objective, main conclusion, implication, counterarguments, and refutations when constructing a SAM scheme?
   - 1c To what extent do undergraduate life science students coordinate supports when constructing a SAM scheme?

Regarding the coordination of the motive, objective, main conclusion, implication, counterargument, and refutation, we can – based on the SAM scheme – distinguish four relations between these moves: (1) the motive needs to be related to objective, (2) the objective needs to be related to the main conclusion, (3) the main conclusion needs to be related to the implication, and (4) the counterarguments/refutations need to be related to the main conclusion. For studying the coordination of supports, we will determine if the supports in the chains are related to each other regarding their content. For instance, does the students connect the right text fragment to a graph? Or a text fragment from the Discussion section with a text fragment from the Results section? This coordination of supports is comparable to the coordination of evidence across epistemic levels.

Furthermore, we will ask students if they feel that using a SAM scheme helps them with understanding a research article. Our second research question was:

2. How do undergraduate life science students appreciate the SAM scheme as a tool for understanding a research article?

6.3 Method

Educational Setting and Participants

As in Chapters 4 and 5, the participants of this study were first-year undergraduate life science students who followed in 2009 the first edition of an eight-weeks module called Reading Research Articles as part of the course Biomedical Research. The details of this module are described in Chapters 4 and 5. The total number of students who entered the module was 125; 55 students were male, 70 were female. They were approximately 18–20 years old. The students were subdivided into 14 tutor groups.

Assignment

At the end of the tutor group meeting in Week 6 of the module, students received a sample scheme: a SAM scheme of a previously discussed research article (see Appendix 2). The tutors told the students that this was an example of an argumentation scheme and explained to them the relations between the rhetorical moves. Then, they received a new research article and an assignment. For the assignment, students had to identify seven rhetorical moves (motive, objective, main conclusion, implication, supports, counterarguments, and refutations) and put them in a SAM scheme.

The tutor groups were divided into Group A (tutor groups 1 until 7, n = 72 students) and B (tutor groups 8 until 14, n = 53 students). The students in Group A received the following assignment:

> Depict the argumentative structure[2] of the article. Put all the argumentative elements[3] in a scheme as explained during the previous tutor meeting. Hand in your answer on paper. Use the authors' original phrases – if necessary use (…) to shorten long quotes. Put behind each fragment a letter that indicates the type of element: (M) motive, (O) objective, (MC) main conclusion, (I) implication, (S) support[4], (C) counterargument, and (R) refutation. Please include for each element the page number.

The students in Group B received the following assignment:

> Depict the argumentative structure of the article. Hand in your answer on paper. Use the authors' original phrases – if necessary use (…) to shorten long quotes. Put behind each fragment a letter that indicates the type of element: (M) motive, (O) objective, (MC) main conclusion,

[2] For the students, we used the term "argumentative structure" instead of "SAM scheme."

[3] For the students, we used the term "argumentative element" instead of "rhetorical move."

[4] As the information sheets (Appendix 1) and the sample scheme make clear that students should compile supports chains, the tutor did not explicitly instructed them to do this.

(I) implication, (S) support, (C) counterargument, and (R) refutation. Please include for each element the page number.

Students handed in their schemes during the subsequent tutor group meeting, seven days later (Week 7). These schemes were collected by the tutors and handed over to us.

We gave students two different assignments to see if students from Group B tended to construct schemes with an alternative layout. This was not the case (data not shown), so from now on we will not make a distinction between students from Groups A and B.

Research Article

The research article used was written by Bauersachs et al. in 1999 and is titled "Endothelial Dysfunction in Chronic Myocardial Infarction Despite Increased Vascular Endothelial Nitric Oxide Synthase and Soluble Guanylate Cyclase Expression: Role of Enhanced Vascular Superoxide Production." The course lecturers selected this article. In the article – which will be described in more detail in the Results section (Section 6.4) – the expression of certain enzymes that play a role in the vasodilation of arteries was studied in rats with endothelial dysfunction. Endothelial dysfunction, a symptom of chronic heart failure, is a state in which the endothelium (the inner lining of blood vessels) does not function normally, resulting in impaired vasodilation and vasoconstriction. The authors found that the production of certain enzymes that have a positive influence on vasodilation (like eNOS, which catalyzes the production of the vasodilator nitric oxide) is increased in rats with endothelial dysfunction. However, the increased expression did not restore the decreased vasodilation. The authors' results suggest that this is caused by the rapid inactivation of nitric oxide by superoxide anions, which are produced by a NAD(P)H-dependent oxidase.

The authors' study consisted of several experiments and many different parameters are described. The article contained two tables and six figures. The main body of the text (including the Abstract, but without the Reference section) contained approximately 3,400 words. Students were familiar with the concepts that are described in the article, because some of them were also mentioned in previous articles used in the module and because the concepts were discussed during lectures beforehand.

Coding Scheme and Data Analysis

For the analysis of students' answers, we gave each of the 80 sentences of the Abstract, Introduction, Results, and Discussion sections of the research article a code consisting of a letter denoting the section and a sentence number. Students did not write down sentences from the Method section, so the sentences in this section were not coded. The Abstract consists of sentences a1 to a9, the Introduction section of i10 to i25, the Results section of r26 to r50, and the Discussion section of d51 to d80. The sentences can be found in Appendix 3. Tables were denoted with the codes t1 and t2; figures with f1, f2, f3, and so forth.

I constructed a SAM scheme of the article. Based on the remarks from a second researcher, this argumentation scheme was revised until we both agreed on its structure (Figure 6.2). The scheme is not complete; for simplicity's sake, we left out text fragments referring to data that are not shown by the authors. Furthermore, we left out supports that were references to previous studies, as students did not include these in their schemes.

For each of the schemes produced by the students (see Figure 6.3 for an example of a student scheme), we assigned the abovementioned codes to the text fragments and inscriptions. To answer Research Questions 1a and 1b, we inventoried students' identification of the motive, objective, main conclusion, implication, counterarguments, and refutations and compared students' answers with our own analysis of the research article. Our own analysis was used as a yardstick for determining students' ability to identify and coordinate rhetorical moves.

To answer Research Question 1c, we analyzed students' support chains. Because some students constructed quite complex support chains – just like we did – we split these into separate ones (Figure 6.4). For the analysis of the chains, we first used descriptive statistics (average number of support chains per student, average number of supports per chain, et cetera). We also determined the frequency of types of supports chains. These types of support chains were based on the location of the supports in the text. For example, one type of support chain was "F/T→R→D-," meaning that

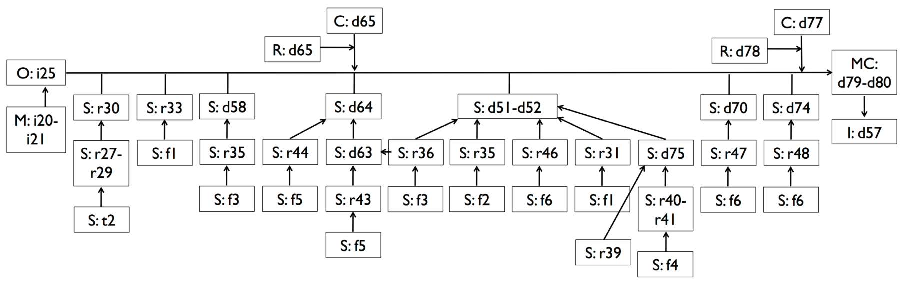


Figure 6.2. The article's SAM scheme used as coding scheme. We used the following letters to denote the different moves: (M) motive, (O), objective, (MC), main conclusion, (I), implication, (S) support, (C) Counter Argument, and (R) refutation.

the chain had a figure or table at the base, followed by a text fragment from the Results section and a text fragment from the Discussion section.

To get a clear overview of the different support chains, we made so-called trees of supports. Because the majority of students' chains had a figure or table at the base (just as our chains in Figure 6.2), we made a tree of supports for each figure and table. Each tree shows how students connect

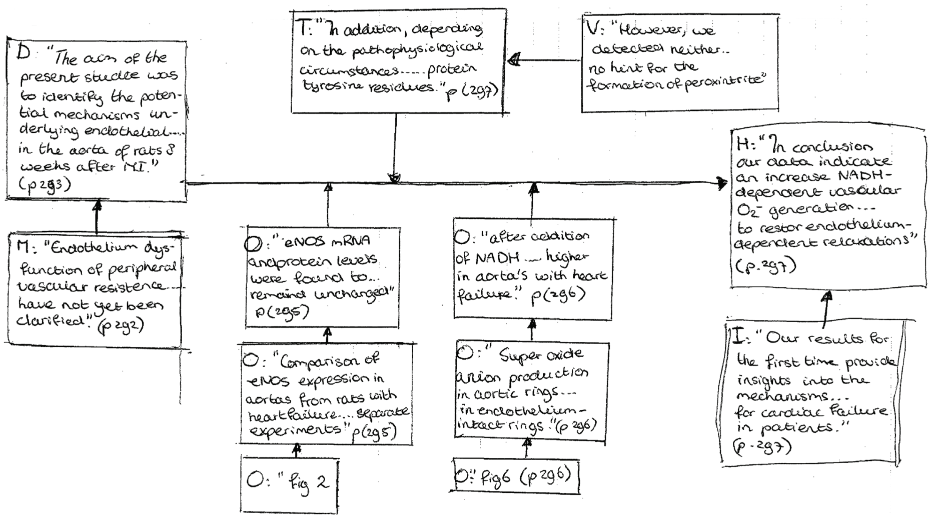


Figure 6.3. A SAM scheme made by a student. The letters are based on Dutch translations of the different rhetorical moves: (M) motive, (D) objective, (H) main conclusion, (I) implication, (O) support, (T) counterargument, and (V) refutation.

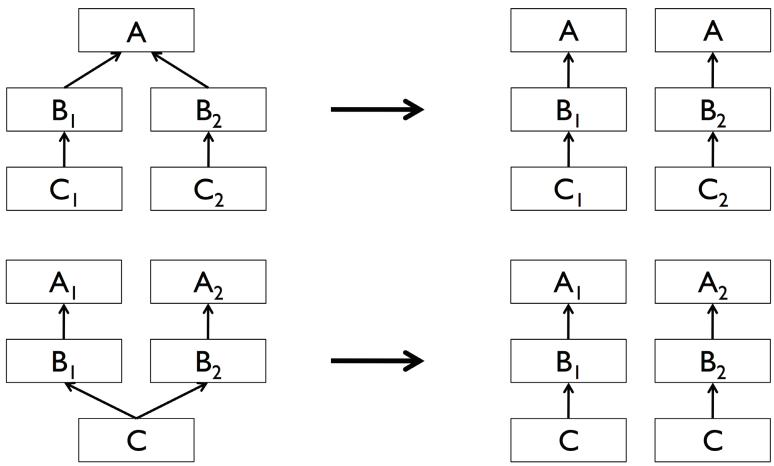


Figure 6.4. Complex support chains were split into separate chains for analysis.

a certain inscription to other supports. We used the open-source computer program Freemind (http://freemind.sourceforge.net) for the construction of these trees. To determine if the supports in the chains were related to each other, we compared the trees of supports with our own scheme in a qualitative way.

Questionnaire

To answer Research Question 2, the students completed at the end of the module a questionnaire containing the following two statements:

- Constructing an argumentation scheme (as explained by the tutor) is a helpful tool for understanding a research article.
- In future, I will more often use an argumentation scheme to depict a research article's arguments.

Via a 5-point rating scale, students could indicate how much they agreed with these statements. Students could also write down comments about their experiences with using the SAM scheme.

6.4 Results

As data source we used the SAM schemes of 73 students. We did not receive the SAM schemes of 28 students, as some tutors did not pass on students' schemes to us. Furthermore, we could not analyze the products of 24 students because they fell into one of these three categories:

- The product consisted of rhetorical moves that were not put in a scheme (n = 12).
- The product consisted of a SAM scheme with paraphrased moves that could not be coded (n = 7).
- The same product was handed in by different students – so we could not identify the author of this scheme (n = 5).

Identification and Coordination of Motive, Objective, Main Conclusion, Implication, Counterarguments, and Refutations

Motive. Regarding the motive, 30 of the 73 students identified sentence a1 or sentences i10 and i11. These sentences represent motives of some sort: they describe a gap in knowledge ("underlying mechanisms have not yet been clarified"). However, this gap is rather non-specific and does not explain why the authors conducted their study the way they did. That is why we designated sentences i20 and i21, which were more specific, as motive:

> i20/i21: With regard to heart failure, no data are available on O_2^- production within the vascular wall or on the potential alterations of the expression of smooth muscle sGC. Moreover, the influence of heart failure on vascular NOS expression is still controversial. (p. 293)

Twenty-six students identified i20/i21 as motive. Six other students identified both a1 or i10/i11 and i20/i21. So, 26 + 6 = 32 students were able to identify the same motive as we did. The remaining 11 students identified a variety of other sentences as motives.

Objective. Regarding the objective, all 73 students identified sentence i25 (just like we did):

> i25: The aim of the present study was therefore to identify the potential mechanisms underlying endothelial dysfunction in heart failure by the simultaneous determination of endothelium-dependent dilator responses, the expression of the key enzymes of the NO/cGMP system, and O_2^- formation in the aorta of rats 8 weeks after myocardial infarction. (p. 293)

So, all 73 students identified the same sentence as objective as we did.

Main conclusion. Regarding the main conclusion, 9 students identified sentences a8 and a9 (from the Abstract section), 8 students identified sentences d51 and d52 (located at the beginning of the Discussion section), and 56 students identified sentences d79 and/or d80 (located at the end of the Discussion section). Sentences d51/d52 are very specific and more like a summary of the results. Sentences a8/a9 and d79/d80 present the authors' findings in a more generalized way (preliminary conclusion), which is why we designated d79/d80 (or alternatively a8/a9) as main conclusion:

> d79/d80: In conclusion, our data indicate that an increased NADH-dependent vascular O_2^- generation represents an important mechanism for the endothelial dysfunction in heart failure by enhancing the inactivation of NO. Even a presumably counterregulatory upregulation of eNOS and sGC is not sufficient to restore endothelium-dependent relaxations. (p. 297)

So, 56 + 9 = 65 students identified the same main conclusion as we did. However, as d51/d52 contain more or less the same information as a8/a9/d79/d80, we are inclined to state that the remaining eight students also identified the main conclusion.

Implication. There was less consensus between the students about which sentence represents the implication. Twenty students did not put an implication in their scheme. Of the 53 students who wrote down an implication, 10 identified sentence d57:

> d57: Our results for the first time provide insights into the mechanisms of the alteration of endothelial function in heart failure after myocardial infarction, which represents the most important cause for cardiac failure in patients. (p. 297)

We also designated this sentence as implication, because it implicitly hints at possible applications of the results: understanding the mechanism may contribute to the treatment of cardiac failure. However, the implication in this article was probably difficult to identify, as d57 does not really mention recommendations, practical applications, or suggestions for future research. The remaining 43 students identified a variety of other sentences. So, 10 students were able to identify the implication.

Counterarguments and refutations. Some students identified more than one counterargument and refutation. And since the answers of the students were very diverse, we cannot present all answers. This is why the numbers given here will not add up to 73. Only two students did not write down a counterargument and refutation. Twenty-one students only identified d64 as counterargument and/or the first part of d65, and as refutation the second part of d65:

> d64–d65: Because sGC activity is susceptible to superoxide[25] and cGMP production was restored by prior treatment with the radical scavenger Tiron, enhanced production of superoxide anions may be responsible for the reduced activity of sGC despite the increase in its expression. Enhanced degradation of cGMP due to increased phosphodiesterase activity[31] is not likely, because our experiments were performed in the continuous presence of a high concentration of a phosphodiesterase inhibitor. (p. 297)

Fifteen students identified only d77 and d78 as respectively a counterargument and refutation:

> d77/d78: In addition, depending on the pathophysiological circumstances, NO and superoxide may react to the powerful oxidant peroxynitrite, which can form hydroxyl radicals and nitrate protein tyrosine residues.[15] However, we detected neither enhanced luminol chemiluminescence nor tyrosine nitration in rats with heart failure, so there was no hint for the formation of peroxynitrite. (p. 297)

One student identified both d65(first part)/d65(second part) and d77/d78 as counterarguments and refutations. So, 37 students identified at least one correct counterargument and refutation.

Nineteen students identified d55 and d56 as respectively a counterargument and refutation. We did not identify d55 as a counterargument, because this sentence referred to NO production, which was not directly measured in this study (only indirectly, via eNOS and iNOS). However, because of the word “however” in sentence d56, it is understandable that students identified this sentence as a refutation.

In summary, students identified the objective and main conclusion in the research article, and to a lesser extent they identified the motive, implication, counterarguments, and refutations.

Regarding the coordination of moves, we can distinguish four important relations (not counting the supports):

- The relation between motive and objective: all students identified the objective. Only 32 students connected to this objective the same motive as we did. Thirty students identified a motive that did not relate to the objective.
- The relation between objective and main conclusion: all students identified the objective and main conclusion, so these moves were in all cases coordinated in a correct way.
- The relation between main conclusion and implication: Only 10 students identified the same implication as we did. Other students did not mention any implication or mentioned a variety of different sentences. So, it is difficult to assess if these two moves were coordinated in a correct way.
- The relation between counterargument/refutation and main conclusion: 37 students identified at least one counterargument and refutation that could be related to the main conclusion. It should be noted that 19 students identified a counterargument and refutation that did relate to other studies, and not the authors’ main conclusion.

Because all students identified the same objective and main conclusion and placed these two elements in their SAM scheme, their chains all support the central arrow consisting of the same objective and main conclusion. This means that we could analyze students’ support chains as one group. Of course, the fact that students’ schemes contain moves that are connected by lines or arrows does not mean that they related these moves consciously.

Table 6.1. Frequency of types of support chains. Support chains can consist of text fragments from the Discussion section (D), text fragments from the Results section (R), or a figure or table (F/T). The dash (–) represents the connection between the uppermost support and the central arrow that connects the objective and main conclusion in the SAM scheme.

Type of support chain	Frequency
F/T→R→D–	105
F/T→R–	65
D–	22
F/T→D–	11
F/T→R→R–	11
F/T→D→D–	7
F/T→D→D→D–	3
R–	3
R→D–	2
D→D→D–	1
F/T→D→F/T→D–	1
F/T→D→F/T→R–	1
D→D–	1
Total	233

Identification and Coordination of Supports

In total, the 73 students wrote down 577 supports[5] in 234 chains. Twenty-five of the 234 chains consisted of one support, 80 chains consisted of two supports, 124 chains consisted of three supports, and five chains consisted of four supports. On average, each student wrote down 3.2 chains (ranging from one to seven) with in each chain 2.5 supports (ranging from one to four). Of the 577 supports, 206 were figure(s) or table(s), 201 were text fragments from the Results section, and 169 were text fragments from the Discussion section. One of the given supports could belong to either the Results or Discussion section; so, the chain in which this support occurs was left out in the subsequent analysis. This left us with 233 chains.

In total, students constructed 13 different types of support chains (Table 6.1). Students most often (n = 105) constructed a support chain with a figure or table at the base, followed by a text fragment from the Results section and then a text fragment from the Discussion section. Sixty-five

[5] An individual support may be a figure, table, or a text fragment containing one or more sentences.

chains consisted of a figure or table at the base followed by a text fragment from the Results section. These chains did not contain a connection with a text fragment from the Discussion section.

Of the 233 chains, 204 had a figure or table at the base, 24 a text fragment from the Discussion section, and 5 a text fragment from the Results section. For our subsequent analysis, we used the 204 support chains that had only one figure or table at the base to construct so-called trees of supports. These trees of supports (see Figures 6.5, 6.6, 6.8, 6.9, 6.10, 6.11, and 6.12) show how students connect other supports to a certain figure or table. Because only a small minority of support chains had a text fragment at the base, we decided to exclude these chains from our analysis. Also, to simplify the analysis, we left out the nine chains that had multiple figures and/or tables at the base (i.e. when an individual support contains a combination of figures and/or tables). This left us with 195 chains. These 195 chains had the following figures and tables at the base: t1 (n = 1), t2 (n = 0), f1 (n = 14), f2 (n = 53), f3 (n = 18), f4 (n = 28), f5 (n = 17), and f6 (n = 64), Below, we will discuss the role of each table and figure in the research article. Then, we will analyze the tree of supports.

T1/T2. Only one student's support chain had t1 at the base (Figure 6.5) and none of the students' support chains had t2 at the base. t1 describes a methodological detail (the primers that were used for PCR) and does not contain any experimental data. For this reason, we did not consider this table a support. T2 contains experimental data showing that rats with a myocardial infarction have (among other things) a lower blood pressure compared to the sham rats (sham rats are rats, which were operated, but which did not get an infarct – i.e. the control group). This is described in sentences r27 till r29. Based on these observations, the authors conclude that the infarct rats "demonstrated heart failure in a compensated stage" (sentence r30). It is therefore likely that the student who constructed this chain meant t2 instead of t1.

F1. The authors obtained aortic rings from the sham and infarct rats to determine the endothelium-dependent and endothelium-independent relaxations. The results of these experiments are depicted in f1a (endothelium dependent) and f1b (endothelium independent). In the Results sec-

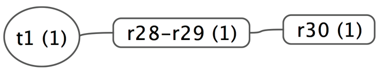


Figure 6.5. Student's support chain with t1 (although this student probably meant t2) at the base. Between brackets is the number of chains containing this connection.

tion, the authors write that figure f1a shows that aortic rings from infarct rats were less able to relax than rings from sham rats (sentence r31). In the Discussion section, they transform this observation into sentence d51: "we observed a pronounced endothelial dysfunction in rats with chronic myocardial infarction." This reasoning step is warranted by the implicit assumption that decreased relaxation is a symptom of endothelial dysfunction. So, f1 is just as t2 used as an experimental check: it shows that the infarct rats have heart failure and endothelial dysfunction. According to the authors, f1b shows that there was only a small difference in relaxation between aortic rings from infarct rats and aortic rings from sham rats (sentence r33). This is not an unexpected result, as endothelial dysfunction does not alter endothelium-independent relaxations. That is probably why the authors do not discuss this result any further in the article.

Students wrote down 14 chains with f1 at the base (Figure 6.6). In 11 of these chains, f1 was connected with sentence r31. Sentence r31 was in none of these chains connected with sentence d51/d52, as in our analysis. Instead, r31 was incorrectly connected with r32 (a statement which describes data that is not shown), d53 (general information about heart failure, based on previous literature), and d57–d58 (sentences which refer to f2 and f3). In our analysis we also connected f1 with sentence r33. Only two students' support chains contained this connection.

So, Figure 6.6 shows that f1 is in 11 of the 14 chains connected to the same text fragment from the Results section as in our analysis (r31). In contrast, none of the chains contained the connection from r31 to d51/52.

F2/F3. The authors determined eNOS protein, eNOS mRNA, iNOS protein, and sGC protein levels in aortas from infarct and sham rats. All these four molecules play a role in the vasodilation process (see Figure 6.7). eNOS protein and mRNA data are depicted in f2. In the Results section,

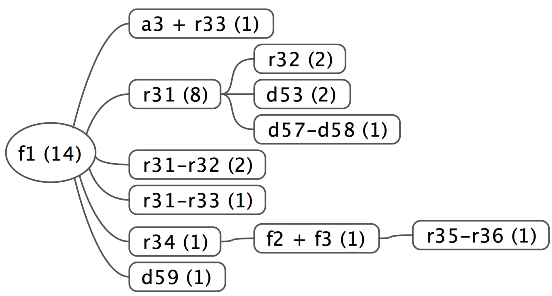


Figure 6.6. Students' support chains with f1a/f1b at the base. Between brackets is the number of chains containing this connection.

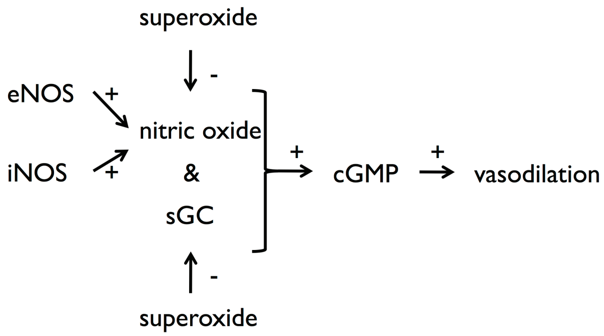


Figure 6.7. The effects of eNOS, iNOS, nitric oxide, sGC, cGMP, and superoxide on vasodilation.

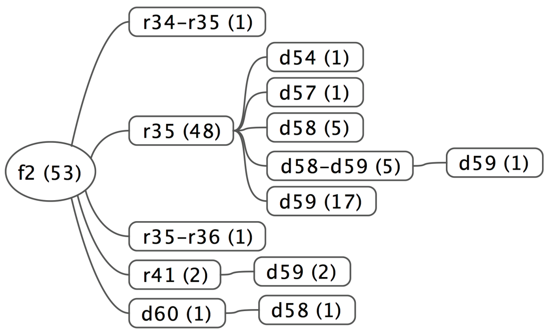


Figure 6.8. Students' support chains with f2 at the base. Between brackets is the number of chains containing this connection.

the authors state that f2 shows that eNOS protein levels are significantly higher in infarct rats compared to sham rats (sentence r35). iNOS and sGC protein data are depicted in f3. In sentence r35 the authors state that in both groups iNOS expression was hardly detectable. This observation is recapitulated in sentence d58, where iNOS expression is compared with eNOS expression. In sentence r36 they refer to f3 and state that sGC protein levels are "markedly enhanced" in infarct rats compared to sham rats. In the Discussion section, the authors summarize these results by stating that they observed a "marked upregulation in the expression of two key enzymes of vasorelaxation: eNOS, regulating the synthesis of the most important vasodilator, NO, and its target enzyme in smooth muscle cells, sGC" (d51). Furthermore, the authors mention the increased sGC expression in d63. In this sentence, the authors speculate about the meaning of this result.

F2 was incorporated into the base of 53 chains (Figure 6.8). F2 was 50 times connected to r35, as in our analysis. But although we connected r35 to d51/d52, the students connected r35 23 times to d59. This is not completely wrong, as the authors mention in d59 the increased eNOS expression when comparing their results to other studies. F3 was incorporated into the base of 18 chains (Figure 6.9). In 12 of these chains, f3 was connected to r36. Then, r36 was connected in seven cases to d63 and/or d64. In none of the chains a connection was made between r36 and d51/d52, as in our analysis.

Figure 6.8 shows that in 50 of the 53 chains f2 was connected to r35. Figure 6.9 shows that f3 was connected respectively 4 and 12 times to r35 and r36. So, Figures 6.8 and 6.9 show that almost all students connected f2 or f3 to a text fragment from the Results section. Respectively 32 and 13 chains in Figures 6.8 and 6.9 contained a text fragment from the Discussion section (although generally not text fragments that corresponded with our analysis).

F4. The influence of superoxide anions on relaxation is depicted in f4. In sentences r40 and r41, the authors state that f4 shows that adding SOD (which removes superoxide anions) enhances the relaxation of infarct rats' aortic rings (r40) and DETC (which inhibits SOD) abolished relaxation (r41). In the Discussion section (sentence d75) they further interpret r40 by stating that in the presence of SOD, there was a "partial restoration of the acetylcholine-induced relaxation." This result can be viewed as an additional support for the authors' claim that there is an increased production of superoxide anions in infarct rats (sentences d51/d52).

F4 was incorporated into the base of 28 chains (Figure 6.10). Interestingly, in 12 chains f4 was connected to r39 – a sentence describing the effects of adding 600 U/ml SOD to the organ baths. However, in f4 the effects of 200 U/ml SOD are shown. So, r39 is not based on f4. In 17 chains

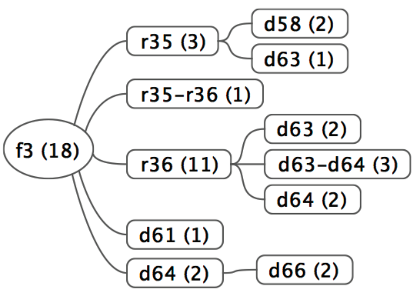


Figure 6.9. Students' support chains with f3 at the base. Between brackets is the number of chains containing this connection.

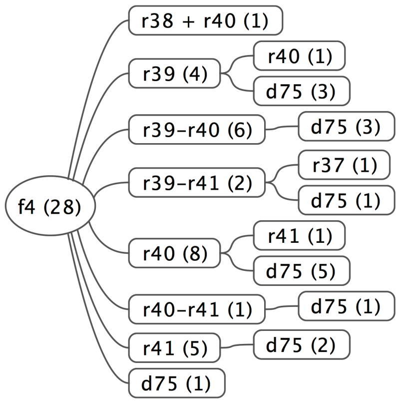


Figure 6.10. Students' support chains with f4 at the base. Between brackets is the number of chains containing this connection.

f4 was connected to r40. In 10 of these chains, r40 was connected to d75, which corresponds with our analysis.

Figure 6.10 shows that students did have difficulties with connecting f4 to the correct text fragment from the Results section, as 12 chains contain a connection to r39. Sixteen chains contain a connection to a text fragment from the Discussion section; this was often d75, which corresponds with our analysis in which f4 is connected to r40–r41 and then d75. In contrast to the students, we connected d75 via d51–d52 to the central arrow.

F5. The authors also determined cGMP formation in aortic rings taken from sham and infarct rats. cGMP levels are depicted in f5. According to the authors, f5 shows that (after adding SNP) cGMP levels are lower in infarct rats than in sham rats (sentence r43). Furthermore, adding Trion (which removes superoxide anions) restored cGMP levels in infarct rats (sentence r44). In the Discussion section, the authors combine r43 and r36 (which states that sGC protein levels were enhanced in infarct rats): "...we observed an upregulation of sGC expression that was associated with a blunted cGMP formation in response to sodium nitroprusside" (sentence d63). Then, the authors use d63, r44, and a reference to previous research to conclude:

> Because sGC activity is susceptible to superoxide[25] and cGMP production was restored by prior treatment with the radical scavenger Tiron, enhanced production of superoxide anions may be responsible for the reduced activity of sGC despite the increase in its expression. (sentence d64)

Here we see that the authors transform “a blunted cGMP formation” (from sentence d63) into “reduced activity of sGC.” This reasoning step is implicitly warranted by the assumption that sGC is responsible for the production of cGMP.

F5 was incorporated into the base of 17 chains (Figure 6.11). In 10 chains f5 was connected to d63 and/or d64, without ($n = 4$) or with ($n = 6$) r43 and/or r44 in between. This corresponded with our analysis. In only one support chain a connection was made between d63 and d64, as the students tended to group them together as one support. In our own analysis, we treated them as individual supports as d64 follows from d63.

Figure 6.11 shows that students could connect f5 to a text fragment from the Results section (r43 and/or r44) and that six students could connect this text fragment to the correct text fragment from the Discussion section (d63/d64).

F6. The authors measured superoxide anion production in aortic rings from sham and infarct rats. This was done in the presence or absence of endothelium to determine if the source of superoxide anion production was the endothelium or vascular smooth muscle cells. The data from this experiment are shown in f6a. Referring to f6a, the authors state that superoxide anion release was “greater in aortas from rats with chronic myocardial infarction” (sentence r46). In the Discussion section, they refer implicitly to this result by mentioning the “increased formation” of super-

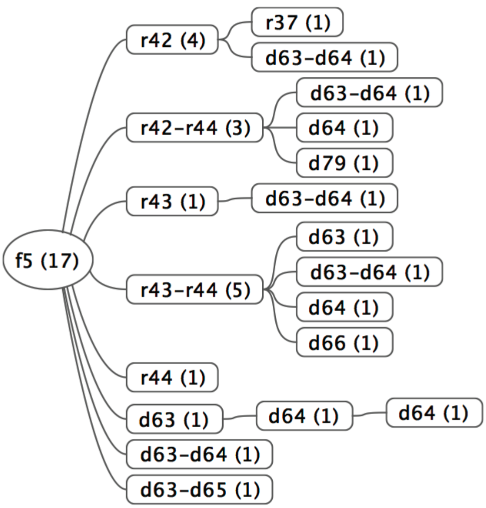


Figure 6.11. Students’ support chains with f5 at the base. Between brackets is the number of chains containing this connection. (The connection between d64 and d64 is due to a student who split up this sentence.)

oxide anions (sentence d52). According to the authors, f6a also shows that "removal of the endothelium slightly but not significantly reduced radical production in both groups" (sentence r47).

Additionally, superoxide anion production was measured in the presence of NADH. In this way, the authors could determine if NADH plays possibly a role in superoxide anion production. The data from this experiment are shown in f6b. Referring to f6b, the authors write that superoxide anion production "was markedly stimulated and significantly higher in

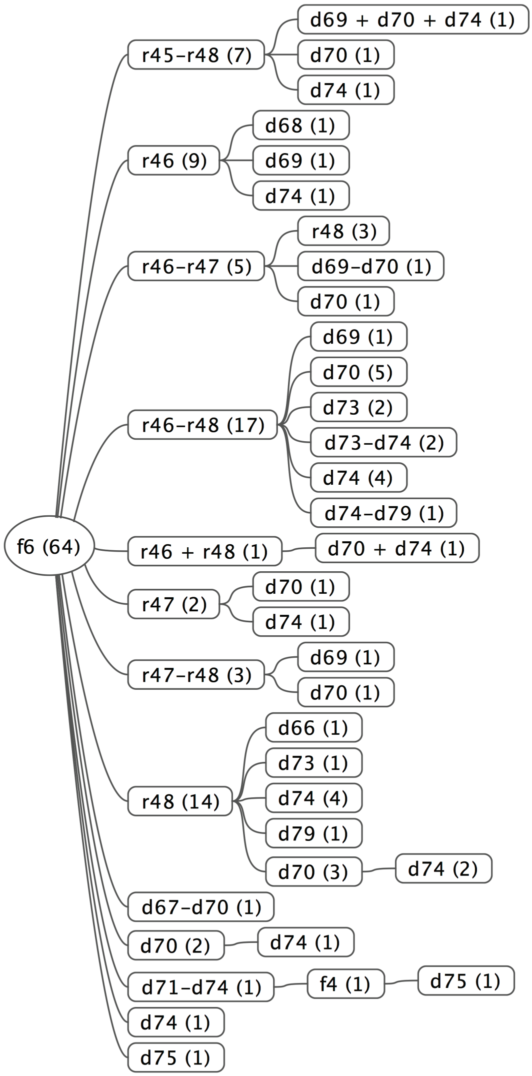


Figure 6.12. Students' support chains with f6a/f6b at the base. Between brackets is the number of chains containing this connection.

aortas from rats with heart failure" (sentence r48). In the Discussion section, the authors use r47 as support for their statement that "the source of superoxide formation appears to be vascular smooth muscle cells" (d70). R48 is used in the Discussion section (together with a number of references to previous studies) to argue that the expression of an NAD(P)H-dependent oxidase is possibly responsible for the superoxide anion production in aortas from infarct rats (sentence d74).

F6 was incorporated into the base of 64 chains (Figure 6.12). f6 was directly connected to r46 (n = 9), r47 (n = 2), r48 (n = 14), or a combination of r45, r46, r47, and/or r48 (n = 33). However, the connection of these sentences with d70 and d74, like we did in our analysis, was only seen in a minority of cases. For example, 17 chains contained a connection between f6 and r46 to r48, but in respectively 5 and 7 of these chains the connection was made between d70 and d74.

Figure 6.12 shows that a majority of the students could connect f6 to a correct text fragment from the Results section. Connecting these fragments to a text fragment from the Discussion section did not always go well. In contrast to our analysis, the students sometimes grouped r46, r47, and r48 together as one support, while these sentences represent three different lines of reasoning.

Students' Appreciation of SAM

Table 6.2 shows the results of the questionnaire. Almost half of the students (49%) believe that the construction of an argumentation scheme helps them with understanding a research article. As one student wrote: "You become acquainted with the structure of an article and the interconnectedness becomes clearer." Another student wrote: "You are forced to study the article more intensively." Roughly a third of the students (28%) did not find the construction of an argumentation scheme helpful. As one student wrote: "I underline the most important [sentences] and make a note of what it is, so a scheme is unnecessary." A minority of the students (15%) planned to construct an argumentation scheme more often. Students complained about the time-consuming nature of constructing an argumentation scheme. This could explain students' hesitancy to construct argumentation schemes in future.

6.5 Discussion

This study indicated that, generally, students were able to construct a SAM scheme after we showed them a sample scheme. Our results show that all students identified the objective and main conclusion and placed them in the right locations of the scheme, thereby linking these two moves. To

Table 6.2. Percentage of students ($n = 122$) who agreed with statements about the SAM scheme.

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Constructing an argumentation scheme (like explained by the tutor) is a helpful tool for understanding a research article.	3	25	23	41	8
In future, I will more often use an argumentation scheme to depict a research article's arguments.	5	37	43	15	0

a lesser extent they identified the motive, implication, counterarguments, and refutations. Regarding the coordination of these moves, we observed that students had particular difficulty with the relation motive-objective and counterargument/refutation-main conclusion. We also analyzed students' support chains. Students wrote down on average 3.2 support chains (ranging from 1 to 7) with on average 2.5 supports per chain (ranging from 1 to 4).

Roughly half of the students believes that the construction of an argumentation scheme (such as the SAM scheme) helps them with understanding a research article, although a minority indicated that they will use the scheme more often. Students felt that they benefited from the scheme, but they saw the time needed to use the scheme as a disadvantage.

Constructing support chains involves the coordination of evidence across epistemic levels. We can make a number of observations about students' coordination of evidence. Relating inscriptions with text fragments from the Results section went often well (as the different trees with support chains show). This could be explained by the authors' use of pointers such as *the results are shown in Figure 1* (Brett, 1994). This corresponds with the results of Chapter 5, in which we show that students rely in great deal on lexical features.

Comparing the trees of supports with our own analysis, we observed that students had difficulty with connecting text fragments from the Results section with the Discussion section. For example, 65 of the 204 chains consisted of a figure/table followed by only a text fragment from the Results section (see Table 6.1). It could be that students saw the connection, but thought that the text fragments in the Discussion section were just repetitions of information. For example, students possibly thought that r40 (SOD enhances the relaxation of infarct rats' aortic rings) and d75 (SOD partially restored acetylcholine-induced relaxation) contained basically the same information (see Figure 6.10). However, r40 and d75 belong to different epistemic levels. In d75 the data are further interpreted by making a causal connection between SOD and its effects on relaxation.

Coordination of evidence also involves the use of inscriptions in support chains. The two figures that were most often used in students' chains were f2 and f6. These are arguably the most important figures of the article, as they show that superoxide anion levels are enhanced while eNOS levels are also enhanced. So, students identified the most important supports. This in contrast to Chapters 2 and 4, where students did not construct a SAM scheme and had difficulties with identifying essential supports. It could be that the SAM scheme helps students with getting a better overview of the article's supports.

T2 and f1 were incorporated the least into chains. However, these supports are rather important in the authors' reasoning: they show that the infarct rats have heart failure and endothelial dysfunction. It seems that students did not see these control experiments as essential supports.

We observed that students tended to group different supports together, which belong to different support chains. For example, the authors use f6 to argue that superoxide anion production is higher in infarct rats (via r46), the source of these superoxide anions are vascular smooth muscle cells (via r47), and that an NAD(P)H-dependent oxidase is responsible for this production (via r48). These supports, however, belong to three different lines of reasoning. This suggests that students have difficulty with another aspect of the coordination of evidence: separating the different lines of reasoning in a research article.

In the Discussion section, the authors regularly referred to previous literature as supports. However, none of the students incorporated in their SAM scheme references as stand-alone supports. This could be explained by the fact that references were not mentioned in the sample scheme either. (Although the information sheets did mention supports that were references to other research.) Students probably did not realize that references could also be incorporated in the scheme.

As shown in Figure 6.2, the research article used in this study contained 13 support chains. Students only put a small selection of these support chains in their SAM schemes: their schemes contained on average 3.2 support chains. This suggests that students did not find all possible support chains, but we have to draw this conclusion with care because other interpretations of this observation are possible. It may be that students had limited space (they had to hand in the scheme on paper). It could also be that they stopped after writing down three support chains and assumed that they did not have to write down more (the sample scheme contained only three support chains). This might be avoided in future by making the assignments and the instructions from the tutor more directive.

The results of this study imply that several adjustments can be made to our teaching strategy. We should encourage the students to identify references to previous studies – for instance by expanding the sample SAM scheme with references to previous studies.

Our results also show that students' awareness of differences in epistemic levels could be further developed. Students seem to be not fully aware of the different epistemic levels in a research article and how these levels are connected to each other. We recommend to give students more training in the identification and the coordination of supports (e.g. by making a SAM scheme in small groups during the tutor group meeting).

Chapter 7
Students' Critical Evaluation of Research Articles

7.1 Introduction

The general aim of our research is developing and testing educational methods for introducing undergraduate life science students to research articles. In the previous chapters, we described a module called Reading Research Articles, which was developed for this purpose. In the module, we taught students how to identify a set of seven rhetorical moves with a heuristic. The development of this heuristic, the Scientific Argumentation Model (SAM), is described in Chapter 3. As shown in Chapter 4, students were more able to identify the rhetorical moves in a research article and they made this assignment (including reading the article) in less time at the end of the module. Furthermore, students adopted more expert-like reading behaviors. In this follow-up study, we extend our research to students' ability to critically evaluate the claims presented in two research articles with a similar topic.

Being able to critically evaluate science texts is an essential part of scientific literacy (Norris & Phillips, 1994; Norris & Phillips, 2003). As Duschl (2008) states, "evaluation of science claims (...) can become a core component of argumentation practices used in science education" (p. 173). This evaluation of scientific claims is a form of critical thinking, as it involves the purposeful judgment by means of reasoning (we elaborate on this below). Before one can evaluate a claim, one needs to be able to find the claim. As SAM describes the claim (in SAM this is called the main conclusion) of a research article and its supports, we think that our heuristic may support students in learning how to critically evaluate research articles.

The development of critical thinking skills is widely seen as an important aim of higher education (Tapper, 2004). However, it is felt that there is not enough emphasis in higher education on critical thinking skills. Already a long time ago, Popper (1970) noted: "The 'normal' scientist, in my view, has been taught badly. I believe, and so do many others, that all teaching on the University level (and if possible below) should be training and encouragement in critical thinking" (p. 52–53). More recently, White et al. (2011) determined the critical thinking skills of science students and scientists. Their results show that there was a significant increasing trend in critical thinking skills with increased academic level. However, they also saw room for improvement. For example, they observed that many senior science undergraduates still had difficulty with critical thinking tasks. They concluded that "science curricula fail to develop essential critical thinking skills in many science students" (p. 106).

Many authors have come up with definitions of critical thinking (e.g. Lipman, 1988; Siegel, 1988). We like to highlight two aspects of these definitions. The first aspect is the purposeful judgment by means of reasoning. For example, Ennis (1985) wrote: "Critical thinking is reflective and reasonable thinking that is focused on deciding what to believe or do" (p. 45). Or, as Facione (1990) stated: "We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based" (p. 3). The second aspect that needs to be highlighted is metacognition. For example, Paul and Elder (2002) defined critical thinking as "that mode of thinking – about any subject, content, or problem – in which the thinker improves the quality of his or her thinking by skillfully taking charge of the structures inherent in thinking and imposing intellectual standards upon them" (p. 15). This study will focus on the first aspect, as metacognition does not fall into the scope of our research.

Reading Research Articles in a Critical Way

Gillen (2006) distinguishes three different aspects of reading research articles. The first aspect is understanding the experimental narrative: the description of the methods and results. In order to do this, the reader has "to comprehend what was done and what was found, without trying to interpret the findings or assess the conclusions" (p. 35). The second aspect is interpreting the text: making sense out of the experimental narrative (e.g. drawing conclusions from the data). The third and final aspect is criticism: judging the strengths and weaknesses of the research article. The difference between interpretation and criticism is sometimes subtle, but "an important difference is that whereas interpretation involves finding meaning in the experimental narrative, criticism involves connecting the research article to other scientific works" (p. 36).

Especially this last aspect forms a demarcation between expert and novice readers of research articles. Andrews et al. (2006) found that 79% of first-year undergraduate students state that they accept the findings of peer-reviewed research articles. They also found that 85% of the students accept research findings if statistical of numerical data are given. Charney (1993) compared graduate students and experts while reading a scientific article, she discovered that the experts "were significantly more often engaged than the graduates in assessing the validity and value of the text" (p. 217). She states: "…the activities of the two groups suggest that while both were capable of understanding the text, the faculty were much more willing or able to enter the rhetorical fray" (p. 217). So, it seems that expert readers tend to read in a more critical way than novice readers. However, even the most distinguished scientist may find it difficult to read a text in a critical way. As Charles Darwin (1887) wrote: "I am (…) a poor critic: a paper or book, when first read, generally excites my admiration, and it is only after considerable reflection that I perceive the weak points" (p. 102).

Bazerman (1988), who interviewed seven physicists about their reading behaviors and observed four of them while reading research articles (see also Section 2.1), describes that if scientists want to mobilize parts of the article for their own work (or, as Bazerman calls it, fitting new material in their schema), they read in a critical way. They are, in other words, evaluating – or judging – the article. Articles are always judged for usefulness and importance by readers; one does not read an article before determining, implicitly or explicitly, if it is useful and important. The general criterion of scientists for determining the importance of a text was "the amount of news contained in the reading – that is, how significantly the article adds to or shakes up the current schema of what is known and how the field should go about knowing more" (p. 247). Bazerman observed that readers sometimes judge also the article's truth or quality. This will not happen when the reader feels that the content of the article lies outside his or her area of expertise. Judging the truth or quality of an article is only done if the reader possesses prior knowledge that is highly focused and articulated. If these judgments are made, readers use internal evidence (e.g. discussing experimental difficulties) and stylistic features (e.g. wording) as clues for determining the article's reliability. According to Bazerman, to evaluate the substance of specific statements, scientists generally rely on their own methodological experiences. For example, experimental techniques are examined to determine their correspondence with the reader's own experience of how such experiments should be run. Furthermore, Bazerman found that scientists also compared the article's results with the existing body of published experimental results by checking if the results fitted their expectations.

It could be that novice readers read less critically than experts because criticizing research articles requires knowledge about the standards of the general scientific community (Gillen, 2006). This knowledge may also be described as concepts of evidence. These are "all the *concepts* which are associated with obtaining evidence, including the relationship between the variables and the reality they represent, the way the data have been collected, represented and interpreted and the notion of fair testing" (Duggan & Gott, 1995, p.139). Novice readers often lack this knowledge (Gillen 2006). Concepts of evidence are related to epistemological criteria: philosophical assumptions, beliefs, and theories about the nature and limits of knowledge and its acquisition, which may serve as standards for assessing knowledge claims (Kitchener, 1983).[1] In science, according to Bailin (2002), epistemological criteria include the accuracy of data, control of experimental variables, reliability of sources, and validity of inferences. Background knowledge and concepts as necessary and sufficient conditions, correlation and causation, and hypothesis and prediction are also important for critical thinking. There also has to be a basic commitment to rational inquiry, such as respect for reasons, an inquiring attitude, open-mindedness, and fair-mindedness. Furthermore, readers also need to know the specific standards of evidence of an article's research field. What in one research field counts as acceptable evidence could be unacceptable in another field (Lewontin, 1991).

Another possible reason for difficulties with critical reading could be novice readers' lack of a so-called "critical spirit": the willingness to exercise one's ability to think critically (Siegel, 1988; Davson-Galle, 2004). The development of a critical spirit is probably closely connected to one's epistemological beliefs. If one sees a research article as an objective description of reality, then one is not likely to engage with the text in a critical way. However, if one sees the research article as a creative product that aims to persuade the reader, then it is much more likely that one develops a critical stance (Bazerman, 1988; Norris & Phillips, 2003). Research on the epistemological views of students in higher education suggests that especially science students tend to regard knowledge as factual and objective (and thus are less inclined to critically evaluate this knowledge). For example, Jehng, Johnson, and Anderson (1993) found that students who study in the "soft" fields (i.e. social science and arts/humanities) are more likely to believe that knowledge is uncertain than students in "hard" fields (i.e. engineering and business). Liu and Tsai (2008) compared science and non-science undergraduate majors and concluded that their results indicate that science majors have less sophisticated beliefs than non-science majors on the theory-laden and cultur-

[1] Epistemological criteria are closely related to epistemological beliefs: beliefs about the nature of knowledge and the process of knowing (Hofer & Pintrich, 1997). These beliefs are discussed in more detail in Section 8.5.

al-dependent aspects of science, although they found no differences in students' views on the certainty of knowledge – in contrast to Jehng et al. (1993).

Previous Studies on Critical Reading

Research suggests that reading primary literature enhances students' critical thinking skills (Hoskins, Stevens, & Nehm, 2007). Nevertheless, studies that describe how university students evaluate the content of an authentic research article seem to be very rare. However, a small number of studies is available describing university students who critically read media reports about scientific subjects (Korpan, Bisanz, Bisanz, & Henderson, 1997; Jiménez-Aleixandre & Federico-Agraso, 2009).

Some authors have formulated recommendations for teaching students how to read research reports in a critical way. For example, Myers (1990) mentions a number of strategies for the critical reading of scientific texts: Look for the rhetorical (i.e. realizing how authors try to make their claims acceptable) and reconstruct the social context (e.g. by determining if there are any controversies in the field). Reconstructing the social context can be done by looking for related texts, by looking for the source of authority (i.e. question the authority with which science is sometimes presented) and by looking for any links between scientific language and everyday uses of language (i.e. realizing that scientific texts contain – just as other types of texts – social and political implications). Hill, Soppelsa, and West (1982) suggested that students can practice their critical reading skills by writing reviews of research papers. Hill et al. propose the following questions students could use for critical reading:

> a) Is there a need?; b) Was the problem testable?; c) Were the methods and design appropriate to answer the research problem?; d) Were the procedures properly carried out?; e) Is the research replicable?; f) Were the findings completely and clearly reported?; g) Were the results valid and reliable?; h) Were the conclusions logical?; i) Were the limitations of the study inclusive?; j) Are the references complete?; k) Is the research meaningful?; l) Was the research article well organized and clearly reported? (p. 343)

Other authors have proposed somewhat similar checklists that students can use for the critical evaluation of research papers (e.g. Janick-Buckner, 1997; Gillen, 2007; Budovec & Kahn, 2010). However, these studies do not contain detailed information about students' performances when reading an article with a checklist.

7.2 Research Questions

In this chapter we describe a baseline measurement of first-year undergraduate life science students' critical evaluation of research articles. Based on these results, we will examine how we can incorporate critical evaluation activities in future versions of our teaching strategy. In the context of this study, we define critical evaluation by students as their ability to:

- Judge the quality of the main conclusion (based on the supports and refutations[2]) in a research article;
- Come up with additional (i.e. not mentioned by the authors) counterarguments;
- Decide about competing knowledge claims from different sources.

The participants were students who followed our module Reading Research Articles in 2010. At the end of the module, we gave the students two articles containing claims that seem to contradict each other about the same biological topic. It is worthwhile to introduce students to these types of problems because it teaches them to work with competing knowledge claims. This is important, because of the pluralistic nature of science:

> As it often happens in science, more than one claim may be an acceptable explanation for the same phenomenon. Arguments are the preferential tool for resolving controversies in science, but when scientists have to make choices, evidence is never totally determinate, nor are arguments overwhelmingly convincing. More than one alternative is not just possible but often also plausible. (Sampson & Clark, 2006, p. 655)

After reading each article, the students made assignments that allowed us to gain some insights into their critical evaluation skills. The assignments contained questions that related to Gillen's (2006) three aspects of reading a research article. To gain insight into the first aspect (understanding the experimental narrative), we asked students questions about the articles' experimental procedures. Furthermore, students identified the different rhetorical moves in the articles, as understanding the rhetorical structure helps with the recall and comprehension of scientific texts (e.g. Samuels et al., 1988). To gain insight into the second and third aspects (interpret-

[2] We focus on refutations, because they refute counterarguments and thus indirectly support the main conclusion.

ing the text and criticism), we asked students evaluative questions about the main conclusions, counterarguments, and refutations. We focused on these moves because they are located in the Discussion section and are therefore the end products of authors' reasoning. The main conclusion forms in a sense the thesis of the authors, while the counterarguments form the antithesis. Because refutations weaken counterarguments, they support the main conclusion/thesis. Finally, we asked students to resolve the contradictory evidence in the articles.

We collected all students' written answers to these assignments and recorded tutor group meetings in which they discussed these assignments. Furthermore, students completed a self-assessment at the beginning and end of the module. Our main research question was:

1. What is students' ability to resolve the contradictory evidence in two articles while following a teaching strategy based on the identification of rhetorical moves?

We also had three additional research questions:

2. What is students' ability to understand the experimental narratives of the articles?
3. What is students' ability to critically evaluate each article?
4. What is students' opinion about their own progression in their ability to critically evaluate a research article?

7.3 Method

Educational Setting and Participants

The data collection took place in 2010, during the course Biomedical Research (see Section 1.9). During the course, students – who were novice readers of research articles – followed the eight-weeks module called Reading Research Articles that aimed to introduce them to research articles by teaching them how to identify rhetorical moves and their relations (see Chapters 4, 5, and 6, which describe the 2009 edition of the module). An important difference between the 2009 and 2010 edition is that the latter did not contain a pre-test and post-test. This gave us two extra weeks, in which the students made the critical evaluation assignments. The 153 students were randomly placed into 18 tutor groups. For this study, we used data from 50 students who belonged to eight randomly selected tutor groups. The 50 students were approximately 18–20 years old. Sixteen students were male and 34 students were female.

Research Articles

The two research articles we used presented claims that seem to contradict each other (on which we will elaborate in the Results section). The articles were – in our opinion – well-reasoned texts that contained no major flaws. Both articles studied the effects of treatment with vitamin C, an anti-oxidant, on respectively chronic heart failure (CHF) and diabetes. The onset and progression of these two diseases are allegedly influenced by oxidative stress. Article 1 ("Vitamin C Improves Endothelial Function in Patients with Chronic Heart Failure") was written by Hornig, Arakawa, Kohler, and Drexler (1998) and published in the journal *Circulation*. Article 2 ("Lack of Effect of Oral Vitamin C on Blood Pressure, Oxidative Stress and Endothelial Function in Type II Diabetes") was written by Darko, Dornhorst, Kelly, Ritter, and Chowienczyk (2002) and published in the journal *Clinical Science*.

The main bodies of Articles 1 and 2 contained, respectively, about 3,600 and 2,000 words. Article 1 contained four tables and two figures, while Article 2 contained only two tables and no figures. Readability of the articles was measured using the Flesch Reading Ease Score (see Section 2.3). Article 1 had a Flesch Reading Ease score of 37. Article 2 had a score of 39. This means that the readability of Articles 1 and 2 were comparable: they were both *difficult to read*.

Assignments

Each week there was a tutor group meeting for which the students had to make an assignment. The assignments described in this chapter were given to the students in the last two weeks of the module. In the previous weeks, students learned how to identify the different rhetorical moves of a research article (see Chapter 5). Students were made familiar with the articles' topics during lectures.

Students received Article 1 and an accompanying assignment in Week 6. This assignment consisted of three parts. Parts A and B were used to determine how well the students understood the articles, part C was used to determine how students evaluated the articles. In part A students identified the different rhetorical moves. They were asked to answer the following question:

A1) Identify in the article the following elements and write them down: motive, objective, main conclusion, implication, important supports, counterarguments, and refutations. Please include for each element the [article's] page number.

Part B of the assignment consisted of questions about some key terms regarding the experimental procedures described in Article 1 (the answers to these questions are not explicitly mentioned in the text):

B1) Why did the authors use (a) L-NMMA and (b) SNP?

B2) Why did the authors also study the "radial artery blood flow"?

In part C of the assignment students evaluated the article in a critical way. This part consisted of the following questions:

C1) Does the main conclusion follow from the presented data?

C2) Do the authors mention all counterarguments? If not, what do you miss?

C3) Are the refutations of sufficient quality? Why?

C4) Does this study warrant the administration of vitamin C pills to patients with chronic heart failure? Why?

Students read Article 1 and made the assignment as homework. They individually wrote down their answers and sent them in via e-mail within five days after the tutor group meeting. During the next tutor group meeting, in Week 7, students discussed the assignment and article. They also received feedback on their answers from the tutor. During the same meeting, students received Article 2 and an accompanying assignment. This assignment consisted of four parts. Parts A and C were identical to parts A and C of the previous assignment (except that in C4 "chronic heart failure" was replaced by "diabetes type II"). Part B consisted of the following question:

B3) Why did the authors measure 8-epiprostaglandin $F_{2\alpha}$-levels ?

Additionally, there was a part D, which was used to determine how students resolve the contradictory evidence in the articles. For part D, the students had to compare the outcomes of Articles 1 and 2. This part consisted of two questions:

D1) How do you explain that the current article does not show an effect of vitamin C while the previous article did show an effect?

D2) Looking at the data of the current and previous article: has vitamin C an effect on oxidative stress or not? Please explain your answer.

Again, the students handed in their answers via e-mail within five days after the tutor group meeting. The assignment and article were discussed during the next tutor group meeting in Week 8. Students also received feedback on their answers from the tutor via e-mail or written down on paper.

Tutor Group Meetings

In Week 7, we recorded a tutor group meeting in which students discussed Article 1 and their answers to the first assignment. A week later, in Week 8, we recorded a tutor group meeting in which the students discussed Article 2 and their answers to the second assignment. Because of logistical considerations, we recorded the meetings of different tutor groups. Respectively eight and ten students participated in the first and second tutor group meetings.

Questionnaires

Self-assessment. At the beginning and end of the module, the 50 students completed an on-line self-assessment in which they indicated how much they agreed to three different statements (see Table 7.8). The statements dealt with students' ability to understand the experimental narrative, their ability to interpret the results, and their ability to criticize a research article.

Evaluation of the group discussion about the two articles. At the end of the module, the 50 students completed a questionnaire to evaluate the module. This questionnaire contained 12 items about the tutor group meetings. One of these items was: I found the discussion about the two articles (which described the effects of vitamin C) during the final tutor group meeting useful. Via a 5-point rating scale, students could indicate how much they agreed with this statement. Students could also leave comments about the discussion during the final tutor group meeting.

Data Analysis

For the analysis of the rhetorical moves that were identified by students (part A), we categorized per move the sentences students mentioned as answer. Students quoted verbatim from the article in their answers, so no interpretation was needed to score the data. A second researcher and I identified the moves in the two articles and discussed them until agreement was reached. Students' answers were compared with the sentences we found. For the analysis of the supports, we determined per student to which tables and/or figures he or she referred explicitly or implicitly. The authors of both articles also used references to previous studies as supports. However, almost none of the students mentioned these references as supports, so we did not include them in our analysis.

The quality of students' answers to the questions of part B were analyzed by marking them as correct or incorrect.

For the analysis of the questions that belonged to parts C and D, I made for each question a coding scheme. I followed a grounded approach (Glaser & Strauss, 1967): during the first phase of the analysis process I made specific categories for the students' answers. These categories were continually refined. Then, during the second phase of the analysis process, I grouped the answers into more general categories. The resulting coding scheme was also used by a second researcher to categorize the answers. This led to some adjustments of the coding schemes. To determine the inter-rater agreement between our codes, we calculated Krippendorff's alpha (Hayes & Krippendorff, 2007). Krippendorff's alpha was 0.83 for Question C1 (based on the analysis of 100 items), 0.96 for C2 (100 items), 0.93 for C3 (100 items), 0.97 for C4 (100 items), 0.90 for D1 (50 items), and 0.98 for D2 (50 items). These numbers indicate a high inter-rater agreement.

For the analysis of the tutor group meetings, we made transcriptions of the two audio-recorded tutor group sessions. Fragments related to students' views on the assignments or the articles were marked. Then, these fragments were summarized and combined into a coherent narrative.

A Wilcoxon signed-rank test was used to compare the self-assessments that were filled out at the beginning and end of the course.

For the analysis of the evaluation questionnaire, we determined the frequencies of students' responses.

7.4 Results

Analysis of Students' Answers to Questions about Article 1

Description of content. Article 1 describes a study in which the effects of vitamin C on the dilation of blood vessels (the flow-dependent dilation or FDD) were studied in patients with CHF. This dilation is decreased in patients with CHF. Decreased dilation is a symptom of endothelial dysfunction. Vitamin C was administered orally (2.0 g daily for 4 weeks) or intra-arterial (25 mg/min over 10 minutes). The authors present their main conclusion at the end of the Discussion section: "…the present study demonstrates that endothelial dysfunction in patients with congestive heart failure can be improved and normalized by acute intra-arterial as well as by chronic oral treatment with (…) vitamin C" (p. 368). The authors also present an additional conclusion (additional because it does not directly relate to the objective): "Furthermore, this study indicates that the beneficial effect of vitamin C on FDD in humans is mediated by an increased availability of NO [nitric oxide], since the portion of FDD mediated by NO was increased by vitamin C" (p. 366). In other words: the improvement they observed is due to the increased availability of nitric oxide (a vasodilator), because the positive effects of vitamin C disappeared when L-NMMA (a nitric oxide synthase inhibitor) was administered. It seems that the authors are not so sure of this statement, as they use the word "indicate," a tentative reporting verb. Additionally, the authors speculate that vitamin C increases the availability of nitric oxide by lowering the levels of oxidative stress (because they assume that oxidative stress degrades nitric oxide).

Identification of rhetorical moves (Part A). Table 7.1 shows which sentences were mentioned most frequently as objective, motive, main conclusion, and implication by the students. These sentences corresponded with our own analysis of the article. It has to be noted that only 31 of the 50 students identified the same sentence as main conclusion as we did. However, the other 19 students almost always identified a sentence that could also be called the main conclusion, although formulated in a less generalized form: "…the impaired FDD in patients with CHF is improved by the antioxidant vitamin C both after intra-arterial administration and 4 weeks of oral therapy, whereas FDD was not affected by vitamin C in healthy volunteers" (p. 366).

The five counterarguments that were most often mentioned by students as answers to the first assignment are shown in Table 7.2 (together with the refutation that was most often mentioned by these students). In our own analysis, we only designated the first two sentences of Table 7.2 as counterarguments ("We cannot exclude…" and "It is also unlikely…"). We found no other counterarguments in the article.

The counterargument that was most often mentioned, by 38 students, describes the possibility that the effects of vitamin C are indirect. Fifteen of the 38 students gave as refutation the five sentences which follow the counterargument and in which the authors argue that it is unlikely that the effect of vitamin C is indirect (this answer corresponds with our analysis). It has to be noted that the other 23 students also gave refutations; however, these students only quoted a selection of the before-mentioned five sentences. Regarding the second counterargument, six of the seven students wrote down as refutation the second part of the sentence that contained the counterargument (which corresponds with our analysis).

Supports were analyzed by determining to which data sources each student referred (e.g. by mentioning a specific figure/table or by mentioning a

Table 7.1. The sentences that were most often mentioned by students for each rhetorical move. The number of students who mentioned these sentences is shown between brackets.

Rhetorical move	Article 1	Article 2
Motive	"In this respect, there is evidence that radical formation is increased in patients with CHF,[8] raising the possibility that endothelial dysfunction in CHF is, at least in part, due to increased inactivation of NO by oxygen free radicals" (p. 363). (*n* = 32)	"It is possible that a water-soluble antioxidant such as vitamin C, which is capable of scavenging superoxide anions ($O^{\bullet}_2{}^{-}$), is more effective than vitamin E in both protecting endothelium-derived nitric oxide and preventing lipid peroxidation" (p. 340). (*n* = 34)
Objective	"Accordingly, the present study was designed to determine the effect of vitamin C on NO-mediated FDD in patients with CHF, both after acute intra-arterial administration and chronic oral treatment with vitamin C" (p. 363). (*n* = 49)	"The object of the present study was to determine whether short-term oral administration of vitamin C reduces oxidative stress, improves endothelial function and lowers blood pressure in patients with uncomplicated Type II diabetes" (p. 340). (*n* = 48)
Main conclusion	"In conclusion, the present study demonstrates that endothelial dysfunction in patients with congestive heart failure can be improved and normalized by acute intra-arterial as well as by chronic oral treatment with the antioxidant vitamin C" (p. 368). (*n* = 31)	"In conclusion, treatment with vitamin C (1.5 g daily) for 3 weeks produced no significant improvements in oxidative stress, blood pressure or endothelial function in patients with Type II diabetes" (p. 343). (*n* = 45)
Implication	"While our initial observations suggest that this beneficial effect may be sustained during longterm supplementation, this finding needs to be confirmed in large-scale clinical trials" (p. 368). (*n* = 29)	No answer. (*n* = 15)

sentence which referred to a specific figure/table). Three students did not refer to any data source at all. The other 47 students primarily referred to Table 2 (*n* = 28), Figure 1 (*n* = 35), and Figure 2 (*n* = 35). These inscriptions (see Section 2.3) show that vitamin C has positive effects on the

FDD. Students did not often refer to Table 1 (n = 2), Table 3 (n = 9), and Table 4 (n = 4). Table 1 contained the characteristics (age, weight, etc.) of the subjects, and Tables 3 and 4 show that vitamin C has no effect on the radial artery blood flow.

Experimental procedures (Part B). Three questions were asked about the content of the article: (B1) Why did the authors use (a) L-NMMA and (b) SNP? (B2) Why did the authors also study the "radial artery blood flow"? The authors used L-NMMA to determine the effects of nitric oxide on vasodilation. SNP was used to determine the effects of vitamin C on endothelial-independent relaxation. The radial artery blood flow was studied to rule out the possibility that the effects of vitamin C are simply caused by an improved blood flow. Although most students gave correct answers to the first two questions, it seems that the last question was quite difficult for the students. Only 6 of the 50 students gave the correct answer. Most other students thought, incorrectly, that the radial artery blood flow was measured to determine the effects of vitamin C. For example, one student wrote: "With this you can see if there is vasodilatation or vasoconstriction and thus measure the effects of the substances you give to the patients."

Evaluation of main conclusion (Part C). When asked if the main conclusion of Article 1 followed from the presented data (Question C1), 41 of the 50 students said yes and four students said no. Five students gave ambivalent or unclear answers. The four students who disagreed with the authors' main conclusion argued that some variables were not influenced at all by vitamin C in the experiment. These students argued that this result does not justify the main conclusion. However, this is not a correct interpretation of the results and underlines our earlier observation that students had difficulty with understanding why the radial arterial blood flow was measured. The fact that this variable does not differ between the experimental groups supports the main conclusion of the authors, because it means that the effects of vitamin C are not caused by a change in blood flow.

The 41 students who did say that the main conclusion followed from the presented data, often referred to the text's inscriptions. For example, one student wrote: "Yes, the main conclusion follows from the presented data. Especially the data that is presented in Figure 2 shows clearly that endothelial dysfunction of patients with chronic heart failure could improve with acute and chronic vitamin C treatment."

Additional counterarguments (Part C). We asked the students if they could come up with additional counterarguments (Question C2). In the case of Article 1, 16 students thought that all relevant counterarguments were

Table 7.2. The five text fragments of Article 1 that were most often mentioned by students as counterargument, and the text fragments that were subsequently most often mentioned as refutation. The number of students who mentioned these text fragments is shown between brackets.

Text fragments mentioned as counterargument	Text fragments mentioned as refutation for this counterargument
"We cannot exclude the possibility that vitamin C might directly improve heart failure (thereby secondarily improving vascular function) and not specifically endothelium-mediated vasodilation" (p. 367). (*n* = 38)	"However, the acute effect of vitamin C on endothelium-mediated vasodilation cannot be explained by changes in the severity of heart failure. (…) It is highly unlikely that vitamin C given acutely into the brachial artery results in a major improvement in central hemodynamics in patients with NYHA class III heart failure. (…) Taken together, we think that the hypothesis that vitamin C improves heart failure per se is unlikely and cannot explain our results" (p. 367). (*n*=15)
"It is also unlikely that a correction of an absolute vitamin C deficiency may explain our findings…" (p. 367). (*n* = 7)	"…because it has recently been shown that there is no correlation between baseline vitamin C plasma levels, endothelial dysfunction, or improvement with treatment.[15]" (p. 367). (*n* = 6)
"These first preliminary observations during long-term supplementation need to be confirmed in a larger group of patients…" (p. 367). (*n* = 4)	"…however, if confirmed, they may have important clinical implications" (p. 367). (*n* = 4)
"The present study was not designed to elucidate the underlying mechanism(s) leading to increased oxidative stress in CHF" (p. 368). (*n* = 4)	"However, there is evidence that angiotensin II, whose plasma and tissue levels are typically elevated in CHF, activates NADH/NADPH-driven oxidases located within the vascular wall[21] that appear to be the main enzymes responsible for vascular synthesis of radicals within the vessel wall. However, other factors may be involved as well, such as increased levels of cytokines such as tumor necrosis factor-α, which in turn may enhance oxidative stress" (p. 368). (*n* = 3)
"In addition, an effect of vitamin C on vascular smooth muscle function rather than the endothelium appears to be unlikely…" (p. 367). (*n* = 3)	"…because vitamin C per se did not affect radial artery diameter and blood flow and the vasodilator response to SNP was similar in patients treated with vitamin C and placebo" (p. 367). (*n* = 2)

Table 7.3. Frequencies of students' answers to question C2 for Articles 1 and 2: "Do the authors mention all counterarguments? If not, what do you miss?" Answers could be classified into more than one category (which explains why the total number of answers per article is greater than 50).

Answer category	Article 1 (*n*)	Article 2 (*n*)
Yes, all counterarguments are mentioned	16	21
No, the sample group is too small or has the wrong composition	13	10
No, there is something wrong with the measuring methods	2	3
No, the influence of diet is not accounted for	4	7
No, the influence of drug use is not accounted for	3	4
No, the influence of other factors (besides drug use/diet) is not accounted for	2	5
No, the authors' explanation for their results could be incorrect	8	0
No: other reasons	6	7
Unclear or ambivalent answer	2	0

mentioned by the authors (Table 7.3). Counterarguments most often mentioned by students related to the used sample groups: these were too small or had the wrong composition. Students also regularly came up with counterarguments that related to the influence of confounding variables (diet, drug use, et cetera). A number of students mentioned a counterargument that sheds doubt on the authors' explanation of their results (nitric oxide availability improved because vitamin C reduced oxidative stress).

Evaluation of the article's refutations (Part C). When asked about the refutations of Article 1 (Question C3), 34 students said that they were of sufficient quality (Table 7.4). These 34 students argued a number of times that the quality was sufficient because the authors refer to other studies. For example, one student wrote: "The counterarguments are refuted by results from earlier studies. This makes the refutations powerful, because you see that it is based on something, and it is not made up by the researchers themselves." Students who were critical about the quality of the refutations most often mentioned that they were not sufficiently supported. As one student wrote: "The refutations are not sufficiently supported. Like in this example: 'It is highly unlikely that vitamin C given acutely into the brachial artery results in a major improvement in central hemodynamics in patients with NYHA class III heart failure.' They only say that it is unlikely but not why it is unlikely."

Administration of vitamin C (Part C). We asked if the study justified the administration of vitamin C to patients with chronic heart failure (Question C5). Thirty-five of the 50 students said the study did not justify the administration of vitamin C (Table 7.5). Many students gave as explanation for their answer that the sample group was too small (*n* = 21), vitamin C could have harmful effects (*n* = 8), or the long-term effects are unknown (*n* = 18). Ten of the 50 students stated that the study did justify the administration of vitamin C (even though the authors themselves state that more research is needed). These students argued that the study showed that vitamin C has a positive effect. Furthermore, the perceived harmlessness of vitamin C was also used to support their answer. As one student wrote: "Yes, this study shows that there are significant differences between vitamin C-users and non-vitamin C-users. There were no negative effects found. This means: [vitamin C] is in any case harmless and it does very likely a lot of good. What's more, vitamin C is for every human important."

Tutor group meeting about Article 1. During the tutor group meeting about Article 1, the tutor discussed each different question with the students. At the beginning of the meeting, the students claimed that Article 1 was not particularly difficult to understand.

A number of students voiced critical remarks about the quality of the article. As one student said: "I found the article not very good." One other

Table 7.4. Frequencies of students' answers to question C3 for Articles 1 and 2: "Are the refutations of sufficient quality? Why?" Answers could be classified into more than one category.

Answer category	Article 1 (*n*)	Article 2 (*n*)
No answer	2	5
Yes, the refutations are of sufficient quality	34	23
No, the refutations are not sufficiently supported	7	8
No, the refutations are not sufficiently mentioned	3	6
No, the authors' words express too much uncertainty	3	5
No, the authors state that there is more research needed	6	0
No: other reasons	4	7

Table 7.5. Frequencies of students' answers to question C4 for Articles 1 and 2: "Does this study warrant the administration of vitamin C pills to patients with [chronic heart failure/diabetes type II]? Why?" Answers could be classified into more than one category.

Answer category		Article 1 (*n*)	Article 2 (*n*)
Ambivalent answer		5	2
Yes, this study warrants the administration of vitamin C ($n = 10$ for Article 1, $n = 1$ for Article 2)	Yes: student does not state why	1	0
	Yes: because vitamin C is not harmful	5	1
	Yes: because vitamin C has a positive effect	8	0
	Yes: other reason	1	0
No, this study does not warrant the administration of vitamin C ($n = 35$ for Article 1, $n = 47$ for Article 2)	No: because the sample group was too small/there is more research needed in a larger group	21	3
	No: because there can be harmful effects/these harmful effects should be investigated first	8	13
	No: long-term effects are unknown/these long-term effects should be investigated first	18	3
	No: intra-arterial administration is not the same as chronic oral treatment (only applicable for Article 1)	3	0
	No: vitamin C had no clear effect	2	44
	No: other reasons	8	1
	No: student does not state why	0	1

student agreed with her. They both thought that the conclusions were not sufficiently justified by the data. They pointed out that the authors state in the Discussion section: "…we think that the hypothesis that vitamin C improves heart failure per se is unlikely and cannot explain our results" (p. 367). According to these two students, this statement contradicts the authors' conclusion. Furthermore, the students thought that the use of hedges (see Section 1.5) made the authors' argument unconvincing. One student said: "They throw a lot of hypotheses in between [the text]. Like: we think this is unlikely. Things why they think it is unlikely. Not really convincing." Another student added: "They say: it could have happened because of this and this. Then I think: yes, but it could also have happened because of that. That occurred a lot."

Students also said that it was difficult to come up with additional counterarguments, especially because they lacked the necessary content knowledge. One student said: "We don't know what all these substances do and which substances can possibly be left out and which other substances can have influence." Students were critical about the refutations. They thought the authors did not justify them sufficiently with evidence. For example, one student remarked about a refutation: "They don't support it. They often use words like may or might." Then, the tutor said: "The use of may and might is almost inevitable." (The tutor probably meant that it is difficult in science to state things with absolute certainty). The student replied: "Yes, okay, but it's not really convincing."

Analysis of Students' Answers to Questions about Article 2

Description of content. In Article 2 the oral administration of vitamin C (1.5 g daily for 3 weeks) to Type II diabetes patients was studied, in particular its effects on blood pressure and levels of oxidative stress. To determine the levels of oxidative stress, the plasma concentrations of 8-epi-prostaglandin $F_{2\alpha}$ (a marker for oxidative stress) were measured. At the beginning of the Discussion section the authors state their main conclusion. At the end of the Discussion section, they repeat this conclusion: "In conclusion, treatment with vitamin C (1.5 g daily) for 3 weeks produced no significant improvements in oxidative stress, blood pressure or endothelial function in patients with Type II diabetes" (p. 343). The authors state that they did not expect that vitamin C had no effect on oxidative stress, because "type II diabetes is well recognized to be associated with increased oxidative stress" (p. 341). That is why the authors try their best to come up with explanations for this unexpected result. For example, they speculate that intracellular oxidative stress is more important than extracellular oxidative stress in diabetes. This could explain the lack of effect, because studies have shown that vitamin C is more effective in reducing extracellular oxidative stress than intracellular oxidative stress. Furthermore, they also list several counterarguments. For example, the authors postulate that the duration of treatment was too short.

Identification of rhetorical moves (Part A). The sentences students mentioned the most as objective, motive, main conclusion, and implication corresponded with our own analysis of the article (Table 7.1). Because Article 2 contained no clear implication, 15 of the 50 students left their answer blank. The other students tried to come up with an answer. The sentence which was most often identified as an implication (by 14 students) was: "We cannot, however, exclude the possibility that a longer duration of treatment with a different dose of vitamin C may be effective in lowering oxidative stress and/or improving endothelial function" (p. 342).

The five counterarguments most often mentioned by students are shown in Table 7.6. We designated only the first two sentences as counterarguments ("We cannot, however, exclude..." and "We cannot exclude a..."). In the case of the first counterargument, 4 of the 11 students were not able to give a refutation. Students were more successful in giving a refutation for the second counterargument: 9 of the 10 students wrote down the second part of the sentence containing the counterargument (which corresponds with our analysis). The other three sentences in Table 7.6 were no counterarguments but an explanation for the results or a statement explaining why the results are unexpected.

We designated two other sentences as counterarguments in the article: one mentioning a discrepancy with the results of a similar study (refutation of the authors: the other study used a different dose), and one mentioning that there still could be a small effect on blood pressure (the authors give no refutation for this counterargument). These two counterarguments were mentioned respectively four and three times by the 50 students.

When asked to identify the supports, 46 of the 50 students referred to Table 2, while 2 students referred to Table 1. Table 1 shows the baseline characteristics of the subjects while Table 2 shows the effects of treatment with vitamin C. It has to be noted that the authors themselves do not refer to Table 1 in the Results section. Two students did not refer to any data source at all when giving supports. One of these students did not answer the question and the other student gave only conclusions.

Experimental procedures (Part B). There was only one question asked about the experimental procedures of Article 2: B3) Why did the authors measure 8-epiprostaglandin $F_{2\alpha}$-levels? These levels were measured because they are a marker for oxidative stress. Forty-seven students gave a correct answer.

Evaluation of main conclusion (Part C). Regarding the evaluation of Article 2's main conclusion, all 50 students agreed that this conclusion followed from the presented data. As one student wrote: "Yes, because the results of Table 2 show no significant difference between placebo and vitamin C."

Table 7.6. The five text fragments of Article 2 that were most often mentioned by students as counterargument, and the text fragments that were subsequently most often mentioned as refutation. The number of students who mentioned these text fragments is shown between brackets.

Text fragments mentioned as counter-argument	Text fragments mentioned as refutation for this counterargument
"We cannot, however, exclude the possibility that a longer duration of treatment with a different dose of vitamin C may be effective in lowering oxidative stress and/or improving endothelial function" (p. 342). (n = 11)	"It is also possible that high doses may have an adverse effect: there is evidence that high concentrations may cause oxidative damage to DNA [29] and may indeed lower NO bioactivity [30]" (p. 342-343). (n = 4) or no refutation (n = 4)
"We cannot exclude a small effect of vitamin C on endothelial function…" (p. 342). (n = 10)	"…but the 95% confidence interval in the present study excluded a difference similar to that observed previously between non-diabetic and diabetic subjects [7,17]" (p. 342). (n = 9)
"Acute administration of oral vitamin C does improve endothelial function in smokers [22] and patients with ischaemic heart disease [23], demonstrating that beneficial effects of oral vitamin C are potentially attainable" (p. 342). (n = 8)	"It is possible, however, that the mechanism of this acute effect does not involve an antioxidant effect [24]" (p. 342). (n = 3)
"The lack of effect of vitamin C on blood pressure in the present study may be explained by relatively well-controlled blood pressure and/or concurrent antihypertensive treatment in our subjects" (p. 343). (n = 8)	"However, it is also consistent with the possibility that vitamin C does not lower oxidative stress in patients with Type II diabetes" (p. 343). (n = 5)
"This finding is unexpected, because Type II diabetes is well recognized to be associated with increased oxidative stress, plasma concentrations of 8-epi-$PGF_{2\alpha}$ are elevated in Type II diabetes [6,7], and vitamin C lowers 8-epi-$PGF_{2\alpha}$ levels in other conditions associated with oxidative stress" (p. 341). (n = 5)	"It is notable that baseline vitamin C concentrations were within the accepted normal range [22] in our subjects, and it is possible that supplementation with vitamin C is only effective in reducing plasma 8-epi-$PGF_{2\alpha}$ when plasma vitamin C concentrations are low" (p. 342). (n = 4)

Additional counterarguments (Part C). Twenty-one of the 50 students thought the authors mentioned all counterarguments. The additional counterarguments of the students were categorized and are shown in Table 7.3. As with Article 1, the counterarguments most often mentioned by students related to the used sample groups and the influence of confounding variables.

Evaluation of the article's refutations (Part C). When asked about the refutations of Article 2, 23 of the 50 students said that they were of sufficient quality (Table 7.4). Again, a number of students stated that the refutations were not backed with evidence. This time, some students also took issue with the choice of words of the authors. As one student wrote: "I find the refutations of poor quality, because they use many times words like may, possible etc. So they are not really certain about their case."

Administration of vitamin C. We asked if the study justified the administration of vitamin C to patients with diabetes (Table 7.5). Only 1 of the 50 students was convinced that the study did justify the administration of vitamin C to patients with diabetes. These numbers are not surprising, because the study showed no positive results regarding the effectiveness of vitamin C. Most students used this reason to justify their answer. A small minority also brought forward the point that vitamin C can have harmful effects.

Tutor group meeting about Article 2. During the tutor group meeting in which the answers to the second assignment were discussed, the students did not report any major difficulties with understanding Article 2. Students all agreed with the authors' main conclusion. The students had no major issues with the article's methodology. When the tutor asked the students if the authors should have come up with additional evidence for their main conclusion, one student thought that it was not necessary for authors to elaborate on negative results: "I can understand that you elaborate if you have positive results. But now it was all… [there was no] real difference." Another student said: "It was a little bit an easy study. Because they… so, they have studied three things. (…) The oxidative stress could have turned out differently if they had looked at another [substance], instead of [8-epi-prostaglandin $F_{2\alpha}$], for example."

As with the other article, students found it difficult to come up with additional counterarguments. As one student said: "Yes, I thought it was difficult, to really [come up with a counterargument] because I thought the article was a little strange." She found it a strange article because the authors list a great number of possible reasons why their results are unexpected. "It seems like they are weakening their own results," she said. Another student suggested: "Maybe [they do this] to avoid a heated discussion. Because they are really contradicting all those other studies."

Analysis of Students' Answers to Questions about the Comparison of Articles 1 and 2

Difference in effects (Part D). We asked students to compare both articles and explain why vitamin C had an effect in Article 1, and no effect in Article 2 (Question D1). Students' written answers could be divided into three groups. Thirteen of the 50 students thought that the different experimental designs (dose, administration method, duration of treatment, etc.) could explain the discrepancy between the two studies. Twenty-two students thought that the difference in results could be explained by the fact that different diseases were studied. Thirteen students thought it was a combination of both. The answers of two students could not be classified.

Effect on oxidative stress (Part D). Finally, we asked students to consider the data of both articles and assess if vitamin C has an effect on oxidative stress (Question D2). Twenty-eight of the 50 students thought that the experimental data in the two articles show that vitamin C had an effect on oxidative stress, even though the study described in Article 1 did not directly measure oxidative stress and the study described in Article 2 could not find an effect. Eleven students thought the opposite, and 11 students gave an unclear/ambivalent answer. The justifications of the students were categorized and are shown in Table 7.7. Of the students who gave an affirmative answer, 11 gave as justification the speculation – mentioned in Article 2 – that vitamin C is only effective in diseases where extracellular stress plays a role (such as CHF). Seven students gave as justification the references that are mentioned in Article 2 (even though we asked the students to base their answer on the data mentioned in the two articles). These are references to studies that show that vitamin C is able to reduce oxidative stress – in contrast to the results of Article 2. Of the students who gave a negative answer, only two gave as justification that oxidative stress was not measured in Article 1 – which seems the correct answer to us.

Self-Assessment

The self-assessment filled out by the 50 students at the beginning of the module was compared to the one they filled out at the end. Students' agreement with the three different statements of the self-assessment increased significantly (Table 7.8).

Table 7.7. Frequencies of students' answers to question D2: "Looking at the data of the current and previous article: has vitamin C an effect on oxidative stress or not? Please explain your answer." Answers could be classified into more than one category.

Answer category		*n*
No answer/ambivalent answer (*n* = 11)		
Vitamin C has an effect on oxidative stress (*n* = 28)	Vitamin C has an effect, but only on extracellular stress/in CHF patients	11
	It follows from the references mentioned in Article 2	7
	It follows from the conclusion/results from Article 1	9
	The two studies had a different design	3
	The effect of vitamin C is sometimes very small	2
	Other justification	7
Vitamin C has no effect on oxidative stress (*n* = 11)	Oxidative stress was not measured in Article 1	2
	Results of Article 1 were inconclusive	5
	Student gave no justification	1
	Other justification	4

Evaluation of the Group Discussion about the Two Articles

Students were asked during the evaluation how much they agreed with the following statement: I found the discussion about the two articles (which described the effects of vitamin C) during the final tutor group meeting useful. Of the 42 students who completed the evaluation, 2 answered *strongly disagree*, 1 answered *disagree*, 9 answered *neither agree nor disagree*, 26 answered *agree*, and 4 answered *strongly agree*.

Some students added that they learned how important it is to read research articles in a critical way. For example, one student wrote: "First I believed everything was true, but after the assignments I looked at conclusions in a more critical way." Another student wrote succinctly: "Academic thinking = critical thinking." Other students argued that the assignments were not really useful, because first-year students lack the specific knowledge that is needed for the critical evaluation of research articles. One student wrote: "As a first-year student you are not able to criticize such an article like a journal does."

Table 7.8. Results of students' self-assessment given at the beginning and end of the module (in percentages).

I am able to…		Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	Signif-icancy (beginning versus end)
…understand how the experiment was done.	Begin-ning	0	18	22	60	0	$z = -2.3$, $p = 0.024$
	End	0	5	13	79	3	
…interpret the presented data by myself.	Begin-ning	0	24	52	21	2	$z = -4.4$, $p < 0.001$
	End	0	3	23	69	5	
…judge the conclusion's supports on their correctness.	Begin-ning	4	49	38	9	0	$z = -4.2$, $p < 0.001$
	End	0	15	54	31	0	

7.5 Discussion

Students read two research articles with contradictory evidence in this study. We investigated how these students evaluated each article and how they resolved the contradictory evidence (the second and third aspects mentioned by Gillen, 2006).

Both the answers to the questions about the content of the articles and students' own judgments indicate that students sufficiently understood the experimental narratives of the articles (the first aspect mentioned by Gillen, 2006). They could give correct answers to questions about the experimental procedures. As stated above, students had difficulty with one question about the experimental procedures described in Article 1. This influenced their interpretation of the results: some students thought that the results contradicted the authors' main conclusion while the opposite was true.

The sentences students identified as motive, objective, main conclusion, and implication corresponded with our own analysis. Regarding the supports, most students found the most important ones, although in the case of Article 1 they tended to ignore the radial blood flow data. We identified in Article 1 two counterarguments. The majority of students identified one of these two counterarguments. We identified in Article 2 four counterarguments. However, only a minority of students identified all these counterarguments. So, students were reasonably successful in identifying the rhetorical moves in the article (with the exception of the counterarguments and refutations). The analysis of the tutor group meetings corroborated our finding that students had probably no major difficulties with understanding the experimental narratives of the two articles.

Our results show that students were more critical about Article 1 than Article 2. Students disagreed more often with the authors' main conclusion of Article 1. Additionally, students came up with more additional counterarguments for Article 1 than for Article 2 (Tables 7.3 and 7.6), probably because the results of the latter seemed less ambiguous. It could also be argued that students were more critical about Article 1 because this text was more understandable for students. However, this is not likely. Article 1 was in our opinion much more complex than Article 2 as it contained more different data sources.

Allchin (2001) describes four error types in science: material (e.g. the used materials were improper), observational (e.g. sampling errors or insufficient controls), conceptual (e.g. reasoning flaws), and discoursive (e.g. fraud). In the case of Article 1 some students were able to come up with counterarguments that related to a specific conceptual error (i.e. the authors' rather weak explanation for their results) but on the whole students' counterarguments related to observational errors. For instance, students were often critical of the small sample size. An explanation for this finding could be that counterarguments related to observational errors probably require the least amount of prior knowledge. It is also questionable if students' criticisms regarding the sample size are valid: a small sample size does not necessarily mean that the study is sub-par.

We observed that a number of students did not appreciate some of the rhetorical strategies used by the authors of the two articles. For example, students were critical about the authors' refutations in Article 2 because of their frequent use of hedges such as "may" or "might." The tutor group meetings showed that students saw this as a weakness. So, it seems that they are not aware of or do not appreciate the vital role of hedges in scientific communication. A possible cause for this finding is that students are not used to hedges, because they seldom encounter them in textbooks and popular-scientific articles. Additionally, students stated during the tutor group meetings that they did not understand why authors mention counterarguments and thus discredit their own study. It seems that they do not see the value of anticipating to possible criticism. These results correspond with our earlier observation – described in Chapter 5 – that students did not understand why authors would use counterarguments if they cannot refute them.

Regarding resolving the contradictory evidence, a small majority of the students thought that the data in the two articles showed that vitamin C reduces levels of oxidative stress, even though the study described in Article 1 did not directly measure oxidative stress and the study described in Article 2 could not find an effect. The authors of Article 1 themselves are very careful in their formulation regarding this subject: "Our results support the concept (…) that vitamin C exerts its antioxidant properties within the vasculature by directly scavenging oxygen derived free radicals such as superoxide anion or hydroxyl radicals" (p. 367-368). It seems that the negative conclusion of Article 2 (which none of the students disputed) made less of an impression than the speculative remarks of Article 1, because the students persisted in their belief that vitamin C has an effect on oxidative stress.

Our results could have been influenced by our wording of Question D2. We asked the students to look at the data of the articles. It seems some students thought that the articles' data included references to previous research (while we meant the experimental data generated by the authors themselves).

Most students stated that these small-scale studies did not justify the administration of vitamin C to patients (see Table 7.6). Students are aware that it is not allowed to extrapolate the results of small-scale studies to clinical practice. This is in line with Article 1, in which the authors state quite explicitly that more research is needed before vitamin C can be given to patients.

Hogan and Maglienti (2001) found that scientists and non-scientists (adults and middle school students) differ in how they judge conclusions drawn from evidence. Scientists thought that conclusions should cohere with all available evidence. Non-scientists also sometimes applied this criterion, but they also thought that conclusions should be plausible. For example, non-scientists were positive about a conclusion because it made sense; they did not give much thought to the evidence that was used to support the conclusion. In other words, scientists and non-scientists have other epistemological criteria (see Section 7.1). Because the effectiveness of vitamin C is plausible, students – who are not (yet) scientists – possibly used the epistemological criterion of plausibility when they had to resolve the contradictory evidence in the two articles.

Research suggests that epistemological beliefs (and thus epistemological criteria) will gradually become more sophisticated during students' training at university. For example, Jehng et al. (1993) found some clear differences between graduate and undergraduate students (from both the hard and the soft fields):

> …students at the graduate level tend to rely on their own independent reasoning, they believe that knowledge is likely to be uncertain, and that learning is not necessarily an orderly process, while undergraduate students tend to believe that knowledge is certain, that knowledge is handed down by authority figures and that the learning process is regular and orderly. (p. 34)

Critically evaluating research articles is a difficult and complex task for novice readers. As stated in Section 7.1, it demands knowledge of epistemological criteria and a so-called critical spirit. Furthermore, prior knowledge is important for the critical evaluation of research articles. As Bazerman (1988) observed, scientists judge the truth or quality of a research article only if they possess relevant prior knowledge. In our study, critical evaluation was even more difficult because the articles did not contain obvious major flaws. It is nevertheless valuable to let undergraduate students work on this type of evaluative assignments. By stimulating students to evaluate research articles, they learn that these articles are persuasive and that it is important to read with a critical stance. In other words, the students will acquire a critical spirit. In addition, they get acquainted with certain epistemological criteria.

We think that the SAM scheme can help students with the critical evaluation of research articles. As this study shows, SAM can direct students to important elements of an article's argument. For example, it is questionable if students would have given much attention to the refutations if they were assigned to answer only questions about the content.

If critical evaluation tasks will be incorporated into future editions of our module, we recommend to stimulate further development of students' epistemological beliefs (e.g. by discussing epistemological criteria: what is acceptable evidence, strength of conclusions, etc.) and the use of certain rhetorical strategies by authors. Furthermore, it would be interesting to investigate in future studies how certain features of research articles (such as the inclusion of counterarguments) influence students' views about the correctness of the main conclusion.

Chapter 8
Discussion

8.1 Introduction

In this chapter, we will begin with a description of the way this study has been conducted (Section 8.2), followed by an overview of our conclusions (Section 8.3). We will combine results from the preceding chapters to answer the primary and secondary research questions that were presented in the first chapter. After that, we will reflect on our research design and teaching strategy (Section 8.4), give recommendations for further research (Section 8.5) and teaching (Section 8.6), and propose an undergraduate curriculum centered on scientific reading and writing (Section 8.7).

8.2 Aim of This Thesis and Research Design

The general aim of this thesis was to develop and evaluate a teaching strategy for introducing first-year undergraduate students to research articles. This was done via an educational design research approach (van den Akker, 1999). Using this approach, we determined the effectiveness of the design principles that underpinned our teaching strategy. As design principles we chose the focus on the rhetorical structure of research articles (teaching students how to identify rhetorical moves) and cognitive apprenticeship, a pedagogical approach that emphasizes the centrality of activity in learning and knowledge and highlights the context-dependent nature of learning (Brown, Collins, & Duguid, 1989; Collins, Brown, & Holum, 1991). This second design principle led to two derived design principles: authenticity (using authentic and non-adapted research articles) and interaction (letting students discuss research articles with each other and a tutor). These design principles were formulated after an preliminary investigation (Chapter 2) in which we compared students' ability to identify two rhetorical moves (conclusions and supports) with experts.

In this teaching strategy, students learned to read research articles with the help of a heuristic called Scientific Argumentation Model (SAM). This heuristic aimed to improve students' rhetorical consciousness: the knowledge about the rhetorical structure of a genre-specific text. SAM is based on concepts from genre analysis and argumentation theory (see Chapter 3). We incorporated a set of seven rhetorical moves in SAM that are typical of research articles: motive, objective, main conclusion, implication, supports, counterarguments, and refutations. In our teaching strategy, students learned to identify these moves in research articles. We provided students with clear descriptions and authentic examples of these seven moves to get students acquainted with them. Additionally, we provided students with a schematic representation of SAM (the SAM scheme), depicting the relations between the abovementioned seven moves. Understanding these relations in a research article will help students with understanding the authors' argument.

The teaching strategy was implemented in the module Reading Research Articles for undergraduate life science students. This module was part of the course Biomedical Research, which took place at the end of the first-year bachelor program. In this module (described in Chapters 4 and 5), students learned step-by-step to identify seven rhetorical moves in a research article (see Chapters 4, 5, and 6). SAM served as a heuristic for the students. In a later edition of the module, we added two assignments in order to study students' critical evaluation of research articles (described in Chapter 7).

In our research, we confined ourselves mainly to testing the first and most important design principle: the focus on rhetorical structure. However, our results allow us to draw tentative conclusions about the derived design principles of interaction and authenticity.

Design research is often iterative: after evaluation the design is revised and tested again. In this thesis, we did not extensively revise our design after evaluation. Our reason for this was that we were primarily interested in testing the validity of our design principles. Fine-tuning our educational materials had not our priority, as these materials are very context-specific: if they will be used in another academic setting, they would need to be substantially adapted anyway.

It was our hypothesis that the ideas from genre analysis and argumentation theory behind our teaching strategy will help students with reading primary literature. We expect that the identification of rhetorical moves is a reading strategy that is transferable to similarly structured research articles from related disciplines. If a student is able to identify the rhetorical moves in a cardiovascular research article, he or she can use this experience for identifying the rhetorical moves in a research article from a related discipline such as neurology.

8.3 Conclusions

The primary research question of this thesis was:

How can undergraduate life science students be taught to read research articles?

As described in Chapter 1, this primary research question led to five secondary research questions:

1. How does our teaching strategy influence students' ability to identify rhetorical moves in research articles?
2. Which features of rhetorical moves do students use for their identification?
3. How do students' reading behaviors change during the implementation of our teaching strategy?
4. What is students' ability to construct an argumentation scheme of a research article?
5. What is students' ability to critically evaluate research articles?

Research Questions 1 and 3 concern learning effects, while Research Questions 2 and 4 concern the ways students use SAM. Research Question 5 is related to critical evaluation, an important part of scientific literacy that was not explicitly taught in our teaching strategy. In the following paragraphs, we will answer these secondary research questions and, finally, the primary research question.

Secondary Research Question 1: How does our teaching strategy influence students' ability to identify rhetorical moves in research articles?

In this thesis, we analyzed students' ability to identify rhetorical moves in research articles (i.e. their rhetorical consciousness) multiple times. However, the purposes of these analyses differed. In a preliminary investigation (Chapter 2) students identified conclusions and supports so that we could characterize students' difficulties with reading research articles. Since these data were collected before the implementation of our teaching strategy, they are not used to answer Secondary Research Question 1. In Chapter 4, students identified rhetorical moves at the beginning (pre-test) and end (post-test) of our module so that we could determine their progression. In Chapter 5, identification of rhetorical moves was studied to determine the features students use for identifying them. In Chapter 6, the identification of rhetorical moves was studied in relation to the construction of an argumentation scheme. In Chapter 7, the identification of rhetorical moves served to determine students' ability to criticize and evaluate a text.

Student performances are summarized in Table 8.1. Chapter 4 (only the pre-test) presents the student performances at the beginning of our module. Chapters 5 and 6 present student performances during the middle of the module. Chapters 4 (post-test) and 7 present student performances at the end of the module.

Table 8.1 should be considered with some prudence, because of the different conditions that may influence the identification of moves by

Table 8.1. Percentages of students who identified rhetorical moves as described in different chapters.

	Motive	Objective	Main conclusion	Implication	Supports	Counterarguments and refutations
Chapter 4 (pre-test)[a]	45	98	24	45	32	27[d]
Chapter 5 (Article 1)[b]	80	100	n.a.	n.a.	n.a.	n.a.
Chapter 5 (Article 2)[b]	n.a.	100	40	80	n.a.	n.a.
Chapter 5 (Article 3)[b]	100	100	100	40	40	60
Chapter 5 (Article 4)[b]	80	80	0[e]	60	60	60
Chapter 6	44	100	100	14	n.a.[c]	51
Chapter 7 (Article 1)	64	98	62	58	n.a.[c]	76[d]
Chapter 7 (Article 2)	68	96	90	30	n.a.[c]	22[d]
Chapter 4 (post-test)[a]	71	99	45	70	37	25[d]

Note. N.a. stands for not applicable.

[a] In the table, we present the percentages of students (the average for Articles 1 and 2) who scored one or more points on a certain move (see Figures 4.2 and 4.3). For the supports, the averages of Tables 4.2 (for Article 1) and 4.3 (for Article 2) were calculated. Refutations were not analyzed.

[b] Percentage of students whose answers were in accordance with our own analysis (based on Table 5.3).

[c] Supports could not be described with one number.

[d] Only counterarguments.

[e] This percentage is 0 because there were four main conclusions: none of the students identified all four.

students. The percentages were determined at different moments in the module, with different assignments, with different methods of analysis, and with different research articles. Because of these contextual differences, the percentages for each move vary a great deal (with the exception of the objective). However, we can draw some cautious conclusions from the data in Table 8.1. For instance, it shows that students were generally able to identify the objective in a research article. They were also reasonable able to identify the motive, main conclusion, and implication. Students had most difficulty with the supports, counterarguments, and refutations.

The differences in conditions make it difficult to observe a progression in student performances during the course. Only in Chapter 4, where we were able to control most of abovementioned conditions, we could detect a progression. The results in Chapter 4 suggest that our teaching strategy has a positive influence on students' ability to identify rhetorical moves. However, identification of supports and counterarguments is still difficult for students at the end of the module.

Secondary Research Question 2: Which features of rhetorical moves do students use for their identification?

In Chapter 2, the task-based/think-aloud interviews with a small number of students showed that they use lexical features (such as reporting verbs), organizational features (such as the place of the sentence in a paragraph), and content-based features (it answers the research question) to identify conclusions.

In Chapter 5, we have used these categories to study students' identification of rhetorical moves in more detail. We observed that students were quite successful in identifying the motive, objective, and main conclusion (moves with pronounced lexical, organizational, and content-based features). In contrast, the moves with less pronounced lexical and organizational features (implications, supports, counterarguments, and refutations) were more difficult to identify for students. So, it seems that students rely mainly on lexical and organizational features and less on content-based features when identifying moves.

Based on these results, we hypothesized that students mainly use lexical (and sometimes organizational) features to identify the move and then use content-based features to check if their assessment is correct. This could explain why students ignored, in some cases, the content-based features of moves that had no pronounced lexical or organizational features (such as counterarguments).

Secondary Research Question 3: How do students' reading behaviors change during the implementation of our teaching strategy?

We looked at two aspects of reading behavior: (non-)sequential and (non-)selective reading. The task-based/think-aloud interviews with four students in our preliminary investigation described in Chapter 2 showed that when students first read a research article, most of them read the articles sequentially and selectively (not reading all parts of an article) and that they began to read more non-sequentially when they subsequently worked on the assignment. In Chapter 4, 15% of the students (n = 110) reported that they read the pre-test article in a non-sequential way. At the end of the module (for the post-test article), this percentage had increased to 32%.

As stated above, the task-based/think-aloud interviews described in Chapter 2 showed that students read research articles selectively. For example, students sometimes skipped or skimmed the Method section, as they thought that this section did not contain information relevant for the assignment. The results in Chapter 4 indicated that at the end of the module, students (according to self-reports) paid less attention to the Abstract, Method, and Results sections compared to the beginning (Table 4.4).

A possible explanation for this shift is that the moves they had to identify were mainly located in the other sections of the research article. As stated at the end of Chapter 4, we surmise that students' reading has become more goal-directed and thus more expert-like. This does not mean that the goals with which students read have become more expert-like. Students reading is primarily assignment-driven, while experts read to further their research.

Secondary Research Question 4: What is students' ability to construct an argumentation scheme of a research article?

The SAM model, which development was described in Chapter 3, consists of descriptions of rhetorical moves and a SAM scheme depicting the relations between these moves. In Chapter 6, students had to construct a whole SAM scheme. For the analysis of these schemes, we focused on the quality of the support chains, as these were not analyzed in detail in the other chapters.

Our results indicated that students who followed our module were generally able to construct a SAM scheme. Students identified the objective and main conclusion of the article and placed them in the correct locations of the scheme, thereby linking these two moves. To a lesser extent they identified the motive, implication, counterarguments, and refutations. Because students often did not identify the same motive and counterarguments/refutations as we did, the relations motive-objective and counterargument/refutation-main conclusion in their schemes did not correspond with our own analysis. Although we analyzed the extent to which students connected moves in the schemes, our methodology did not allow us to determine to what extent students are conscious of the relations between moves they identified.

The students constructed support chains and placed these under the central arrow that connects the objective to the main conclusion. Our analysis showed that students identified the most important inscriptions and referred to them in the scheme, but they did not present all the article's supports. Further, we found that students connected supports in a way that did not correspond with our own analysis. For example, linking supports from the Discussion section with supports from other parts of the article (i.e. the coordination of evidence across epistemic levels) was problematic for students.

Secondary Research Question 5: What is students' ability to critically evaluate research articles?

In Chapter 2, we found that students relied in great deal on the authors' own interpretation of the results. This indicates that students are not inclined to look critically at the way authors underpin their conclusions.

In Chapter 7, we described a baseline measurement study about students' ability to critically evaluate research articles. Students were assigned to read two research articles and identify their rhetorical moves. We also asked them to come up with additional counterarguments (i.e. counterarguments not mentioned by the authors), because this is – in our opinion – one aspect of critical evaluation. Additionally, students had to resolve the contradictory evidence in the two articles.

The results showed that students formulated new counterarguments, although these were often rather non-specific and related to observational errors (e.g. sampling errors or insufficient controls). Students also had difficulty with resolving contradictory evidence. A small majority of the students thought that the data in the two articles showed that vitamin C reduces levels of oxidative stress, even though the study described in the first article did not directly measure oxidative stress and the study described in the second article could not find an effect. We drew the conclusion that students' ability to evaluate research articles is probably influenced by their epistemological beliefs.

As stated previously, our primary research question was:

How can undergraduate life science students be taught to read research articles?

Comparing pre-test and post-test scores, we found that students' identification of the motive, objective, main conclusion, and implications improved. Their ability to identify supports and counterarguments did not improve significantly. These observations from Chapter 4 were supported by the results in other chapters. Furthermore, based on self-reported data, we found that students changed their reading behaviors.

We conclude that the teaching strategy, which was based on two design principles (the focus on rhetorical structure and cognitive apprenticeship), is useful for teaching students how to read research articles.

Our central design principle was the focus on the rhetorical structure of research articles. The combination of rhetorical moves and argumentation theory, operationalized in SAM, may be a powerful tool for giving students insight in the rhetorical structure of research articles.

Despite the benefits of SAM for improving students' rhetorical consciousness, we found that students' reading abilities are restricted by their epistemological beliefs (especially about the use of counterarguments and refutations as rhetorical strategies), and their prior knowledge. It is difficult for students to judge the validity and significance of the conclusions in a research article, particularly because this requires knowledge of the research area that is not presented in the text. In future, this should be studied in more detail (see Section 8.5).

We should keep in mind that gaining insight into the rhetorical structure is only a first step into students' development to capable readers of research articles. Ultimately, readers should be able to (among other things) connect their prior knowledge to new information in the text, monitor their comprehension, and draw inferences during and after reading (Pearson, Roehler, Dole, & Duffy, 1992). Identifying rhetorical moves may help them with these processes.

Our derived design principles (authenticity and interaction) were inspired by cognitive apprenticeship (Collins, Brown, & Holum, 1991). The derived design principle of authenticity meant that we used research articles that were not translated or adapted. This was important, because an authentic context ensures that students understand the relevance of the tasks they receive.

Lee and Butler (2003) mention four methods of adding authenticity to tasks in the science classroom: (1) the students work on real-world problems that are comparable to the ones dealt with by scientists, (2) students find solutions for problems from their own lives, (3) students work together with scientists, and (4) the tasks involve activities such as argumentation, presentation, or communication. In our study, we made tasks authentic by using the last two of these four methods. Regarding the first method, there is a distinct difference in the types of problems scientists work on and the types of problems our students had to solve. In this sense, the student tasks were not authentic. This corresponds with Freedman and Medway (1994), who wrote that in science writing it is not possible for students to be fully inducted into the genres of working science while still in school, because the demands that motivate the writing strategies of scientists are not the same as the demands of the writing strategies of students. However, this does not mean that school writing "cannot grant experience of the *ways of thinking* or procedures of handling concepts and styles of deployment of argument, that are employed in the professional domain" (p. 14). The same can be said of reading. So, even while students read with a different purpose than scientists, it is still valuable for them to read original research articles. By reading research articles, students will be exposed to the ways of thinking of scientists – even though the tasks are not fully authentic. For example, they may learn about the use of data to support conclusions, how to set up experiments, the use of references in a text, and so forth. (Although one has to keep in mind that research articles only give a partial account of the thinking processes of scientists, as is explained in Section 1.4.) In Section 8.6, we will give some suggestions on how to further improve the authenticity of student activities.

The derived design principle of interaction was important because it ensures appropriation. The reader learns via this process which features of a text are important, how to interpret specific semiotic constructions, how to evaluate the content, and so forth. Bhatia (1997) – following Bakhtin (1981) – describes how members of a particular discourse community – the insiders – share "the knowledge of the genre, which includes the knowledge of its construction, interpretation and use" (p. 364). This shared genre knowledge is not routinely available to outsiders. Interaction between students and tutor also stimulates the practice of talking about science. As Sutton (2003) notes: "Learners should experience language as a medium for conversation about ideas, not just for receiving 'the truth'. Students should re-work scientific ideas and practise using those ideas in argument and discussion" (p. 36). Especially a research-tutored environment is suited for this. In our research, the tutor functions as intermediary between scientists/lecturers (who are the insiders or experts) and students (who are more or less outsiders or novices). However, our findings raise the question if our tutors had enough experience to function as intermediaries. We will reflect on this in Section 8.4 and give some recommendations in Section 8.6.

8.4 Reflections on the Study

In this section we will critically evaluate and indicate the limitations of our study by commenting on the selection of participants, data sources, the heuristic (i.e. SAM), the design of the module (choice of articles and the role of prior knowledge), and the role of the tutors.

Participants

Because of practical considerations, the participants of our study all followed the same course (Biomedical Research). They were all first-year undergraduate life science students who had chosen to major in biomedical sciences or behavioral and neurosciences. Some of our analyses involved the whole group of students (with a size of approximately 100–125) who attended the course. In other cases, we used data from single tutor groups, especially for our qualitative analyses. Tutor groups were randomly composed and the selection of tutor groups for our analyses was also done randomly.

Although the sample used in this research was very specific, freshmen in other disciplines share the same characteristics regarding reading experience and their level of prior knowledge. This means that the basis of our strategy may be adequately used in freshmen courses of other bachelor programs. Of course, this requires the use of other research articles. We will discuss this issue further in Section 8.5.

Data Sources

We used several data sources in our study: written answers to assignments, questionnaires, and interviews and group discussions.

Written answers. In our research, we have focused primarily on students' written answers to questions of the articles' content. The main focus in our analysis was students' identification of rhetorical moves. For this purpose, we explicitly asked students to identify specific rhetorical moves. This allowed us to analyze the products of a large number of students. Alternatively, we could have assigned students to write summaries or reviews of the articles and extracted the identified moves from these texts as others have done (e.g. Koeneman, Goedhart, & Ossevoort, 2013). However, it is more difficult to determine to which move sentences belong if summaries are used. Moreover, the quality of summaries and reviews is not only determined by students' reading skills, but also by their writing skills.

We compared students' answers with our own analysis and in one case – in Chapter 2 – with the analysis of experts. In this case, the analysis of the experts was not radically different from our own analysis. Our own analysis of the articles was often performed by two researchers, who analyzed the articles independently and later agreed on the location(s) of each move. We used different assignments asking students to identify moves. In some cases, students were assigned to write down text fragments verbatim; in other cases, students could paraphrase text fragments. Although the paraphrased answers were somewhat more difficult to analyze, it is our impression that this did not influence the results.

In the cases where students were assigned to identify supports and counterarguments (Chapters 2, 4, 5, 6, and 7), we observed that students only mentioned a small number of sentences (in Chapter 6 they identified on average 3.2 support chains, while our analysis indicated that there were at least 13 present in the research article). This may mean that their ability to identify these moves was limited, but it is also possible that students stopped working on the task after having found a certain number of moves. To discriminate between these options we should make the assignments more directive (e.g. ask students to write down all supports they can find) or we could use task-based/think-aloud interviews.

Questionnaires. We used questionnaires to determine (among other things) students' perception of their own abilities, reading behaviors, students' appreciation of the course, and tutors' deviations from the teaching strategy. In Chapter 4, we found that students' answers to questionnaire items on their perceived abilities did not correspond with our more objective measurements of students' abilities to identify moves. However, questionnaires were used as additional data sources to support our conclusions and, therefore, we did not validate the questionnaire items.

Interviews and group discussions. Students' written answers often do not reveal the precise reasons why they identified certain text fragments as certain moves (even if asked to explain their answers). We therefore used interviews and recordings of tutor group meetings to gather more information about students' reasoning behind their answers.

In Chapter 2, we conducted task-based interviews (combined with the think-aloud method) with a small number of students to observe their reading behaviors. During the interviews we used the requirements described by van Someren, Barnard, and Sandberg (1994) and Goldin (2000) to guarantee the replicability of the study. Since the number of interviewees was only four, we realize the limited value of these interviews. More interviews could have provided us with more information on students' reading behaviors, as well as the features they use to identify rhetorical moves. In further studies in this area we recommend to use task-based/think-aloud interviews to gain more insights into students' interpretation of scientific texts.

We observed and recorded several tutor group meetings to follow the discussions between tutor and students. Our research was done in a naturalistic setting: tutors were not instructed to behave different from their usual teaching. Although the amount of information we got from the group discussions was limited, recordings of group discussions did supply additional insights into the role of the tutor and students' understanding and interpretation of the articles.

Heuristic

Students were quite able to work with our heuristic, the Scientific Argumentation Model. Nevertheless, we recognize that SAM has some limitations regarding how well it represents the content of a research article. Firstly, there is only one central line/arrow (the connection between objective and main conclusions). However, some research articles have multiple objectives and main conclusions. When this situation arises, teachers could consider instructing students to make SAM schemes with multiple central lines. However, this would make the SAM schemes rather complicated and less clear.

Secondly, we observed that research articles often contained explanations: these are statements that aim to make sense of a phenomenon (the explanandum) based on established scientific facts (Osborne & Patterson, 2011). The role of an explanation in SAM is not clear if it is not a main conclusion. We suggest that an explanation could be included as a support, since it adds credibility to the main conclusion.

Design of the Module

Choice of articles. Students' identification of rhetorical moves is certainly influenced by the research articles they read. The selection of articles may explain the fluctuation in students' ability to identify rhetorical moves (see Table 8.1). The lecturers of the course selected the articles. Their selection was based on the subject of the articles (this should match the subjects of the course) and their perception of the level of difficulty. We checked if the articles adhered to the criteria developed by Muench (2000) and this was generally the case. The selection was neither based on the clarity of authors' argument nor the presence of certain rhetorical moves. In hindsight, we are able to give further recommendations on the selection of articles for introductory courses on scientific reading. These will be given in Section 8.6.

Prior knowledge. During the module, the subjects of the lectures students attended corresponded with the subjects of the research articles. We observed that this had two major benefits. Firstly, it helped students with understanding the research articles. Secondly, it encouraged the students to work through the articles, as they had to pass a written exam about the content at the end of the course.

Based on recordings of the tutor group meetings, questionnaires, and the analysis of students' written answers, we conclude that students were generally able to understand the concepts mentioned in the articles. Therefore, students' identification of the moves was probably not considerably influenced by a lack of understanding of the experimental narrative. However, students do not have the same level of prior knowledge as experts. The identification of certain moves might require a deeper understanding of the article at the levels of interpretation and criticism (Gillen, 2006). For example, the identification of supports may be easier if the reader knows the intricacies of how certain data warrants the main conclusion.

The Role of the Tutors

In educational design research it is important that the design is implemented as intended by the researcher. We recorded only a small number of tutor group meetings, so relied mainly on the questionnaires the tutors filled out (see Chapter 4). These suggest that tutors generally followed the intended scenario but sometimes implemented the teaching strategy in a different way. This concerned particularly the modeling phase of the cognitive apprenticeship approach. It was intended that tutors – in the role of experts – showed students how they identified moves. However, tutors sometimes chose a different strategy in which they asked students to identify the moves after reading the information sheets.

Further, it appeared from the transcripts of tutor group meetings that it was sometimes difficult for the tutor to instigate discussions in which students could, for instance, exchange their criteria for identifying the moves. This shows that tutors need more knowledge on using educational approaches where students are stimulated to interact.

The deviations affect our conclusions on the utilization of the cognitive apprenticeship approach. The deviations between the intended curriculum and the realized curriculum (Goodlad, 1979) may be diminished by proper tutor instructions before the start of the course. Nevertheless, we should not forget that tutors do not have much teaching experience and have problems with teaching according to a predetermined strategy.

8.5 Recommendations for Further Research

In this thesis we, like Hill, Soppelsa, and West (1982), Samuels et al. (1988), Blanton (1990), Swales (1990), and du Boulay (1999), have assumed that there is a clear relation between reading ability and knowledge about the structural characteristics of a text. As a consequence, we expect that more knowledge about text structure will have a positive effect on text comprehension. O'Reilly and McNamara (2007) found that increased reading skill helps learners to partially compensate for lower knowledge when comprehending a text. In future, it would be interesting to determine how text comprehension, prior knowledge, and rhetorical consciousness influence each other. For example, in Chapter 4 students reported an increase in text comprehension (they found the article more understandable) and their rhetorical consciousness improved. Is there a causal relationship between these two variables? If so, it would further validate our teaching strategy.

Another suggestion for further research is to investigate how we can improve students' identification of supports, counterarguments, and refutations. Our research in Chapters 5, 6, and 7 showed that students have difficulty with identifying these moves. In the following section we will formulate recommendations that can help students with the identification of these moves. We could also study how the design of our heuristic influences the identification of moves. With this information we could adapt the heuristic to improve students' identification of certain moves.

Another option is to investigate the use of our heuristic in other scientific disciplines besides the life sciences. We think that SAM is useful in disciplines that rely in great deal on experimental work and expect that articles from these disciplines contain more or less the same moves, as an experiment has to be introduced with a motive and an objective, a conclusion has to be drawn based on the experimental data, and so forth. As Knorr-Cetina (1981) stated, research articles "tend to be rhetorically standardized with regard to paragraph organization, choice of vocabulary and grammatical means of expression" (p. 95). However, the frequency of rhetorical moves may differ between disciplines. Peacock (2002) found that research articles in physics and environmental science contain fewer limitations. This could indicate that it is more difficult to find counterarguments in research articles from these disciplines. So, it is necessary to analyze research articles from other disciplines and determine if they are compatible with SAM.

The results of Chapter 6 show that students found the construction of a SAM scheme very time consuming. Only a small majority stated that they planned to construct a SAM scheme more often. It should be investigated if it is beneficial for students to construct a SAM scheme when reading a research article on a regular basis. It could be the case that constructing a SAM scheme has no added value once students are conscious of the rhetorical structure of research article. Otherwise, we should investigate how we can stimulate students to keep using the SAM scheme when reading research articles.

Our research shows the importance of improving students' personal epistemologies (or epistemological/epistemic beliefs). These are beliefs about the nature of knowledge and the process of knowing (Hofer & Pintrich, 1997). Hofer (2004) distinguishes four approaches for describing personal epistemologies. The first of these four is epistemological development (e.g. King & Kitchener's Reflective Judgment Model, 1994). This approach assumes that a person's epistemological beliefs go through certain stages of development. The second approach describes personal epistemologies as a collection of beliefs about knowledge (e.g. Schommer, 1990). In contrast to the first approach, these beliefs are independent; they do not develop along certain lines.[1] The third approach – proposed by Hofer and Pintrich (1997) – is centered on epistemological theories. This approach assumes that epistemological beliefs are organized into theory-like structures. The fourth approach – developed by Hammer and Elby (2002) – uses so-called epistemological resources. These resources may be described as context specific epistemological beliefs.

According to Hofer (2004), the different approaches all share a set of four underlying dimensions: certainty of knowledge (is knowledge fixed or more fluid?), simplicity of knowledge (is knowledge viewed as a set of discrete facts or contextual and contingent concepts?), source of knowledge (does someone view knowledge as external or as something which meaning you actively construct?), and justification for knowing (what makes a sufficient knowledge claim? – e.g. the epistemic criteria mentioned in Chapter 7). The first two dimensions relate to the nature of knowledge while the latter two relate to the nature of knowing. A notable difference between these abovementioned models is their domain specificity. For instance, King and Kitchener's (1994) and Schommer's (1990) models are domain general. In contrast, Hofer and Pintrich's (1997) epistemological theories may be domain general or domain specific.

An example of a domain specific model that is focused on scientific epistemology is Carey and Smith's (1993) widely cited three-level description of students' understanding of the nature of science (NOS). In this model, students with a level 1 understanding believe that scientific models more or less resemble reality. Students with a more sophisticated level 3 understanding view models as instruments that can be used to develop and test scientific theories. However, some argue that there is no evidence that scientific epistemologies are structured in such a coherent way (e.g. Elby, 2010).

In Chapters 5 and 7 we noted that students' reading is influenced by their epistemological beliefs. For example, some students did not appre-

[1] As Hofer (2004) points out, the first two approaches "are potentially compatible views, as beliefs, organized and structured as theories, might be expected to develop over time in somewhat predictably patterned ways" (p.45).

ciate the rhetorical strategies used by authors in research articles. These rhetorical strategies are probably more valued when students are more conscious of the persuasive nature of research articles. This suggests to us that epistemological beliefs play an important role when reading research articles. As suggested in Section 7.5, it would be worthwhile to investigate how certain features of research articles (like the inclusion of counterarguments) influence students' views on the correctness of the main conclusion.

Nussbaum, Sinatra, and Poliquin (2008) found that there is a relation between students' epistemological beliefs and their ability to construct scientific arguments. Additionally, Khishfe (2012) found a relation between students' epistemological beliefs and their ability to generate arguments, counterarguments, and rebuttals. This could indicate that there is a relation between epistemological beliefs and rhetorical consciousness. Further research could elucidate this.

8.6 Recommendations for Teaching

General Recommendations

We can make a number of recommendations for implementing our teaching strategy in courses for undergraduate science students. We have shown that our teaching strategy was effective and we can recommend implementing the strategy in other courses. The success of the strategy depends on the following conditions:

- The concept of rhetorical moves is rather abstract for novice readers. Therefore, it is advisable to use information sheets with clear descriptions and authentic examples of rhetorical moves. Our results indicate that students appreciated these information sheets.
- Subjects of the lectures and lab work should correspond with the subjects of the research articles.
- Students should work in small tutor groups. Tutors act as experts who show how they identify rhetorical moves and tutors should stimulate discussion between students on their identification of moves and their use of the SAM scheme.
- Proper scheduling of the activities in the course is essential. Lab work, lectures, and tutor group meetings should be adjusted to each other. Further, students should be given enough time to prepare for the tutor meetings, to read the research articles and make the assignments.

As mentioned in Section 8.3, students' reading activities were authentic with regard to a number of aspects. For instance, we used original research articles that were not translated or adapted. However, authenticity may be improved if students' reading is combined with research activities, for instance in inquiry-oriented research projects. Then, reading will prepare them for their lab work. Because information from research articles is used to advance students' research, reading will not only be assignment-driven. This results in more goal-oriented reading, thus affecting their performances in a positive way.

As our research shows, students have difficulty with the identification of supports, counterarguments, and refutations. This has probably to do with the fact that these moves are mostly identifiable by content-based features, and not lexical or organizational features. In Chapters 5 and 6, we offered suggestions that could help students with the identification of these moves:

1. Students' epistemological beliefs should be adjusted, so that they will have a better understanding of the role of rhetorical strategies in research articles (e.g. the inclusion of hedges and counterarguments). Understanding these strategies will make students more appreciative about authors' use of counterarguments and refutations (Chapter 5). This argument may also be reversed: identification of these elements may adjust students' epistemological beliefs.
2. Students should be provided with opportunities to exchange their answers and their views on the features of rhetorical moves (e.g. by letting them discuss their answers to the assignment in pairs during the tutor group meetings). These kinds of exchanges between students were now relatively rare (Chapter 5).
3. The instructions we gave to the tutors should be more explicit. In our research, they were too limited. As a result the tutors did not use the opportunities to discuss more extensively the identification of moves. A better preparation could improve their ability to discuss the features of rhetorical moves. This would help to bring suggestion 2 into practice and it would also ensure that the tutors stay close to the move descriptions given in the information sheets (Chapter 5).
4. We should extend the example SAM scheme by including all different categories of supports (Chapter 6).
5. Students' awareness of differences in the epistemic levels of supports should be further developed, for instance by letting students make SAM schemes during the tutor group meetings (Chapter 6).

Regarding the second and third suggestions, research has shown that (graduate) student teaching assistants have difficulties with their role as a teacher: most of them lack pedagogical skills and sometimes they have only a superficial knowledge of the content (Luft, Kurdziel, Roehrig, &

Turner, 2004). This reinforces our argument for providing tutors with more instructions about effective teaching strategies.

Improving Epistemological Beliefs

How can we further develop students' epistemological beliefs? Jehng, Johnson, and Anderson (1993) suggest that the education level plays an important role. At the undergraduate level, introductory courses contain content that is systematically organized. Also, problems that students have to solve are well-structured. This may cause students to believe that knowledge is certain. At the graduate level, the content of courses is less structured. By reading research articles, students learn that theories regularly conflict with each other and that the questions seldom have one absolute answer. This leads to a more sophisticated view on the nature of knowledge.

According to a review article by Abd-El-Khalick and Lederman (2000) on teaching the nature of science to prospective science teachers, enhancing students' epistemological views can be done implicitly or explicitly. An implicit approach assumes that understanding the nature of science "is a learning outcome that can be facilitated through process skill instruction, science content coursework, and 'doing science'" (p. 673). An explicit approach utilizes "elements from history and philosophy of science and/or instruction geared towards the various aspects of NOS" (p. 673). After a meta-analysis, Abd-El-Khalick and Lederman conclude that the explicit approach is the most effective. Therefore, it could be worthwhile to combine our teaching strategy with implicit and explicit approaches that aim to enhance students' epistemological beliefs.

A possible approach could involve the use of a "collective classic" (Goodney & Long, 2003). A collective classic is a set of related articles that represents a scientific revolution (e.g. the discovery of the structure of DNA). Because of their revolutionary character, these articles are somewhat atypical. However, discussing a collective classic can give students insight into how major discoveries are communicated and the role of language in this process.

Recommendations Regarding the Choice of Articles

We can also make recommendations with regard to the choice of research articles that can be used in our teaching strategy. Muench (2000) described three criteria for selecting suitable primary literature for novice readers: (1) the experiments in the article should be visualized easily, (2) the results should be unambiguous, and (3) the relationship between the conclusion and the data should be relatively simple. In the context of our teaching strategy, we would like to add three additional criteria for selecting articles for novice readers:

- To teach students the role of supports, choose articles with a limited amount of supports. If there are too many, students will lose overview. This will inhibit them to see the connections between the supports and the main conclusion. When students gain more experience, they can be given articles with more supports.
- Choose articles with a clear objective and a clear main conclusion. The connection between these two moves forms the central line/arrow of argumentation in an article to which all the other elements are connected. If the beginning and end point of this central line of argumentation are unclear, it will also be difficult to place the other rhetorical moves. This criterion is somewhat akin to Muench's third criterion.
- Choose articles that contain the moves students need to learn to identify at that specific time. For example, some articles do not contain counterarguments or an implication. If articles do not contain these moves, students cannot gain experience in identifying them. Later on, when students have more knowledge of the different moves, it is not an issue that an article does not contain all the moves.

8.7 A Reading and Writing Science Curriculum

In this thesis, we have studied students' reading of research articles. However, scientific literacy also encompasses writing skills (Norris & Phillips, 2003). It has been suggested that many scientists find writing research articles a difficult chore. As Montgomery (1996) notes, there is a dichotomy between the act of doing science – which is often exciting and interesting – and the dry, unemotional prose of research articles: "How can [technical writing] not be a chore for so many, a task for even the most competent?" (p. 23). This suggests to us that it is important to give sufficient attention to scientific writing in a curriculum.

Prain (2004) distinguishes two approaches regarding the role of writing in science education. The first approach may be characterized as learning how to write science. This approach "assert[s] that students primarily must learn to understand and reproduce the traditional written discourses of the science community if they are to become scientifically literate" (p. 34). These discourses are regular scientific reports, such as lab reports, conference papers, or journal articles. The second approach asserts that students should learn to use a diversified range of science writing types (e.g. writing a scientific report in such a way that your family is able to understand it). These writing types may be formal or informal:

> The diversified writing approach assumes that written language can be a valuable resource for various aspects of learning science and science literacy, and proposes a broader role for writing than mainly as a record of past learning or as an imitation of the 'official' writing formats of practicing scientists. (p. 35)

In the context of higher education, we think that both approaches are important. The first approach is valuable because the production of well-written texts – not only research articles, but also research proposals, management reports, and popular-scientific texts – is nowadays very important for scientists. As Duff (2010) notes: "Assessments for scholarships, grants, degrees, and jobs require more strategic and visible output with greater perceived impact than ever before" (p. 186). The second approach is valuable because it enhances "students' engagement in, and understanding of, science concepts and practices" (Prain, 2004, p. 42).

We want to present some ideas about a scientific literacy curriculum in which students learn to read and write scientific texts. In this curriculum reading, writing, argumentation, and epistemological beliefs should form an integrated whole. This curriculum should follow the abovementioned two approaches. So, besides learning to read and write research articles, they should also read and write other types of writing.

Martin (1999) describes a set of related teaching models that aim to give students control of a genre. A prototypical example is the 1989 DSP (Disadvantaged School Program) Primary Curriculum Model (Callaghan & Knapp, 1989, as cited in Martin, 1999). This model contains three phases: modeling, joint construction of text, and independent construction of text. During the modeling phase, students are introduced to an example of a text genre, they discuss the function of the genre, and examine its structure. During the joint construction phase, students develop a text together with the teacher. For example, a teacher may develop a text on large sheets of paper in response to students' suggestions. During the last phase, the independent construction of text, students will produce a text on their own. Creative exploration of the genre and its possibilities is encouraged during this phase.

We would like to suggest a similar approach for teaching undergraduate students how to read and write research articles and other types of writing. The modeling phase corresponds mainly with the teaching strategy described in this thesis and is centered on the rhetorical structure of a research article. This phase takes place – in our case – in the first year of the bachelor program. The next phase, the joint construction of text, could take place in the second year of the bachelor program. Together with more experienced writers (e.g. tutors), students could develop a relatively simple research article, possibly based on an existing data set or data collected by students themselves. During the last phase, students write scientific texts by themselves.

These activities should be as authentic as possible and should be integrated with students' own research activities. So, students should write texts for real-life purposes. Research has shown that this is a very effective way of learning to write in a certain genre (Purcell-Gates, Duke, & Martineau, 2007). For example, students could use their own data when writing a research article and publish their texts in a *faux* journal. As a guideline, students should focus on writing a text containing the seven moves that are described by SAM.

We think that SAM may not only be useful in the modeling phase, but also in the other two phases. Using SAM could help students to improve their own writing.[2] They could first develop an argument by constructing a scheme, and then write the text.

Additionally, students should be introduced to other aspects of scientific writing, such as incorporating hedges, the effective use of references, how to write the Method section, and so forth. In other words, we want to extend students' genre knowledge. Genre knowledge does not provide students with a clear-cut and simplistic template for writing research articles. Instead, it allows them to explore the possibilities of a genre. As Cooper (1998) wrote:

> The genre knowledge that comes from discussing [model texts] (…) provides a powerful heuristic for writing, opening up many possibilities for students and leaving countless decisions for them to make as they develop and shape their arguments. (p. 48)

Following this curriculum, we should strive to integrate reading and writing as much as possible, as these two activities reinforce each other. When students know where to find a main conclusion when reading a research article, they will also better be able to place the main conclusion when writing a research article themselves (and vice versa).

Reading, interpreting, and discussing research articles are research-based and research-tutored teaching activities (Jenkins & Healey, 2010), as students undertake inquiry and engage in research discussions. The use of such research-based and research-tutored activities may strengthen the research-teaching nexus, in which research and teaching are connected to

[2] Alternately, CAS (Comprehensive Argumentation Scheme) instead of SAM could be used. In contrast to SAM, CAS depicts the rhetorical moves that are mentioned in the Method section.

each other. As stated in Chapter 1, connecting teaching and research has become an important goal of many universities (Elsen, Visser-Wijnveen, van der Rijst, & van Driel, 2009). It is thought that strengthening the research-teaching nexus in higher education will prepare students more effectively to the demands of the modern knowledge society, as the students will know how to manage knowledge, how knowledge is generated, and so forth. We hope that our research will contribute to this process.

References

Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of nature of science: A critical review of the literature. *International Journal of Science Education*, *22*(7), 665–701.

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher*, *61*(5), 364–373.

Alexandrov, A. V. (2004). How to write a research paper. *Cerebrovascular Diseases*, *18*(2), 135–138.

Allchin, D. (2001). Error types. *Perspectives on Science*, *9*(1), 38–58.

Almeida, C. A., & Liotta, L. J. (2005). Organic chemistry of the cell: An interdisciplinary approach to learning with a focus on reading, analyzing, and critiquing primary literature. *Journal of Chemical Education*, *82*(12), 1794–1799.

Amsterdamska, O., & Leydesdorff, L. (1989). Citations: Indicators of significance? *Scientometrics*, *15*(5), 449–471.

Andrews, R. (2010). *Argumentation in higher education: Improving practice through theory and research*. New York: Routledge.

Andrews, R., Bilbro, R., Mitchell, S., Peake, K., Prion, P., Robinson, A., … Torgerson, C. (2006). *Argumentative skills in first year undergraduates: A pilot study*. York, U.K.: The Higher Education Academy.

Bailin, S. (2002). Critical thinking and science education. *Science & Education*, *11*(4), 361–375.

Bakhtin, M. M. (1981). *The dialogic imagination: Four essays*. (C. Emerson & M. Holquist, Trans., M. Holquist, Ed.). Austin: University of Texas Press.

Baram-Tsabari, A., & Yarden, A. (2005). Text genre as a factor in the formation of scientific literacy. *Journal of Research in Science Teaching*, *42*(4), 403–428.

Barthes, R. (1977). *Image music text*. (S. Heath, Trans.). London: Fontana Press.

Bas, O., Songur, A., Sahin, O., Mollaoglu, H., Ozen, O. A., Yaman, M., … Yagmurca, M. (2007). The protective effect of fish n-3 fatty acids on cerebral ischemia in rat hippocampus. *Neurochemistry International*, *50*(3), 548–554.

Bauersachs, J., Bouloumié, A., Fraccarollo, D., Hu, K., Busse, R., & Ertl, G. (1999). Endothelial dysfunction in chronic myocardial infarction despite increased vascular endothelial nitric oxide synthase and soluble guanylate cyclase expression: Role of enhanced vascular superoxide production. *Circulation*, *100*(3), 292–298.

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

Benelux Bologna Secretariat. (2009). Bologna beyond 2010: Report on the development of the European Higher Education Area. Retrieved from http://www.ond.vlaanderen.be/hogeronderwijs/bologna/conference/documents/Beyond_2010_report_FINAl.pdf

Berkenkotter, C., & Huckin, T. (1995). *Genre knowledge in disciplinary communication: Cognition/culture/power*. Hillsdale, NJ: Lawrence Erlbaum.

Bhatia, V. K. (1997). The power and politics of genre. *World Englishes*, *16*(3), 359–371.

Bhatia, V. K. (2004). *Worlds of written discourse: A genre-based view*. London: Continuum.

Biomedisch onderzoek. (n.d.). Retrieved from http://www.rug.nl/ocasys/rug/vak/show?code=WLP10B20

Björk, B. C., Roos, A., & Lauri, M. (2009). Scientific journal publishing–yearly volume and open access availability. *Information Research*, *14*(1), paper 391.

Blanton, W. E. (1990). The role of purpose in reading instruction. *The Reading Teacher*, *43*(7), 486–493.

Bloch, J. (2010). A concordance-based study of the use of reporting verbs as rhetorical devices in academic papers. *Journal of Writing Research*, *2*(2), 219–244.

Booth, W. C., Colomb, G. G., & Williams, J. M. (2003). *The craft of research* (2nd ed.). Chicago: University of Chicago Press.

Brett, P. (1994). A genre analysis of the results section of sociology articles. *English for Specific Purposes*, *13*(1), 47–59.

Brill, G., Falk, H., & Yarden, A. (2004). The learning processes of two high-school biology students when reading primary literature. *International Journal of Science Education*, *26*(4), 497–512.

Brown, J. R. (2001). *Who rules in science? An opinionated guide to the wars*. Cambridge: Cambridge University Press.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32–42.

Budovec, J. J., & Kahn, C. E., Jr. (2010). Evidence-based radiology: A primer in reading scientific articles. *American Journal of Roentgenology*, *195*(1), W1–W4.

Burrough-Boenisch, J. (1999). International reading strategies for IMRD articles. *Written Communication*, *16*(3), 296–316.

Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist*, *28*(3), 235–251.

Cela-Conde, C. J., Ayala, F. J., Munar, E., Maestú, F., Nadal, M., Capó, M. A., … Marty, G. (2009). Sex-related similarities and differences in the neural correlates of beauty. *Proceedings of the National Academy of Sciences*, *106*(10), 3847–3852.

Charney, D. (1993). A study in rhetorical reading: How evolutionists read “The Spandrels of San Marco”. In J. Selzer (Ed.), *Understanding scientific prose* (pp. 203–231). Madison: University of Wisconsin Press.

Chen, P. C., Vargas, M. R., Pani, A. K., Smeyne, R. J., Johnson, D. A., Kan, Y. W., & Johnson, J. A. (2009). Nrf2-mediated neuroprotection in the MPTP mouse model of Parkinson's disease: Critical role for the astrocyte. *Proceedings of the National Academy of Sciences*, *106*(8), 2933–2938.

Chen, Q., & Donin, J. (1997). Discourse processing of first and second language biology texts: Effects of language proficiency and domain-specific knowledge. *Modern Language Journal*, *81*(2), 209–227.

Clanchy, J., & Ballard, B. (1995). Generic skills in the context of higher education. *Higher Education Research and Development*, *14*(2), 155–166.

Clark, B. R. (1997). The modern integration of research activities with teaching and learning. *The Journal of Higher Education*, *68*(3), 241–255.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9.

Coil, D., Wenderoth, M. P., Cunningham, M., & Dirks, C. (2010). Teaching the process of science: faculty perceptions and an effective methodology. *CBE–Life Sciences Education*, *9*(4), 524–535.

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator*, *15*(3), 6–11, 38–46.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences*, *13*(1), 15–42.

Connor, U., Upton, T. A., & Kanoksilapatham, B. (2007). Introduction to move analysis. In D. Biber, U. Connor, & T. A. Upton (Eds.), *Discourse on the move: Using corpus analysis to describe discourse structure* (pp. 23–41). Amsterdam: John Benjamins.

Cooper, C. R. (1998). What we know about genres, and how it can help us assign and evaluate writing. In C. R. Cooper & L. Odell (Eds.), *Evaluating writing: The role of teachers' knowledge about text, learning, and culture* (pp. 23–52). Urbana, IL: National Council of Teachers of English.

Crabbe, J. C., Wahlsten, D., & Dudek, B. C. (1999). Genetics of mouse behavior: Interactions with laboratory environment. *Science*, *284*(5420), 1670–1672.

Crookes, G. (1986). Towards a validated analysis of scientific text structure. *Applied Linguistics*, *7*(1), 57–70.

da Silva Soares, P. P., Porto, C. S., Francis Abdalla, F. M., De La Fuente, R. N., Moreira, E. D., Krieger, E. M., & Irigoyen, M. C. (2006). Effects of rat sinoaortic denervation on the vagal responsiveness and expression of muscarinic acetylcholine receptors. *Journal of Cardiovascular Pharmacology*, *47*(3), 331–336.

Darko, D., Dornhorst, A., Kelly, F. J., Ritter, J. M., & Chowienczyk, P. J. (2002). Lack of effect of oral vitamin C on blood pressure, oxidative stress and endothelial function in type II diabetes. *Clinical Science*, *103*(4), 339–344.

Darwin, C. (1887). *The life and letters of Charles Darwin, including an autobiographical chapter*. (F. Darwin, Ed.) (Vols. 1-3, Vol. 1). London: John Murray.

Davson-Galle, P. (2004). Philosophy of science, critical thinking and science education. *Science & Education*, *13*(6), 503–517.

Day, R. A. (1998). *How to write and publish a scientific paper* (5th ed.). Phoenix, AZ: Oryx.

Dee-Lucas, D., & Larkin, J. (1988). Novice rules for assessing importance in scientific texts. *Journal of Memory and Language*, *27*(3), 288–308.

du Boulay, D. (1999). Argument in reading: What does it involve and how can students become better critical readers? *Teaching in Higher Education*, *4*(2), 147–162.

Dudley-Evans, T. (1994). Genre analysis: An approach to text analysis for ESP. In M. Coulthard (Ed.), *Advances in written text analysis* (pp. 219–228). London: Routledge.

Duff, P. A. (2010). Language socialization into academic discourse communities. *Annual Review of Applied Linguistics*, *30*(1), 169–192.

Duggan, S., & Gott, R. (1995). The place of investigations in practical work in the UK National Curriculum for Science. *International Journal of Science Education*, *17*(2), 137–147.

Duncan, D. B., Lubman, A., & Hoskins, S. G. (2011). Introductory biology textbooks under-represent scientific process. *Journal of Microbiology & Biology Education*, *12*(2), 143–151.

Duschl, R. A. (2008). Quality argumentation and epistemic criteria. In S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in science education* (pp. 159–175). London: Springer.

Elby, A. (2010). Coherence vs. fragmentation in student epistemologies: A reply to Smith & Wenk. *Electronic Journal of Science Education*, *14*(1). Retrieved from http://ejse.southwestern.edu/article/view/7324/5618

Elsen, M. G. M. F., Visser-Wijnveen, G. J., van der Rijst, R. M., & van Driel, J. H. (2009). How to strengthen the connection between research and teaching in undergraduate university education. *Higher Education Quarterly*, *63*(1), 64–85.

Engeström, Y. (2001). Expansive learning at work: Toward an activity-theoretical reconceptualization. *Journal of Education and Work*, *14*(1), 133–156.

Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educational Leadership*, *43*(2), 44–48.

European Ministers of Higher Education. (1999). The Bologna declaration. Retrieved from http://ec.europa.eu/education/policies/educ/bologna/bologna.pdf

European Ministers of Higher Education. (2012). Making the most of our potential: Consolidating the European Higher Education Area. Retrieved from http://www.minedu.fi/OPM/Koulutus/artikkelit/bologna/liitteet/Bucharest_Communique.pdf

Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations. Retrieved from http://www.eric.ed.gov/PDFS/ED315423.pdf

Fahnestock, J. (1986). Accommodating science: The rhetorical life of scientific facts. *Written Communication*, *3*(3), 275–296.

Fang, Z. (2005). Scientific literacy: A systemic functional linguistics perspective. *Science Education*, *89*(2), 335–347.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*(3), 221–233.

Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: Countering criticisms against corpus-based methodologies. *English for Specific Purposes*, *24*(3), 321–332.

Freedman, A., & Medway, P. (1994). Introduction: New views of genre and their implications for education. In A. Freedman & P. Medway (Eds.), *Learning and Teaching Genre* (pp. 1–22). Portsmouth, NH: Boynton/Cook.

Gilbert, G. N. (1977). Referencing as persuasion. *Social Studies of Science*, *7*(1), 113–122.

Gilbert, G. N., & Mulkay, M. J. (1984). *Opening Pandora's box: A sociological analysis of scientists' discourse*. Cambridge: Cambridge University Press.

Gillen, C. M. (2006). Criticism and interpretation: Teaching the persuasive aspects of research articles. *CBE–Life Sciences Education*, *5*(1), 34–38.

Gillen, C. M. (2007). *Reading primary literature: A practical guide to evaluating research articles in biology*. San Francisco: Pearson/Benjamin Cummings.

Gillen, C. M., Vaughan, J., & Lye, B. R. (2004). An online tutorial for helping nonscience majors read primary research literature in biology. *Advances in Physiology Education*, *28*(3), 95–99.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine de Gruyter.

Goldin, G. A. (2000). A scientific perspective on structured, task-based interviews in mathematics education research. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 517–545). Mahwah, NJ: Lawrence Erlbaum.

Goldman, S. R., & Bisanz, G. L. (2002). Toward functional analysis of scientific genres: Implications for understanding and learning processes. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 19–50). Mahwah, NJ: Lawrence Erlbaum.

Goodlad, J. I. (1979). *What schools are for*. Bloomington, IN: Phi Delta Kappa Educational Foundation.

Goodney, D. E., & Long, C. S. (2003). The collective classic: A case for the reading of science. *Science & Education*, *12*(2), 167–184.

Gould, S. J. (1993). Fulfilling the spandrels of world and mind. In J. Selzer (Ed.), *Understanding scientific prose* (pp. 310–336). Madison: University of Wisconsin Press.

Griffiths, R. (2004). Knowledge production and the research–teaching nexus: The case of the built environment disciplines. *Studies in Higher Education*, *29*(6), 709–726.

Gross, A. G., Harmon, J. E., & Reidy, M. (2002). *Communicating science: The scientific article from the 17th century to the present*. Oxford: Oxford University Press.

Halliday, M. A. K. (1993). On the language of physical science. In M. A. K. Halliday & J. R. Martin (Eds.), *Writing science: Literacy and discursive power* (pp. 54–68). London: Falmer Press.

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169–190). Mahwah, NJ: Lawrence Erlbaum.

Hammond, N. (2007). Preface: Plutarch's fire. In A. Jenkins, M. Healey, & R. Zetter (Eds.), *Linking teaching and research in disciplines and departments* (pp. 3–4). York, U.K.: The Higher Education Academy.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89.

Healey, M. (2005). Linking research and teaching: Exploring disciplinary spaces and the role of inquiry-based learning. In R. Barnett (Ed.), *Reshaping the university: New relationships between research, scholarship and teaching* (pp. 67–78). Maidenhead, U.K.: Open University Press.

Henry, A., & Roseberry, R. L. (1998). An evaluation of a genre-based approach to the teaching of EAP/ESP writing. *TESOL Quarterly*, *32*(1), 147–156.

Herman, C. (1999). Reading the literature in the jargon-intensive field of molecular genetics. *Journal of College Science Teaching*, *28*(4), 252–253.

Hill, S. S., Soppelsa, B. F., & West, G. K. (1982). Teaching ESL students to read and write experimental-research papers. *TESOL Quarterly*, *16*(3), 333–347.

Hofer, B. K. (2004). Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist*, *39*(1), 43–55.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, *67*(1), 88–140.

Hoffmann, R. (2002). Writing (and drawing) chemistry. In J. Monroe (Ed.), *Writing and revising the disciplines* (pp. 29–53). Ithaca, NY: Cornell University Press.

Hogan, K., & Maglienti, M. (2001). Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions. *Journal of Research in Science Teaching*, *38*(6), 663–687.

Holmes, R. (1997). Genre analysis, and the social sciences: An investigation of the structure of research article discussion sections in three disciplines. *English for Specific Purposes*, *16*(4), 321–337.

Hornig, B., Arakawa, N., Kohler, C., & Drexler, H. (1998). Vitamin C improves endothelial function of conduit arteries in patients with chronic heart failure. *Circulation*, *97*(4), 363–368.

Hoskins, S.G., Lopatto, D., & Stevens, L. M. (2011). The CREATE approach to primary literature shifts undergraduates' self-assessed ability to read and analyze journal articles, attitudes about science, and epistemological beliefs. *CBE–Life Sciences Education*, *10*(4), 368–378.

Hoskins, S. G., Stevens, L. M., & Nehm, R. H. (2007). Selective use of the primary literature transforms the classroom into a virtual laboratory. *Genetics*, *176*(3), 1381–1389.

Houde, A. (2000). Student symposia on primary research articles: A window into the world of scientific research. *Journal of College Science Teaching*, *30*(3), 184–87.

Hunston, S. (1989). *Evaluation in experimental research articles* (Doctoral dissertation). University of Birmingham. Retrieved from http://etheses.bham.ac.uk/912/

Hunston, S. (1994). Evaluation and organization in a sample of written academic discourse. In M. Coulthard (Ed.), *Advances in written text analysis* (pp. 191–218). London: Routledge.

Hyland, K. (1998a). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics*, *30*(4), 437–455.

Hyland, K. (1998b). *Hedging in scientific research articles*. Amsterdam: John Benjamins.

Hyland, K. (1999). Persuasion in academic articles. *Perspectives*, *11*, 73–103.

Hyland, K. (2007). Applying a gloss: Exemplifying and reformulating in academic discourse. *Applied Linguistics*, *28*(2), 266–285.

Hyon, S. (1996). Genre in three traditions: Implications for ESL. *TESOL Quarterly*, *30*(4), 693–722.

Ioannidis, J. P. A. (2007). Limitations are not properly acknowledged in the scientific literature. *Journal of Clinical Epidemiology*, *60*(4), 324–329.

Jacques-Fricke, B. T., Hubert, A., & Miller, S. (2009). A versatile module to improve understanding of scientific literature through peer instruction. *Journal of College Science Teaching*, *39*(2), 24–32.

Janick-Buckner, D. (1997). Getting undergraduates to critically read and discuss primary literature. *Journal of College Science Teaching*, *27*(1), 29–32.

Jehng, J. J., Johnson, S. D., & Anderson, R. C. (1993). Schooling and students' epistemological beliefs about learning. *Contemporary Educational Psychology*, *18*(1), 23–35.

Jenkins, A., & Healey, M. (2010). Undergraduate research and international initiatives to link teaching and research. *Council on Undergraduate Research Quarterly*, *30*(3), 36–42.

Jiménez-Aleixandre, M. P., & Federico-Agraso, M. (2009). Justification and persuasion about cloning: Arguments in Hwang's paper and journalistic reported versions. *Research in Science Education*, *39*(3), 331–347.

Johns, A. (2002). Destabilizing and enriching novice students' genre theories. In A. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 237–246). Mahwah, NJ: Lawrence Erlbaum.

Kanoksilapatham, B. (2005). Rhetorical structure of biochemistry research articles. *English for Specific Purposes*, *24*(3), 269–292.

Kelly, G., & Takao, A. (2002). Epistemic levels in argument: An analysis of university oceanography students' use of evidence in writing. *Science Education*, *86*(3), 314–342.

Kelly, G., Regev, J., & Prothero, W. (2008). Analysis of lines of reasoning in written argumentation. In S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in science education* (pp. 137–157). London: Springer.

Khishfe, R. (2012). Relationship between nature of science understandings and argumentation skills: A role for counterargument and contextual factors. *Journal of Research in Science Teaching*, *49*(4), 489–514.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey-Bass.

Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. *Human Development*, *26*(4), 222–232.

Klasson, L., Westberg, J., Sapountzis, P., Näslund, K., Lutnaes, Y., Darby, A. C., … Andersson, S. G. E. (2009). The mosaic genome structure of the Wolbachia wRi strain infecting Drosophila simulans. *Proceedings of the National Academy of Sciences*, *106*(14), 5725–5730.

Kleiger, R. E., Miller, J. P., Bigger, J. T., Jr., & Moss, A. J. (1987). Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *The American Journal of Cardiology*, *59*(4), 256–262.

Knorr-Cetina, K. (1981). *The manufacture of knowledge: An essay on the constructivist and contextual nature of science*. Oxford: Pergamon Press.

Koeneman, M., Goedhart, M., & Ossevoort, M. (2013). Introducing pre-university students to primary scientific literature through argumentation analysis. *Research in Science Education*. Advance online publication. doi: 10.1007/s11165-012-9341-y

Koffie, R. M., Meyer-Luehmann, M., Hashimoto, T., Adams, K. W., Mielke, M. L., Garcia-Alloza, M., … Spires-Jones, T. L. (2009). Oligomeric amyloid β associates with postsynaptic densities and correlates with excitatory synapse loss near senile plaques. *Proceedings of the National Academy of Sciences*, *106*(10), 4012–4017.

Korpan, C. A., Bisanz, G. L., Bisanz, J., & Henderson, J. M. (1997). Assessing literacy in science: Evaluation of scientific news briefs. *Science Education*, *81*(5), 515–532.

Kozeracki, C. A., Carey, M. F., 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. *CBE–Life Sciences Education*, *5*(4), 340–347.

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

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, *77*(3), 319–337.

Kuldell, N. (2003). Read like a scientist to write like a scientist: Using authentic literature in the classroom. *Journal of College Science Teaching*, *33*(2), 32–35.

Larivière, V., Archambault, É., & Gingras, Y. (2008). Long-term variations in the aging of scientific literature: From exponential growth to steady-state science. *Journal of the American Society for Information Science and Technology*, *59*(2), 288–296.

Larson, M., Britt, M. A., & Larson, A. A. (2004). Disfluencies in comprehending argumentative texts. *Reading Psychology*, *25*(3), 205–224.

Latour, B. (1987). *Science in action*. Milton Keynes, U.K.: Open University Press.

Latour, B., & Woolgar, S. (1979). *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Lee, H. S., & Butler, N. (2003). Making authentic science accessible to students. *International Journal of Science Education*, *25*(8), 923–948.

Levine, E. (2001). Reading your way to scientific literacy. *Journal of College Science Teaching*, *31*(2), 122–125.
Lewontin, R. C. (1991). Facts and the factitious in natural sciences. *Critical Inquiry*, *18*(1), 140–153.
Li, L.-J., & Ge, G.-C. (2009). Genre analysis: Structural and linguistic evolution of the English-medium medical research article (1985–2004). *English for Specific Purposes*, *28*(2), 93–104.
Lipman, M. (1988). Critical thinking – what can it be? *Educational Leadership*, *46*(1), 38–43.
Liu, S. Y., & Tsai, C. C. (2008). Differences in the scientific epistemological views of undergraduate students. *International Journal of Science Education*, *30*(8), 1055–1073.
Lu, T. K., & Collins, J. J. (2009). Engineered bacteriophage targeting gene networks as adjuvants for antibiotic therapy. *Proceedings of the National Academy of Sciences*, *106*(12), 4629–4634.
Luft, J. A., Kurdziel, J. P., Roehrig, G. H., & Turner, J. (2004). Growing a garden without water: Graduate teaching assistants in introductory science laboratories at a doctoral/research university. *Journal of Research in Science Teaching*, *41*(3), 211–233.
Markel, M. (1993). Induction, social constructionism, and the form of the science paper. *Journal of Technical Writing and Communication*, *23*(1), 7–22.
Marshall, S. (1991). A genre-based approach to the teaching of report-writing. *English for Specific Purposes*, *10*(1), 3–13.
Martin, J. R. (1999). Mentoring semogenesis: “Genre-based” literacy pedagogy. In F. Christie (Ed.), *Pedagogy and the shaping of consciousness: Linguistic and social processes* (pp. 123–155). London: Continuum.
Meadows, A. (1985). The scientific paper as an archaeological artefact. *Journal of Information Science*, *11*(1), 27–30.
Medawar, P. B. (1996). *The strange case of the spotted mice and other classic essays on science*. Oxford: Oxford University Press.
Mermin, N. D. (1990). *Boojums all the way through: Communicating science in a prosaic age*. Cambridge: Cambridge University Press.
Miller, C. R. (1979). A humanistic rationale for technical writing. *College English*, *40*(6), 610–617.
Miranda, A., Costa-e-Sousa, R. H., Werneck-de-Castro, J. P., Mattos, E. C., Olivares, E. L., Ribeiro, V. P., … Campos-de-Carvalho, A. C. (2007). Time course of echocardiographic and electrocardiographic parameters in myocardial infarct in rats. *Anais da Academia Brasileira de Ciências*, *79*(4), 639–648.
Montgomery, S. L. (1996). *The scientific voice*. New York: Guilford Press.
Muench, S. B. (2000). Choosing primary literature in biology to achieve specific educational goals. *Journal of College Science Teaching*, *29*(4), 255–260.
Mulnix, A. B. (2003). Investigations of protein structure and function using the scientific literature: An assignment for an undergraduate cell physiology course. *Cell Biology Education*, *2*(4), 248–255.
Myers, G. (1990). *Writing biology: Texts in the social construction of scientific knowledge*. Madison: University of Wisconsin Press.

Myers, G. (1992). "In this paper we report…" – speech acts and scientific facts. *Journal of Pragmatics*, *17*(4), 295–313.
Neumann, R., & Lindsay, A. (1987). Excellence at risk? The future of research and research training in Australian universities. *The Australian Quarterly*, *59*(2), 199–209.
Noguchi, J. (2006). *The science review article: An opportune genre in the construction of science*. Bern: Peter Lang.
Norris, S. P., & Phillips, L. M. (1994). Interpreting pragmatic meaning when reading popular reports of science. *Journal of Research in Science Teaching*, *31*(9), 947–67.
Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, *87*(2), 224–240.
Nussbaum, E. M., Sinatra, G. M., & Poliquin, A. (2008). Role of epistemic beliefs and scientific argumentation in science learning. *International Journal of Science Education*, *30*(15), 1977–1999.
Nwogu, K. N. (1997). The medical research paper: Structure and functions. *English for Specific Purposes*, *16*(2), 119–138.
O'Reilly, T., & McNamara, D. S. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal*, *44*(1), 161–196.
Opleidingscommissie Biologie. (2005). Richtlijnen opbouw bachelorscriptie.
Osborne, J. F., & Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education*, *95*(4), 627–638.
Ossevoort, M., & Voskamp, K. (Eds.). (2008). *Syllabus communicatie*. Groningen, The Netherlands: Rijksuniversiteit Groningen.
Ozen, O. A., Cosar, M., Sahin, O., Fidan, H., Eser, O., Mollaoglu, H., … Songur, A. (2008). The protective effect of fish n-3 fatty acids on cerebral ischemia in rat prefrontal cortex. *Neurological Sciences*, *29*(3), 147–152.
Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, *19*(3), 228–242.
Paltridge, B. (1994). Genre analysis and the identification of textual boundaries. *Applied Linguistics*, *15*(3), 288–299.
Paul, R. W., & Elder, L. (2002). *Critical thinking: Tools for taking charge of your professional and personal life*. Upper Saddle River, NJ: Financial Times Prentice Hall.
Peacock, M. (2002). Communicative moves in the discussion section of research articles. *System*, *30*(4), 479–497.
Pearson, P. D., Moje, E., & Greenleaf, C. (2010). Literacy and science: Each in the service of the other. *Science*, *328*(5977), 459–463.
Pearson, P. D., Roehler, L. R., Dole, J. A., & Duffy, G. G. (1992). Developing expertise in reading comprehension. In S. J. Samuals & A. Farstrup (Eds.), *What research has to say about reading instruction* (2nd ed., pp. 145–199). Newark, DE: International Reading Association.
Peck, W. H. (2004). Teaching metastability in petrology using a guided reading

from the primary literature. *Journal of Geoscience Education*, *52*(3), 284–288.

Penrose, A. M., & Katz, S. B. (1998). *Writing in the sciences: Exploring conventions of scientific discourse*. New York: Longman.

Pinch, T. (1985). Towards an analysis of scientific observation – the externality and evidential significance of observational reports in physics. *Social Studies of Science*, *15*(1), 3–36.

Popper, K. (1970). Normal science and its dangers. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 51–58). Cambridge: Cambridge University Press.

Prain, V. (2004). The role of language in science learning and literacy. In C. S. Wallace, B. Hand, & V. Prain (Eds.), *Writing and learning in the science classroom* (pp. 33–46). Dordrecht, The Netherlands: Kluwer Academic.

Prunier, F., Pfister, O., Hadri, L., Liang, L., Monte, F. D., Liao, R., & Hajjar, R. J. (2007). Delayed erythropoietin therapy reduces post-MI cardiac remodeling only at a dose that mobilizes endothelial progenitor cells. *American Journal of Physiology – Heart and Circulatory Physiology*, *292*(1), 522–529.

Purcell-Gates, V., Duke, N. K., & Martineau, J. A. (2007). Learning to read and write genre-specific text: Roles of authentic experience and explicit teaching. *Reading Research Quarterly*, *42*(1), 8–45.

Rajagopalan, S., Kurz, S., Münzel, T., Tarpey, M., Freeman, B. A., Griendling, K. K., & Harrison, D. G. (1996). Angiotensin II-mediated hypertension in the rat increases vascular superoxide production via membrane NADH/NADPH oxidase activation. Contribution to alterations of vasomotor tone. *Journal of Clinical Investigation*, *97*(8), 1916.

Ren, X. R., Reiter, E., Ahn, S., Kim, J., Chen, W., & Lefkowitz, R. J. (2005). Different G protein-coupled receptor kinases govern G protein and β-arrestin-mediated signaling of V2 vasopressin receptor. *Proceedings of the National Academy of Sciences*, *102*(5), 1448–1453.

Reston, J. (1994). *Galileo: A life*. New York: HarperCollins.

Roberts, J. (2009). An undergraduate journal club experience: A lesson in critical thinking. *Journal of College Science Teaching*, *38*(3), 28–31.

Roth, W. M., Bowen, G. M., & McGinn, M. K. (1999). Differences in graph-related practices between high school biology textbooks and scientific ecology journals. *Journal of Research in Science Teaching*, *36*(9), 977–1019.

Rovito, S. M., Parra-Olea, G., Vásquez-Almazán, C. R., Papenfuss, T. J., & Wake, D. B. (2009). Dramatic declines in neotropical salamander populations are an important part of the global amphibian crisis. *Proceedings of the National Academy of Sciences*, *106*(9), 3231–3236.

Russell, D. R. (1997). Rethinking genre in school and society: An activity theory analysis. *Communication Abstracts*, *21*(2), 504–554.

Sampson, V. D., & Clark, D. B. (2006). Assessment of argument in science education: A critical review of the literature. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), *Proceedings of ICLS 2006* (pp. 655–661). Presented at the The International Conference of the Leaning Sciences, Bloomington, IN: International Society of the Learning Sciences.

Sampson, V. D., & Clark, D. B. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education*, *92*(3), 447–472.

Samuels, S. J., Tennyson, R., Sax, L., Mulcahy, P., Schermer, N., & Hajovy, H. (1988). Adults' use of text structure in the recall of a scientific journal article. *The Journal of Educational Research*, *81*(3), 171–174.

Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition & Instruction*, *23*(1), 23–55.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, *82*(3), 498.

Shirtcliff, E. A., Coe, C. L., & Pollak, S. D. (2009). Early childhood stress is associated with elevated antibody levels to herpes simplex virus type 1. *Proceedings of the National Academy of Sciences*, *106*(8), 2963–2967.

Shymansky, J. A., Yore, L. D., & Good, R. (1991). Elementary school teachers' beliefs about and perceptions of elementary school science, science reading, science textbooks, and supportive instructional factors. *Journal of Research in Science Teaching*, *28*(5), 437–454.

Siegel, H. (1988). *Educating reason: Rationality, critical thinking, and education*. New York: Routledge.

Smith, G. R. (2001). Guided literature explorations: Introducing students to the primary literature. *Journal of College Science Teaching*, *30*(7), 465–69.

Suppe, F. (1998). The structure of a scientific paper. *Philosophy of Science*, *65*(3), 381–405.

Sutton, C. (2003). New perspectives on language in science. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education: Part 1* (pp. 27–38). Dordrecht, The Netherlands: Kluwer Academic.

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

Swales, J., & Najjar, H. (1987). The writing of research article introductions. *Written Communication*, *4*(2), 175–191.

Tapper, J. (2004). Student perceptions of how critical thinking is embedded in a degree program. *Higher Education Research & Development*, *23*(2), 199–222.

Tenopir, C., King, D. W., Edwards, S., & Wu, L. (2009). Electronic journals and changes in scholarly article seeking and reading patterns. *Aslib Proceedings*, *61*(1), 5–32.

Teufel, S. (1999). *Argumentative zoning: Information extraction from scientific articles* (Doctoral dissertation). University of Edinburgh. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.105.485

Thompson, D. K. (1993). Arguing for experimental "facts" in science: A study of research article Results sections in biochemistry. *Written Communication*, *10*(1), 106–128.

Topping, K. J. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education*, *32*(3), 321.

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

Toulmin, S. E., Rieke, R. D., & Janik, A. (1979). *An introduction to reasoning*. New York: Macmillan.

van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker, R. M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 1–14). Dordrecht, The Netherlands: Kluwer Academic.

van Lacum, E., Ossevoort, M., Buikema, H., & Goedhart, M. (2012). First experiences with reading primary literature by undergraduate life science students. *International Journal of Science Education*, *34*(12), 1795–1821.

van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. London: Academic Press.

von Humboldt, W. (1903). Über die innere und äussere Organisation der höheren wissenschaftlichen Anstalten in Berlin. In A. Leitzmann (Ed.), *Gesammelte Schriften X* (pp. 250–260). Berlin: Behr.

Watson, J. D., & Crick, F. H. C. (1953). A structure for deoxyribose nucleic acid. *Nature*, *171*(4356), 737–738.

Wenk, L., & Tronsky, L. (2011). First-year students benefit from reading primary research articles. *Journal of College Science Teaching*, *40*(4), 60–67.

Westfall, R. S. (1983). *Never at rest: A biography of Isaac Newton*. Cambridge: Cambridge University Press.

White, B., Stains, M., Escriu-Sune, M., Medaglia, E., Rostamnjad, L., Chinn, C., & Sevian, H. (2011). A novel instrument for assessing students' critical thinking abilities. *Journal of College Science Teaching*, *40*(5), 103–107.

Williams, I. A. (1999). Results sections of medical research articles: Analysis of rhetorical categories for pedagogical purposes. *English for Specific Purposes*, *18*(4), 347–66.

Witte, A., Fobker, M., Gellner, R., Knecht, S., & Flöel, A. (2009). Caloric restriction improves memory in elderly humans. *Proceedings of the National Academy of Sciences*, *106*(4), 1255–1260.

Yudkin, B. (2006). *Critical reading: Making sense of papers in life science and medicine*. London: Routledge.

Ziman, J. M. (1969). Information, communication, knowledge. *Nature*, *224*(5217), 318–324.

Summary

The aim of this thesis is to develop and evaluate a teaching strategy for introducing first-year undergraduate students to research articles. In science, the research article is the most important type of primary literature. By primary literature we mean reports of original observations, theories, or opinions, written for peers in the scientific community. The ability to read research articles, an important aspect of scientific literacy, is particularly beneficial for students who will become scientists. The ability to read research articles may also be useful in other careers where professionals serve as intermediaries between science and society at an academic level.

However, reading research articles is a demanding task for novice readers. This is caused by specifics of the genre, such as the language (e.g. the frequent use of technical terms), persuasiveness (students are used to textbooks in which factual statements are presented; in contrast, research articles aim to persuade the reader that its claims are correct), and contextual features (the relevance of research is often not explicitly mentioned). This suggests that higher education institutions should use specific teaching strategies that will introduce science students to primary literature and will improve their level of scientific literacy.

We described several studies in which college and university students learn to read research articles. In almost every study students are guided through research articles via guided reading (e.g. answering questions about certain aspects of the article), followed by a group discussion. However, the observed progression in reading ability is often based on questionnaires in which students assess themselves. Furthermore, the educational aims are often poorly defined. Therefore, it is unclear what kind of skills or abilities of students the authors want to improve. This prompted us to use a more systematic approach.

The primary research question of this thesis was: *How can undergraduate life science students be taught to read research articles?* To answer

this question, we followed an educational design research approach. Our goal with this approach was to generate, articulate, and test design principles. The research was performed in the course Biomedical Research. This course was part of the last quarter of the first-year programs of Biology and Life Science and Technology at the University of Groningen. The subjects of the course were related to physiology and pharmacology (with a focus on the cardiovascular system). The course was built around lectures, lab work, and tutor group meetings.

In a preliminary investigation we focused on students' ability to identify rhetorical moves. A rhetorical move is a section of a text that performs a specific communicative function. We call the arrangement of rhetorical moves in a text the rhetorical structure. We presumed, based on ideas from the field of genre analysis, that there is a positive relation between reading ability and rhetorical consciousness (recognizing the rhetorical structure of a genre-specific text). So, in our preliminary investigation we determined if life science students are able to identify two important rhetorical moves: conclusions and the supports that are used to justify them. This preliminary investigation took place during the 2008 edition of the course Biomedical Research.

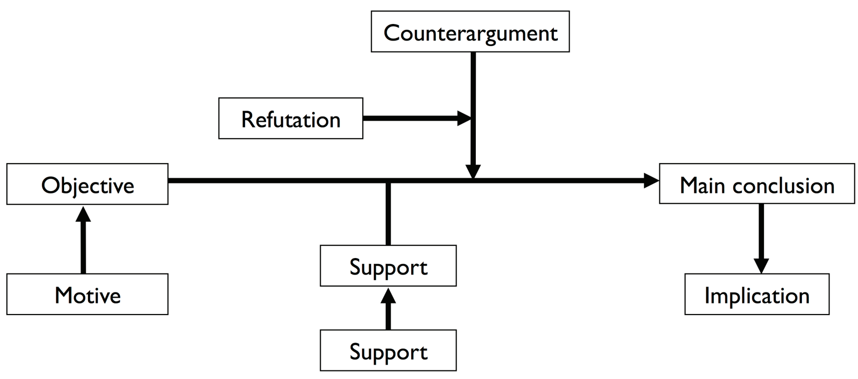
We analyzed in detail the identification of conclusions and supports in three different research articles by 20 students. We also compared their answers with the answers of two expert readers. Students and experts agreed on the most important conclusions of the articles they read. However, students identified a wide range of sentences that were not seen as conclusions by the experts. The supports students mentioned mostly matched the conclusions they identified. Students often failed to identify important supports for a particular conclusion. Furthermore, we conducted task-based/think-aloud interviews with four students to gain more insight into their reading behaviors and to determine the criteria they used to identify conclusions and supports. Our results showed that students and experts used different criteria to identify conclusions and supports. We also found that the interviewed students read their articles selectively. They skipped or skimmed for example the Method section, because they thought it was not relevant for the assignment. Our students – with one exception – read their articles sequentially at first (reading the parts of a text in order). But when making the assignment, students read the article non-sequentially.

The results of our preliminary investigation led to a teaching strategy that aimed to improve students' ability to identify rhetorical moves. The strategy was based on two design principles: a focus on rhetorical structure and cognitive apprenticeship. Cognitive apprenticeship involves (among other things) making processes of a task visible and using authentic contexts. The design principle of cognitive apprenticeship led to two derived design principles: authenticity and interactivity.

To realize the first design principle, we developed (mainly based on genre analysis studies) a set of seven rhetorical moves, which describe the most important aspects of the rhetorical structure of a research article: motive (why was the research done?), objective (what do the authors want to know?), main conclusion (what is the main outcome of the research?), implication (what are the consequences of the research?), supports (all the elements the authors use to justify their main conclusion), counterarguments (statements that weaken or discredit the main conclusion), and refutations (statements that weakens or refute a counterargument). We combined the concept of rhetorical moves with ideas from argumentation theory to construct an argumentation scheme that depicts the seven rhetorical moves and their relations. This scheme forms together with the set of seven rhetorical moves the Scientific Argumentation Model (SAM).

To realize the first derived design principle, we used authentic, non-adapted research articles. To realize the second derived design principle, we used tutor groups as our educational setting. In such a setting, students are stimulated to discuss what they have read with others. Also, the tutor can show the students how to read research articles.

Then, the teaching strategy was implemented in the course module Reading Research Articles (which was part of the 2009 edition of the course Biomedical Research). Each week, students received instructions from the tutor (a senior student) and an assignment. For the assignment, students – among other tasks – identified certain moves in a research article. The research articles' concepts were discussed during preceding lectures. The tutor group meetings were intended to let students discuss the content of the articles and the answers to their assignments.



The SAM scheme.

To test the effectiveness of our teaching strategy, we made use of a pre-test and post-test in which we determined the ability of 108 students to identify rhetorical moves (i.e. their rhetorical consciousness). Furthermore, we determined via questionnaires if there was a change in students' reading behaviors during the module.

The results showed that our teaching strategy improves undergraduate students' rhetorical consciousness of research articles. More students identified the motive, the objective, the main conclusion, and the implication at the end of the module. Regarding the identification of counterarguments and supports, there was, in general, no improvement between the pre-test and post-test. According to their own estimates, students made the post-test assignment (including reading the article) in less time and they adopted more expert-like reading behaviors: they read more selectively and non-sequentially.

We also investigated during the 2009 edition how certain features of rhetorical moves influence their identification by students. For this purpose, we distinguished three types of features: content-based, organizational, and lexical features. Content-based features relate to the function of a move and may be similar to the abovementioned descriptions of our seven moves. Organizational features describe the location of the move in the research article. For example, some moves (e.g. the objective) can always be found in the Introduction section. Lexical features are words or phrases that may trigger the reader to identify a certain statement as move. We analyzed written answers of five students to four assignments about the identification of rhetorical moves. We also analyzed transcriptions of recorded conversations from tutor group meetings in which 10 students discussed these assignments with the tutor.

We found that students were quite successful in identifying the motive, the objective, and the main conclusion (moves with pronounced lexical, organizational, and content-based features). In contrast, the identification of moves with less pronounced lexical and organizational features (implications, supports, counterarguments, and refutations) was more problematic for students. So, it seems that students rely mainly on lexical and organizational features and less on content-based features when identifying moves. Based on these results, we hypothesized that students mainly use lexical (and sometimes organizational) features to identify the move and then use content-based features to check if their assessment is correct.

During the same edition of the module, we also determined students' ability to construct a SAM scheme. We were primarily interested in students' presentation of supports. In the SAM scheme, supports are represented in chains. These support chains visualize the different epistemic levels by which scientific data can be represented. For example, the base of a support chain is often an inscription (a figure, table, or other non-textual representation). This support may be connected to a text fragment from the Results section that contains a relatively simple interpretation of the inscription. This simple interpretation may be connected to a text fragment containing a more complex interpretation or generalization.

Students were assigned to construct a SAM scheme. We collected SAM schemes of 73 students and then analyzed them by comparing them with our analysis of the research article. Based on this comparison, we concluded that students were generally able to construct a SAM scheme. Relating inscriptions with text fragments from the Results section went often well. Students also identified the most important figures as supports. We observed that students were less successful in connecting text fragments from the Results section with the Discussion section. Often, connections were not made or the connections did not correspond with our own analysis. Students seem to be not fully aware of the different epistemic levels in a research article and how these levels are connected to each other. Our results suggest that students' awareness of differences in epistemic levels could be further developed. Roughly half of the students indicated that the construction of an argumentation scheme (such as the SAM scheme) helps them with understanding a research article. Only a minority of the students indicated that they will use the scheme more often.

During the next edition of our module, in 2010, we did a baseline measurement of students' ability to critically evaluate research articles. We gave 50 students two articles containing contradictory evidence about the relation between treatment with vitamin C and the reduction of oxidative stress. Students individually made a number of assignments and discussed the articles and assignments during tutor group meetings. For the assignment, students were asked to evaluate the articles critically and resolve the contradictory evidence. The analysis of students' answers showed that a majority critically evaluated the articles by providing additional counterarguments not given by the authors. However, these counterarguments were often rather non-specific and related to supposed observational errors (e.g. sampling errors). Most students thought that the data in the two articles showed that vitamin C reduces levels of oxidative stress, even though both studies did not directly observe such effect. We observed that a number of students did not appreciate some of the rhetorical strategies used by the authors of the two articles (e.g. the use of words such as "may" and "might" to express (un)certainty). This suggests that students' ability to evaluate research articles is influenced by their epistemological beliefs (beliefs about the nature of knowledge and the process of knowing).

We concluded that our teaching strategy (based on two design principles: the focus on rhetorical structure and cognitive apprenticeship) is useful for teaching students how to read research articles. In further studies, it is recommended to investigate how text comprehension, prior knowledge, and rhetorical consciousness influence each other. The relation between rhetorical consciousness and epistemological beliefs could also be elucidated. We propose a curriculum in which students learn to read and write research articles. This curriculum consists of three phases: modeling, joint construction of text (together with more experienced writers), and independent construction of text. In this curriculum reading, writing, argumentation, and the development of epistemological beliefs should form an integrated whole.

Samenvatting

Het doel van dit proefschrift is het ontwikkelen en evalueren van een onderwijsstrategie die eerstejaarsstudenten levenswetenschappen laat kennismaken met onderzoeksartikelen. Het onderzoeksartikel is in de wetenschap het belangrijkste type van primaire literatuur. Onder primaire literatuur verstaan we verslagen van originele observaties, theorieën of opinies die geschreven zijn voor andere wetenschappers. Vaardig zijn in het lezen van onderzoeksartikelen, een belangrijk aspect van wetenschappelijke geletterdheid, is in het bijzonder nuttig voor studenten die wetenschapper worden. Vaardig zijn in het lezen van wetenschappelijke artikelen kan ook nuttig zijn in beroepen waarbij afgestudeerden op een academisch niveau als intermediair fungeren tussen wetenschap en samenleving.

Het lezen van wetenschappelijke artikelen is echter een uitdagende taak voor onervaren lezers. Dit wordt veroorzaakt door eigenschappen van het genre, zoals de taal (bijvoorbeeld het regelmatige gebruik van jargon), persuasiviteit (studenten zijn gewend aan leerboeken waarin feitelijke beweringen worden gedaan; onderzoeksartikelen hebben daarentegen als doel om de lezer te overtuigen dat de conclusies correct zijn) en contextuele eigenschappen (de relevantie van het onderzoek wordt vaak niet expliciet gemaakt). Dit suggereert dat hoger-onderwijs-instellingen specifieke onderwijsstrategieën moeten gebruiken die studenten primaire literatuur leren lezen en hun wetenschappelijke geletterdheid verbetert.

We beschreven verschillende studies waarin studenten van *colleges* (bepaalde Amerikaanse onderwijsinstellingen) en universiteiten onderzoeksartikelen leren lezen. In bijna elke studie lezen studenten het onderzoeksartikel via *guided-reading* (het beantwoorden van vragen over bepaalde aspecten van het artikel), gevolgd door een groepsdiscussie. Maar de vooruitgang in leesvaardigheid is vaak gebaseerd op vragenlijsten waarin studenten zichzelf beoordelen. Ook worden de onderwijsdoelen vaak slecht gedefinieerd. Hierdoor is het onduidelijk welke vaardigheden van studenten de auteurs willen verbeteren. Dit zette ons er toe aan om een meer systematische benadering toe te passen.

De primaire onderzoeksvraag van dit proefschrift was: *Hoe kunnen we eerstejaarsstudenten levenswetenschappen onderzoeksartikelen leren lezen?* Voor het beantwoorden van deze vraag maakten we gebruik van educatief ontwerponderzoek. Ons doel hierbij was het genereren, articuleren en testen van ontwerpprincipes. Het onderzoek is uitgevoerd binnen het vak Biomedisch onderzoek. Dit vak was onderdeel van het laatste kwartaal van de eerstejaars programma's van de bacheloropleidingen Biologie en *Life Science and Technology* van de Rijksuniversiteit Groningen. De onderwerpen van de cursus hadden betrekking op fysiologie en farmacologie (met een focus op het cardiovasculaire systeem). De cursus bestond uit colleges, practica en tutorgroepbijeenkomsten.

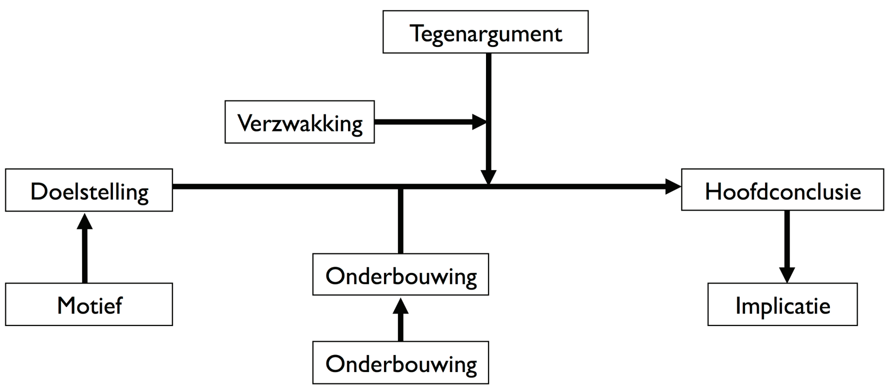
In een vooronderzoek hebben we gekeken naar de vaardigheid van studenten in het identificeren van retorische *moves*. Een retorische *move* is een tekstfragment dat een specifieke communicatieve functie heeft. We noemen de schikking van retorische *moves* in een tekst de retorische structuur. We veronderstelden, gebaseerd op ideeën uit de genreanalyse, dat er een positieve correlatie bestaat tussen leesvaardigheid en retorisch inzicht (het herkennen van de retorische structuur van een genre-specifieke tekst). Daarom hebben we in ons vooronderzoek bepaald of studenten levenswetenschappen in staat zijn om twee belangrijke retorische *moves* te identificeren: conclusies en de daarbij behorende onderbouwingen. Dit vooronderzoek vond plaats binnen de 2008-editie van de cursus Biomedisch onderzoek.

We analyseerden bij 20 studenten in detail de identificatie van conclusies en onderbouwingen in drie verschillende onderzoeksartikelen. Ook vergeleken we hun antwoorden met de antwoorden van twee expert-lezers. Studenten en experts waren het eens over de meest belangrijke conclusies van de gelezen artikelen. Studenten identificeerden echter een groot aantal zinnen die door de experts niet beschouwd werden als conclusies. De onderbouwingen die studenten noemden, pasten bij de conclusies die ze identificeerden. Studenten slaagden er vaak niet in om belangrijke onderbouwingen voor een bepaalde conclusie te identificeren. We voerden ook *task-based/think-aloud*-interviews uit met vier studenten om meer inzicht te krijgen in hun leesgedrag en om criteria te bepalen waarmee ze conclusies en onderbouwingen identificeerden. Onze resultaten toonden aan dat studenten en experts verschillende criteria gebruikten om conclusies en onderbouwingen te identificeren. We ontdekten ook dat de geïnterviewde studenten hun artikelen op een selectieve manier lazen. Ze scanden bijvoorbeeld de Methode-sectie of sloegen deze over. Onze studenten – op één uitzondering na – lazen in eerste instantie hun artikelen sequentieel (het op volgorde lezen van de onderdelen van een tekst). Maar bij het maken van de opdracht lazen de studenten non-sequentieel.

De resultaten van ons vooronderzoek leidden tot een onderwijsstrategie die als doel had om de vaardigheid van studenten in het identificeren van retorische *moves* in onderzoeksartikelen te verbeteren. Deze strategie was gebaseerd op twee ontwerpprincipes: een focus op retorische structuur en *cognitive apprenticeship*. *Cognitive apprenticeship* houdt onder andere het zichbaar maken van de processen van een taak en het gebruik van authentieke contexten in. Het ontwerpprincipe van *cognitive apprenticeship* leidde tot twee afgeleide ontwerpprincipes: interactiviteit en authenticiteit.

Om het eerste ontwerpprincipe te realiseren, ontwikkelden we (hoofdzakelijk gebaseerd op genreanalyse-studies) een verzameling retorische *moves* die de meest belangrijke aspecten van de retorische structuur van een onderzoeksartikel beschrijven: motief (waarom is het onderzoek uitgevoerd?), doelstelling (wat willen de auteurs te weten komen?), hoofdconclusie (wat is de belangrijkste uitkomst van het onderzoek?), implicatie (wat zijn de consequenties van het onderzoek?), onderbouwingen (alle elementen die gebruikt worden om de hoofdconclusie te onderbouwen), tegenargumenten (beweringen die de hoofdconclusie verzwakken of ondergraven) en verzwakkingen (beweringen die tegenargumenten verzwakken of weerleggen). We combineerden het concept van retorische *moves* met noties uit de argumentatietheorie voor het construeren van een argumentatieschema waarin de zeven retorische *moves* en hun relaties worden weergegeven. Dit schema vormt samen met de serie van zeven retorische *moves* het *Scientific Argumentation Model* (SAM).

Om het eerste afgeleide ontwerpprincipe te realiseren, gebruikten we authentieke, niet-aangepaste onderzoeksartikelen. Om het tweede afgeleide ontwerpprincipe te realiseren, werden tutorgroepen gebruikt als



Het SAM-schema.

onderwijsomgeving. In tutorgroepen worden studenten gestimuleerd om datgene wat ze gelezen hebben te bediscussiëren met anderen. Ook kan de tutor aan de studenten laten zien hoe onderzoeksartikelen gelezen moeten worden.

De onderwijsstrategie werd vervolgens geïmplementeerd in de cursusmodule Onderzoeksartikelen lezen (onderdeel van de 2009-editie van het vak Biomedisch onderzoek). Elke week ontvingen studenten instructies van de tutor (een ouderejaarsstudent) en een opdracht. Voor de opdracht moesten studenten onder andere retorische *moves* identificeren in een onderzoeksartikel. De concepten die in de onderzoeksartikelen voorkwamen werden in voorafgaande colleges behandeld. De tutorgroepbijeenkomsten waren bedoeld om studenten te laten discussiëren over de inhoud van de artikelen en hun antwoorden op de opdrachten.

Om de effectiviteit te testen van onze onderwijsstrategie maakten we gebruik van een pre-test en post-test waarbij we bij 108 studenten bepaalden in hoeverre ze retorische *moves* kunnen identificeren (oftewel hun retorisch inzicht). Met vragenlijsten bepaalden we ook of het leesgedrag van studenten veranderde tijdens de module.

De resultaten toonden aan dat onze onderwijsstrategie het retorisch inzicht van eerstejaarsstudenten verbetert wat betreft onderzoeksartikelen. Meer studenten slaagden er in het motief, de doelstelling, de hoofdconclusie en de implicatie te identificeren aan het eind van de module. Er was over het algemeen geen verschil tussen de pre-test en post-test bij de identificatie van tegenargumenten en onderbouwingen. Volgens hun eigen inschattingen maakten de studenten de post-test-opdracht (inclusief het lezen van het artikel) in minder tijd en vertoonden ze meer expertachtig leesgedrag: ze lazen meer selectief en non-sequentieel.

Tijdens de 2009-editie onderzochten we eveneens hoe bepaalde kenmerken van retorische *moves* hun identificatie door studenten beïnvloedden. Met dit doel voor ogen maakten we onderscheid tussen drie soorten kenmerken: inhoudelijke, organisatorische en lexicale kenmerken. Inhoudelijke kenmerken zijn gerelateerd aan de functie van een *move* en kunnen overeenkomen met de bovengenoemde beschrijvingen van onze zeven *moves*. Organisatorische kenmerken beschrijven de locatie van een *move* in het onderzoeksartikel. Sommige *moves*, zoals de doelstelling, komen bijvoorbeeld altijd voor in de Introductie-sectie. Lexicale kenmerken zijn woorden of frasen die de lezer helpen bij het identificeren van een bepaalde bewering als retorische *move*. We analyseerden de schriftelijke antwoorden van vijf studenten op vier opdrachten over het identificeren van retorische *moves*. We analyseerden ook transcripties van opnames van gesprekken in tutorgroepbijeenkomsten waarin 10 studenten deze opdrachten bediscussieerden met hun tutor.

We constateerden dat studenten vrij succesvol waren in het identificeren van het motief, de doelstelling en de hoofdconclusie (*moves* met uitgesproken lexicale, organisatorische en inhoudelijke kenmerken). De identificatie van *moves* met minder uitgesproken lexicale en organisatorische kenmerken (implicaties, onderbouwingen, tegenargumenten en verzwakkingen) was echter problematischer voor studenten. Het lijkt er daarom op dat studenten bij het identificeren van *moves* vooral gebruik maken van de lexicale en organisatorische kenmerken en minder van inhoudelijke kenmerken. Op basis van deze resultaten veronderstellen we dat studenten vooral lexicale (en soms organisatorische) kenmerken gebruiken om de *move* te identificeren en dan inhoudelijke kenmerken gebruiken om te controleren of hun beoordeling juist is.

Tijdens dezelfde editie van de module bepaalden we ook in hoeverre studenten een SAM-schema kunnen construeren. We waren voornamelijk geïnteresseerd in de weergave van onderbouwingen door studenten. Onderbouwingen worden in het SAM-schema weergegeven in ketens. Deze ketens van onderbouwingen visualiseren de verschillende epistemologische niveaus waarop wetenschappelijke data weergegeven kunnen worden. De basis van een keten van onderbouwingen is bijvoorbeeld vaak een inscriptie (een figuur, tabel of een andere non-tekstuele representatie). Deze onderbouwing kan verbonden worden met een tekstfragment uit de Resultaten-sectie dat een relatief simpele interpretatie bevat van de inscriptie. Deze simpele interpretatie kan verbonden worden met een tekstfragment dat een meer complexe interpretatie of generalisatie bevat.

Studenten werd de opdracht gegeven om een SAM-schema te construeren. We verzamelden SAM-schema's van 73 studenten en analyseerden deze vervolgens door ze te vergelijken met onze eigen analyse van het onderzoeksartikel. Op basis van deze vergelijking concludeerden we dat studenten over het algemeen in staat waren om een SAM-schema te construeren. Het verbinden van inscripties met tekstfragmenten uit de Resultaten-sectie ging vaak goed. Studenten identificeerden als onderbouwing ook de belangrijkste figuren. We constateerden dat studenten minder succesvol waren met het verbinden van tekstfragmenten uit de Resultaten-sectie met de Discussie-sectie. Verbindingen werden vaak niet gemaakt of ze kwamen niet overeen met onze analyse. Studenten leken zich niet helemaal bewust te zijn van de verschillende epistemologische niveaus in een onderzoeksartikel en hoe deze niveaus met elkaar samenhangen. Onze resultaten suggereren dat studenten meer bewust gemaakt kunnen worden van de verschillen in epistemologische niveaus. Ruwweg de helft van de studenten gaf aan dat het construeren van een argumentatieschema (zoals het SAM-schema) hen hielp met het begrijpen van een onderzoeksartikel. Een minderheid van de studenten gaf aan dat ze het schema vaker gaan gebruiken.

Tijdens de daaropvolgende editie van onze module in 2010, voerden we een nulmeting uit waarbij we bepaalden in hoeverre studenten onderzoeksartikelen kritisch kunnen beoordelen. We gaven 50 studenten twee artikelen die strijdig bewijs bevatten over de relatie tussen het toedienen van vitamine C en vermindering van oxidatieve stress. Studenten maakten individueel een aantal opdrachten en bediscussieerden de onderzoeksartikelen en antwoorden op de opdrachten tijdens tutorgroepbijeenkomsten. Bij de opdrachten werd de studenten gevraagd om de onderzoeksartikelen kritisch te beoordelen en het strijdige bewijs op te helderen. De analyse van de antwoorden van de studenten toonde aan dat een meerderheid van de studenten de artikelen kritisch beoordeelde door het geven van extra tegenargumenten die niet genoemd werden door de auteurs. Deze tegenargumenten waren echter vaak niet heel specifiek en hadden betrekking op vermeende observationele fouten (bijvoorbeeld steekproeffouten). De meeste studenten zeiden dat de data in de twee artikelen lieten zien dat oxidatieve stress verminderd wordt door vitamine C, terwijl in beide studies dit effect niet direct waargenomen is. We zagen dat een aantal studenten het gebruik van bepaalde retorische strategieën door de auteurs van beide artikelen niet naar waarde konden schatten (zoals het gebruik van woorden zoals “zou kunnen” en “wellicht” om (on)zekerheid uit te drukken). Dit suggereert dat de vaardigheid van studenten in het kritisch beoordelen van onderzoeksartikelen beïnvloed wordt door hun epistemologische opvattingen (opvattingen over de aard van kennis en het proces van weten).

We concludeerden dat onze onderwijsstrategie (gebaseerd op twee ontwerpprincipes: een focus op retorische structuur en *cognitive apprenticeship*) nuttig is bij het leren lezen van onderzoeksartikelen aan studenten. Als vervolgonderzoek stellen we voor om te bestuderen hoe het begrijpen van tekst, voorkennis en retorisch inzicht elkaar beïnvloeden. De relatie tussen retorisch inzicht en epistemologische opvattingen kan eveneens onderzocht worden. We stellen een curriculum voor waarin studenten leren hoe ze onderzoeksartikelen moeten lezen en schrijven. Dit curriculum bestaat uit drie fasen: modelleren, de gezamenlijke constructie van tekst (samen met meer ervaren schrijvers), en het zelfstandig construeren van een tekst. In dit curriculum zouden lezen, schrijven, argumenteren en de ontwikkeling van epistemologische opvattingen een geïntegreerd geheel moeten vormen.

Appendix 1
Information Sheets

This appendix contains the English translation of the information sheets. The original information sheets were in Dutch. Examples are taken from Ren et al. (2005), Cela-Conde et al. (2009), Chen et al. (2009), Klasson et al. (2009), Koffie et al. (2009), Lu and Collins (2009), Rovito, Parra-Olea, Vásquez-Almazán, Papenfuss, and Wake (2009), Shirtcliff, Coe, and Pollak (2009), and Witte, Fobker, Gellner, Knecht, and Flöel (2009).

MOTIVE Can be found in: Introduction	
• Why is the research important according to the authors? • The motive leads to the objective. • Possible motives: » There is something unknown » There are conflicting results » There is a specific problem that needs a solution	Examples: *" …potential benefits of specific "brain-healthy diets" have been proposed, but have not been confirmed (…) by animal experiments and human epidemiological studies. Evidence (…) in humans is still missing"* *"Although a few studies have indicated preferential phosphorylation and desensitization of one or another 7TM receptor by a particular GRK, specialized functions of these enzymes have not been clearly defined."* *"Work by several groups has shown a decrease in dendritic spine density and synaptophysin-positive synapses radiating out from the surface of plaques in mouse models of [Alzheimer's disease]. Whether this is caused by fibrillar plaques or soluble oligomeric Aβ is controversial."* *"Bacterial infections are responsible for significant morbidity and mortality in clinical settings (1). Many infections that would have been cured easily by antibiotics in the past now are resistant, resulting in sicker patients and longer hospitalizations."*

OBJECTIVE Can be found in: Introduction	

• What do the authors want to know?
• Indicates the function of the research. Possible functions: giving a description of a phenomenon, comparing, determining if something falls into a certain category, evaluating a theory or design, finding an explanation for a certain phenomenon, et cetera.
• May be formulated as a:
  » Research aim
  » Research question
  » Hypothesis
  » Design aim
	Examples:
	
	"Therefore, the aim of the present study was to elucidate cognitive effects of a diet low in calories (…) in healthy elderly individuals." (Research aim)
	
	"…in the present study we set out to answer the following important question: Once a 7TM receptor is phosphorylated by a GRK, is an invariant program of signaling events set in motion, or do different GRKs, perhaps by phosphorylating distinct sites on a receptor, engender distinct programs of regulatory activity?" (Research question)
	
	"Here, we test the hypothesis that oligomeric Aβ is directly synaptotoxic." (Hypothesis)
	
	"Therefore, by using a combination of engineered antibiotic-enhancing phage and antibiotics, we hoped to reduce the incidence of antibiotic resistance and enhance bacterial killing." (Design aim)

<table>
<tr><th colspan="2">MAIN CONCLUSION
Can be found in: Discussion</th></tr>
<tr><td>• The <u>main conclusion</u> relates to the <u>objective</u> of the study: authors answer the research question, they state to what extent the research aim was met, they decide if the hypothesis was correct or not, or they determine if the design has the desired characteristics.
• Often recognizable by signal words: “we conclude,” “our results show,” et cetera.
• The authors often indicate how certain they are of their <u>main conclusion</u>. This is done with words like: “might,” “suggest,” “prove,” “may,” et cetera.</td><td>Examples:

“To our knowledge, the current results provide first experimental evidence in humans that caloric restriction improves memory in the elderly.”

“…there might be different processing strategies for beauty in women and men.”

“The results of this study point to widespread and severe declines of upland salamanders at multiple sites in Guatemala and Mexico, including the most intensively-studied salamander transect in the neotropics.”

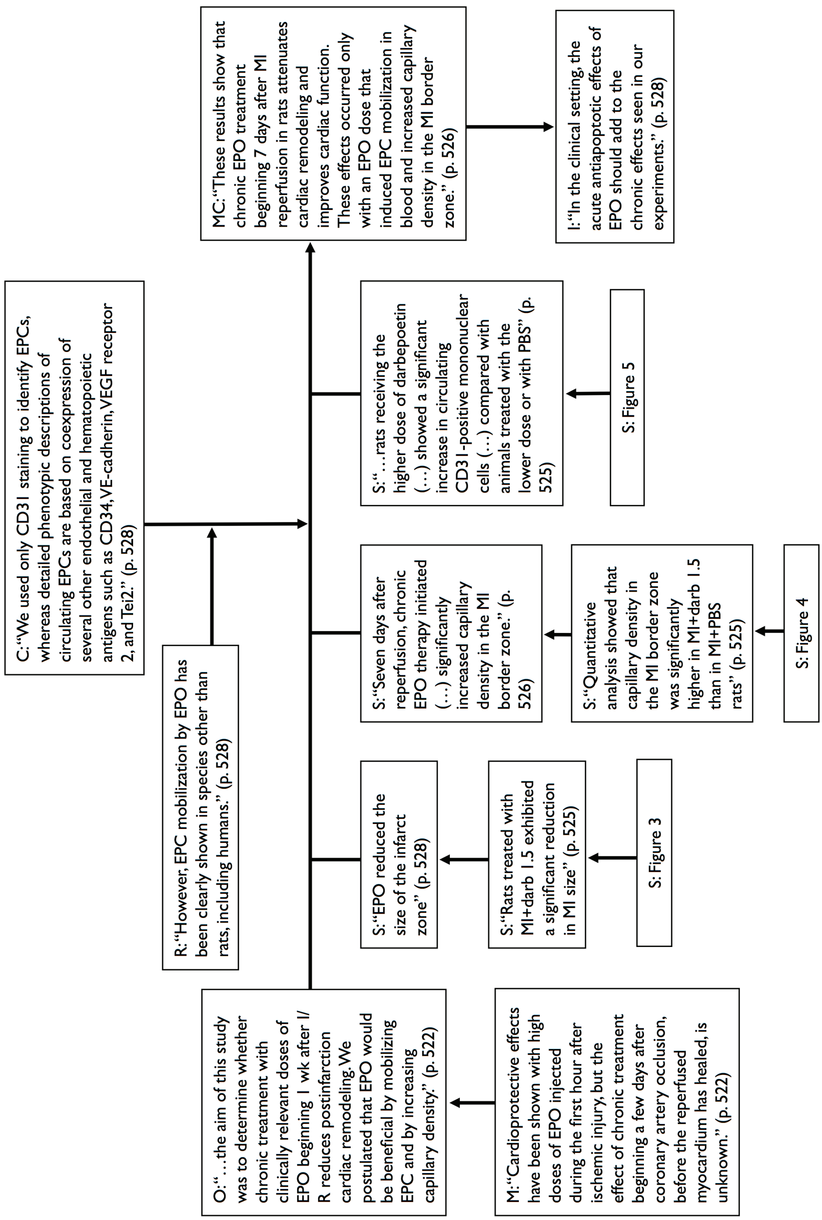
“Our general conclusion is that early life events are critical for creating the healthy foundation on which both emotional and physical well being is established.”</td></tr>
</table>

IMPLICATION Can be found in: Discussion	
• What is the meaning of the <u>main conclusion</u> in the study's context? • Possible <u>implications</u>: » Recommendation » Practical applications » Suggestions for future research	Examples: *"…the salamander populations of many upland species are in need of protection. Until the forces causing these declines are identified, however, an effective conservation strategy cannot be devised."* *"The present findings may help to develop new prevention and treatment strategies for maintaining cognitive health into old age."* *"An exciting avenue for future research is to identify the interacting endosymbiont-host proteins and determine whether these evolve by purifying, positive, or diversifying selection within* Wolbachia *subpopulations."*

SUPPORTS Can be found in: Results & Discussion	
• All the elements that the authors use to support their main conclusion. • Possible supports: » Own results (data, interpretations) » References to other research • Supports often occur in chains. A main conclusion is often supported by multiple support chains.	Example: *Table 1/Figure 2* ↓ *"During the 300- to 700-ms interval, activity was greater in parietal regions for stimuli rated as beautiful than for those rated as not beautiful. Moreover, whereas in women this activity was found in both hemispheres, in men it was mainly located in the right hemisphere."* ↓ *"Hence, it appears that women and men engage different strategies of spatial analysis during aesthetic preference activity."* (This chain supports the following main conclusion: "*...there might be different processing strategies for beauty in women and men.*")

COUNTERARGUMENTS & REFUTATIONS Can be found in: Discussion	
• Counterarguments argue against the main conclusion • Possible counterarguments: » Potential errors » Problems with the theory that underlies the research » Problems with the methodology » The main conclusion or supports are contradicting other research » An alternative explanation for the main conclusion » Limitations of the main conclusion • Counterarguments are sometimes called "limits" or "limitations." • Authors almost always try to refute counterarguments. They do this with refutations.	Examples: Counterargument: *"…dietary habits were self-reported only and thus prone to over- or underestimation"* Refutation: *"However, in the CR group, weight loss and BMI reduction demonstrated adherence to the intended dietary regime."* Counterargument: *"One alternative explanation for our findings is that institutional and abusive family settings may have resulted in higher levels of infection during childhood."* Refutation: *"Yet, our analyses indicated a similar overall prevalence of HSV infection across the 3 groups (…) This prevalence rate of HSV suggests that differences in levels of antibody more likely reflect functional differences in immune competence rather than differences in rates of exposure in postinstitutionalized or physically-abused youth."* Counterargument: *"It also appears that Nrf2$^{-/-}$ mice have lower basal levels of DA than the Nrf2$^{+/+}$ mice. This contradicts work of Pacchioni* et al. *(40), who showed no basal difference between Nrf2$^{+/+}$ and Nrf2$^{-/-}$ mice."* Refutation: *"This discrepancy may in part be caused by genetic background differences."*

Appendix 2
Example SAM Scheme



SAM scheme based on Prunier et al. (2007). Used as sample scheme in Chapter 6.

Appendix 3
Sentences in Bauersachs et al. (1999)

Abstract

a1 Background-Endothelial dysfunction of the peripheral vasculature is a well-known phenomenon in congestive heart failure that contributes to the elevated peripheral resistance; however, the underlying mechanisms have not yet been clarified.

a2 Methods and Results-Dilator responses, the expression of protein and mRNA of the endothelial nitric oxide synthase (eNOS), inducible NOS (iNOS), and soluble guanylate cyclase (sGC), and superoxide anion (O_2^-) and peroxynitrite production were determined in aortic rings from Wistar rats 8 weeks after myocardial infarction and compared with those in sham-operated animals.

a3 In rats with heart failure, the concentration-response curve of the endotheliumdependent vasodilator acetylcholine (after preconstriction with phenylephrine) was significantly shifted to the right, and the maximum relaxation was attenuated.

a4 Determination of expression levels of the 2 key enzymes for NO-mediated dilations, eNOS and sGC, revealed a marked upregulation of both enzymes in aortas from rats with heart failure, whereas iNOS expression was not changed.

a5 Pretreatment with exogenous superoxide dismutase partially restored the acetylcholine-induced relaxation in aortas from rats with heart failure.

a6 Aortic basal and NADH-stimulated O_2^- production assessed by use of lucigenin-enhanced chemiluminescence was significantly elevated in rats with chronic myocardial infarction.

a7 Peroxynitrite-mediated nitration of protein tyrosine residues was not different between the 2 groups of rats.

a8 Conclusions-These results demonstrate that endothelial dysfunction in ischemic heart failure occurs despite an enhanced vascular eNOS and sGC expression and can be attributed to an increase in vascular O_2^- production by an NADH-dependent oxidase.

a9 By inactivation of NO, O_2^- production appears to be an essential mechanism for the endothelial dysfunction observed in heart failure.

Introduction

i10 Endothelial dysfunction of the peripheral vasculature contributes to the elevated peripheral vascular resistance in patients with heart failure,[1-3] as well as several animal models of cardiac dysfunction.[4-6]

i11 However, the underlying mechanisms may be complex and have not yet been clarified.

i12 One attractive hypothesis appears to be a decrease in the production of endothelium-derived nitric oxide (NO).

i13 In a heart failure model of ventricular pacing in dogs, an endothelial hyporesponsiveness in the coronary circulation and an attenuated expression of the endothelial NO synthase (eNOS) in the aorta have been described.[7,8]

i14 In contrast, other studies reported enhanced basal production of NO in heart failure,[6,9] which might originate from the inducible NOS (iNOS) in the vasculature, because the expression of this high-output NO generating enzyme has been shown in hearts from patients with dilated cardiomyopathy.[10]

i15 In other pathophysiological states, such as hypercholesterolemia and hypertension, compelling evidence suggests that endothelial dysfunction results from increased vascular production of superoxide anion (O_2^-).[11-13]

i16 Because O_2^- rapidly scavenges NO within the vascular wall, a reduction of bioactive NO might occur despite an increased NO generation. [14,15]

i17 In patients suffering from heart failure, elevated levels of plasma lipid peroxides as a marker of oxidative stress have been observed.[16]

i18 This is further supported by the fact that the impaired flow-induced NO-mediated dilation in patients with heart failure can be restored by short-term treatment with high doses of the antioxidant vitamin C.[17]

i19 In addition, more recently, alterations of the effector system of NO, in particular a reduced expression of the cGMP-forming soluble guanylate cyclase (sGC), were identified as an important mechanism of dilator dysfunction in hypertension.[18]

i20 With regard to heart failure, no data are available on O_2^- production within the vascular wall or on the potential alterations of the expression of smooth muscle sGC.

i21 Moreover, the influence of heart failure on vascular NOS expression is still controversial.

i22 Chronic myocardial infarction in the rat is considered to be a useful model to study the pathophysiological sequelae of heart failure.

i23 Indeed, the beneficial effects of ACE inhibitors were predicted from results obtained in this experimental model.[19]

i24 In rats with heart failure due to myocardial infarction, endotheliumdependent relaxations, still normal at 1 week after coronary ligature, are reduced at 4 weeks and progressively worsen with time.[20]

i25 The aim of the present study was therefore to identify the potential mechanisms underlying endothelial dysfunction in heart failure by the simultaneous determination of endothelium- dependent dilator responses, the expression of the key enzymes of the NO/cGMP system, and O_2^- formation in the aorta of rats 8 weeks after myocardial infarction.

Results

r26 Global parameters of heart failure rats and sham-operated animals are shown in Table 2.

r27 Infarct size was 45±1%.

r28 Mean arterial blood pressure, left ventricular systolic pressure, and dP/dt_{max} were significantly lower in rats with chronic myocardial infarction, whereas left ventricular end-diastolic pressure was elevated.

r29 Plasma renin activity was significantly higher in rats with heart failure.

r30 Therefore, these rats demonstrated heart failure in a compensated stage.

r31 In phenylephrine-preconstricted aortic rings, acetylcholine elicited a concentration-dependent relaxation that was blunted in aortas from rats with cardiac dysfunction (Figure 1A).

r32 Acetylcholine-induced relaxations were mediated by NO because they were abolished after incubation with the NOS inhibitor N^G-nitro-L-arginine 0.3 mmol/L for 30 minutes (data not shown).

r33 Endothelium-independent relaxations induced by sodium nitroprusside were slightly but not significantly attenuated at lower concentrations in rats with heart failure, and maximum relaxation was not different (100%) in the 2 groups of rats (Figure 1B).

r34 To elucidate whether the attenuation of endotheliumdependent relaxation is the result of an alteration in the expression of the key enzymes of NO-mediated dilation, the expression of protein and mRNA of both eNOS and iNOS as well as sGC was determined in aortic segments from rats with heart failure and sham-operated animals by Western blot and RT-PCR.

r35 As shown in Figure 2, eNOS mRNA and protein levels were found to be significantly increased in aortas from rats with heart failure compared with sham-operated animals (2.9- and 2.1-fold increase, respectively, $P<0.05$, n=4), whereas the iNOS expression in the thoracic aorta, hardly detectable by Western blot analysis (Figure 3A), remained unchanged.

r36 In addition, Western blot analysis performed on whole aortic protein extracts showed that the protein level of the b1-subunit of the sGC was markedly enhanced in rats with myocardial infarction (Figure 3B, 2.5-fold increase, $P<0.05$, n=4), whereas the RT-PCR analysis failed to detect significant differences between infarcted and sham-operated animals (data not shown).

r37 Because NO- and cGMP-generating enzymes were found to be upregulated in rats with heart failure, we investigated the potential involvement of reactive oxygen species in the alteration of the endothelial function.

r38 The effects of radical scavengers were studied on the vascular reactivity and cGMP production.

r39 In phenylephrine-constricted rings, addition of exogenous SOD 600 U/mL elicited a relaxation that was significantly enhanced in aortic rings from rats with heart failure (83±3% versus 56±4%, $P<0.01$).

r40 Furthermore, in the presence of exogenous SOD 200 U/mL, the relaxation induced by submaximal concentrations of acetylcholine in aortas from rats with chronic cardiac dysfunction was significantly enhanced (Figure 4).

r41 Conversely, after inhibition of the endogenous SOD by use of DETC 1 mmol/L for 40 minutes, the acetylcholine-induced relaxation in aortic rings was markedly depressed in aortas from sham-operated rats and abolished in animals with chronic myocardial infarction (Figure 4).

r42 Basal levels of cGMP in aortas from rats with heart failure (2.4±0.4 pmol/mg protein) were not different from those in sham-operated animals (1.5±0.3 pmol/mg protein, n=6, Figure 5).

r43 Stimulation with sodium nitroprusside induced a marked increase in cGMP formation, and cGMP levels were lower in rats with cardiac dysfunction than in sham-operated animals.

r44 However, in the presence of the radical scavenger Tiron 10 mmol/L, sodium nitroprusside-induced cGMP formation was significantly enhanced in aortas from rats with heart failure (Figure 5).

r45 Finally, we assessed the production of O_2^- and of peroxynitrite generated by aortic rings by lucigenin- and luminolenhanced chemiluminescence, respectively.

r46 O_2^- release was greater in aortas from rats with chronic myocardial infarction (Figure 6A).

r47 Removal of the endothelium slightly but not significantly reduced radical production in both groups (Figure 6A).

r48 After addition of NADH 100 mmol/L, O_2^- formation was markedly stimulated and significantly higher in aortas from rats with heart failure (Figure 6B).

r49 The luminol-mediated luminescence as indicator of peroxynitrite formation was hardly detectable and not different in aortas from rats with chronic myocardial infarction and those from sham-operated animals.

r50 Moreover, Western blot analysis of aortic proteins with a specific antibody against nitrotyrosine to detect peroxynitrite-mediated nitration of tyrosine residues showed no increase in nitrotyrosine in rats with heart failure compared with sham-operated animals (data not shown).

Discussion

d51 In the present study, we observed a pronounced endothelial dysfunction in rats with chronic myocardial infarction despite a marked upregulation in the expression of 2 key enzymes of vasorelaxation: eNOS, regulating the synthesis of the most important vasodilator, NO, and its target enzyme in smooth muscle cells, sGC.

d52 Our data suggest that even this upregulation is not sufficient to compensate for the increased formation of O_2^-, which rapidly inactivates NO.

d53 Heart failure is associated with an endothelial dysfunction of coronary arteries as well as large conductance and peripheral arteries, with considerable implications for myocardial perfusion, cardiac workload, and peripheral vascular resistance. [1,7,26]

d54 From these functional studies, the mechanism underlying the reduction of agonist-stimulated dilator responses in heart failure has been proposed to be a defective production of endothelium-derived NO, and in a heart failure model of ventricular pacing in dogs as well as in monocrotaline-induced cardiac failure, a reduction of endothelial NO release was associated with an attenuated expression of eNOS.[8,27]

d55 Data on basal production of NO in heart failure have been controversial: using the amount of constriction in response to an NOS inhibitor as an indirect measure for basal NO release, some investigators found an increase[6,9] and speculated that expression of iNOS in the vasculature may be induced, as has been shown in hearts from patients with dilated cardiomyopathy.[10]

d56 However, other reports found no difference or even a decrease of basal NO formation in patients with heart failure.[28,29]

d57 Our results for the first time provide insights into the mechanisms of the alteration of endothelial function in heart failure after myocardial infarction, which represents the most important cause for cardiac failure in patients.

d58 Although in agreement with the results obtained in monocrotaline-induced heart failure,[27] iNOS expression was hardly detectable in rats with chronic myocardial infarction, we observed an unexpected marked increase in the expression of eNOS.

d59 The association of an increased eNOS expression with a marked attenuation of endothelium-dependent relaxation adds to the mounting evidence that enhanced NO formation or NOS expression does not necessarily imply improved dilator function but rather may even be detrimental[15] or at least a failed counterregulatory mechanism.

d60 In parallel with our results obtained in the aorta, in the myocardium of spontaneously hypertensive genetically heart failure-prone rats, an upregulation of cardiac eNOS expression has been observed; however, the functional consequences of this were not investigated.[30]

d61 The second key enzyme for endothelium-dependent dilation, the sGC in smooth muscle cells, is activated after binding of endothelium-derived NO to generate large amounts of cGMP.

d62 Recently, an attenuation of aortic sGC expression was recognized as a potential mechanism of reduced dilator response in aged spontaneously hypertensive rats.[18]

d63 In rats with chronic myocardial infarction, however, we observed an upregulation of sGC expression that was associated with a blunted cGMP formation in response to sodium nitroprusside.

d64 Because sGC activity is susceptible to superoxide[25] and cGMP production was restored by prior treatment with the radical scavenger Tiron, enhanced production of superoxide anions may be responsible for the reduced activity of sGC despite the increase in its expression.

d65 Enhanced degradation of cGMP due to increased phosphodiesterase activity[31] is not likely, because our experiments were performed in the continuous presence of a high concentration of a phosphodiesterase inhibitor.

d66 An enhanced O_2^- formation in rats with heart failure appears to account for the paradoxical attenuation of cGMP accumulation despite increased sGC expression.

d67 Elevated levels of plasma lipid peroxides in patients suffering from heart failure provide clear evidence of an enhanced oxidative stress under this condition.[16,32]

d68 In addition, the transition from hypertrophy to heart failure in coarctationinduced hypertension was associated with an increased oxidative stress and could be prevented by treatment with the antioxidant vitamin E, thus indicating a pathophysiological role for oxidative stress in the pathogenesis of heart failure.[33]

d69 High doses of vitamin C were able to restore the impaired NO-mediated dilation in patients with heart failure,[17] and in line with these observations, our results provide the first direct experimental evidence for an enhanced release of reactive oxygen species from the vasculature in chronic ischemic cardiac dysfunction.

d70 The source of superoxide formation appears to be vascular smooth muscle cells, because removal of the endothelium did not significantly attenuate radical production.

d71 Cultured and native vascular smooth muscle cells are able to generate superoxide in response to the vasoconstrictor peptide angiotensin II, which stimulates the expression of an NAD(P)H-dependent oxidase.[13,34]

d72 Plasma renin activity as well as tissue ACE activity is markedly elevated in heart failure.[35]

d73 Therefore, an enhanced formation of angiotensin II may lead to an enhanced vascular superoxide formation through the expression of an NAD(P)H-dependent oxidase in aortic smooth muscle cells.[13,34]

d74 Indeed, the observed upregulation of NADHdependent O_2^- formation in aortas from rats with chronic myocardial infarction suggests that this mechanism may be operative in ischemic heart failure.

d75 The deleterious role of O_2^- formation for endothelial function in ischemic cardiac dysfunction is further strengthened by the observation that exogenous SOD exerted a significantly greater relaxation in rats with chronic myocardial infarction and by the partial restoration of the acetylcholine-induced relaxation in the presence of SOD.

d76 An imbalance between NO and superoxide production with enhanced inactivation of NO, leading to a reduction of bioactive NO despite a normal or even increased generation of NO, has been associated with endothelial dysfunction and appears to be a common feature of many cardiovascular diseases, such as hypercholesterolemia and hypertension.[11-13,15,18]

d77 In addition, depending on the pathophysiological circumstances, NO and superoxide may react to the powerful oxidant peroxynitrite, which can form hydroxyl radicals and nitrate protein tyrosine residues.[15]

d78 However, we detected neither enhanced luminol chemiluminescence nor tyrosine nitration in rats with heart failure, so there was no hint for the formation of peroxynitrite.

d79 In conclusion, our data indicate that an increased NADHdependent vascular O_2^- generation represents an important mechanism for the endothelial dysfunction in heart failure by enhancing the inactivation of NO.

d80 Even a presumably counterregulatory upregulation of eNOS and sGC is not sufficient to restore endothelium-dependent relaxations.

Appendix 4
Assignments of the Module Reading Research Articles

This appendix contains the English translation of all the assignments used in the 2010 edition of the module Reading Research Articles. The regular assignment was made by all students. Each presentation assignment was assigned to one student.

Week 1

Assigment for every student:

Read the article of the week.

1. What is in research articles (in general) the function of the…
    - Abstract
    - Introduction
    - Method
    - Results
    - Discussion/Conclusion
2. Summarize each of the sections that are mentioned above in two or three sentences.
3. The article contains five figures and two tables. Which of these is the most important? Explain your answer.
4. Why do research articles always have the same structure?
5. What criteria should a good title of a research article meet? Does the title of the article of week meet these criteria?
6. The article of the week also contains a so-called “short/running title.” Try to identify this title and write it down.
7. How is study that is described in the article paid for?
8. How much time was there between the moment when the journal received the article and the moment when the journal decided to publish the article? How do you explain this lag?

9. There are also other types of articles published in scientific journals besides research articles (in which the results of a specific study are reported). Give three examples.
10. Examine the references in the article. What do you see? Do the authors cite mainly old literature or new literature? Do they cite publications they have written themselves? What is the benefit of self-citation?

Presentation assignments:

1. What does biomedical research precisely entail? What is the aim of biomedical research? Who are doing biomedical research? Which types of institutions are doing biomedical research? In which disciplines may biomedical research be divided? Give a short overview of the biomedical research that is done at the University of Groningen.
2. Most journals have an impact factor. What is an impact factor? How is an impact factor measured? What is in the cardiovascular field the journal with the highest impact factor? Show how you have determined this. Which of these journals are European and which of these journals are American? Compare the impact factors of these journals with the impact factors of journals from another research field (e.g. nephrology, pharmacology).
3. Who are the article of the week's authors? How many articles have they each published? Show how you have determined this. What is in general the role of the first, second, and third author? Is it possible to have two first authors? Finally, explain the Hirsch-index (or h-index). Try to calculate the h-index of the article of the week's third author.

Week 2

Assigment for every student:

Read the article of the week.

1. Choose five important concepts that are mentioned in the Introduction section and write down their meaning.
2. Describe in a couple of sentences the research field in which the authors place they study (you do not have to use external sources for this question).
3. There are many references to previous studies in the Introduction section. What is in general the function of these references?
4. Summarize in a couple of sentences what the authors say about these previous studies (you do not have to look up these studies).
5. What is the article's motive? Explain your answer.

6. What is the article's objective? Explain your answer.
7. Does the article's objective relate to the motive? Explain your answer.

Presentation assignments:

1. Choose three articles to which the authors refer in the Introduction section. Summarize these three articles. You do not have to read these articles in their entirely; reading the abstracts is enough. Try to show what the relation is between these three articles and the article of the week.
2. What is peer review? How does the peer review process take place? What is the role of respectively the author, editor, and reviewer? Who will be anonymous? Give two advantages of peer review and two disadvantages.
3. Summarize the complete article. Discuss the Introduction, Method, Results, and Discussion sections.

Week 3

Assigment for every student:

Read the article of the week.

1. What is/are the article's objective(s)?
2. What is/are the article's motive(s)?
3. What is/are the article's main conclusion(s)? Explain your answer.
4. How certain are the authors of their main conclusion? How did you infer this?
5. Can you find other conclusions in the Results or Discussion sections besides the main conclusion? Write these conclusions down.
6. What is/are the article's implication(s)? Explain your answer.
7. Does the article's objective relate to the motive? Explain your answer.
8. Choose a reference to a previous study that is mentioned in the Discussion section and check if the authors correctly represent the content of this article. You may find the article in PubMed. For this question you only have to read the article's Abstract. Explain why the authors represent the content of the article correctly (or not).

Presentation assignments:

1. Give a presentation about the (1) experimental design and (2) the materials and method. Discuss the following points. What was the studied population? Which variables did the authors study? What is HR variability and why do the authors measure it? What is the difference

between spectral and non-spectral analysis? Why is it important to decide beforehand which correlations you want to determine? Explain “cherry picking” and “data dredging.” Could you think of an experiment to evaluate the authors’ hypothesis in another way?
2. Present the results by means of the three tables and the figure. Describe for each table and figure what it depicts and what it means.
3. Present the article’s Discussion section. Discuss the main conclusion(s) and other conclusions, how the authors support their conclusions, if you agree with the authors, references to previous studies, and the implications.

Week 4

Assigment for every student:

Read the article of the week.

1. What is/are the article’s objective(s)?
2. What is/are the article’s motive(s)?
3. What is/are the article’s main conclusion(s)?
4. What is/are the article’s implication(s)?
5. With which supports from the Results and Discussion sections is/are the main conclusion(s) justified? Try to depict the supports in at least three chains (an example of a chain is given in the information sheets).
6. The authors use in the Discussion section also other studies as supports. Give two examples of these types of supports.
7. What is in your opinion the most important figure or table? Explain your answer.
8. Write down the counterargument(s) that the authors mention in the article. Write next to each counterargument – if mentioned – the corresponding refutation. Do you think that the authors sufficiently refute their counterarguments? Explain your answer.

Presentation assignments:

1. Give a presentation about organizing experiments with animals or humans. Can you just do experiments with animals or humans as subjects? Which organization(s) (in the Netherlands) can give you permission to do these kinds of experiments? What are the criteria that are used in considering your application? Regarding these criteria, what are the differences between human and animal subjects?
2. Present the Introduction and Method sections.
3. Present the Results and Discussion sections.

Week 5

Assigment for every student:

Read the article of the week.

1. What is/are the article's objective(s)?
2. What is/are the article's motive(s)?
3. What is/are the article's main conclusion(s)?
4. What is/are the article's implication(s)?
5. With which supports from the Results and Discussion sections is/are the main conclusion(s) justified? Try to depict the most important supports in at least two chains (an example of a chain is given in the information sheets).
6. Write down the counterargument(s) that the authors mention in the article. Write next to each counterargument – if mentioned – the corresponding refutation.
7. What were in this article the experimental groups?
8. The Method section consists of the following sections:
    - Estimation of vascular $\cdot O_2^-$ production
    - Examination of sources of $\cdot O_2^-$ in vascular homogenates
    - Isolated vascular ring experiments
    - Liposomal-encapsulated superoxide dismutase

   Summarize in no more than three sentences the essence of the experimental procedures for each of these sections. Furthermore, try to describe in no more than three sentences how each of the different sections relates to the objective(s) of the authors.
9. Three different experimental procedures may be distinguished:
    - Standard experimental procedures that are used by everybody in the field.
    - Experimental procedures that are developed/improved in previous studies.
    - An experimental procedure that was developed by the authors themselves.

   Identify in the article for each category an example.

Presentation assignment:

1. Give a presentation about methodologies. Discuss the following questions: What is the difference between correlative and causative research? (Give examples for each type of research.) What is a variable? What are the differences between independent, dependent, and controlled variables? (Give examples for each type of variable.) Which types of variables were examined in the article of the week? What are

between-group, within-group, and cross-over experiments? What is a "control system"?

Week 6

Assigment for every student:

Read the article of the week.

1. Depict the argumentative structure of the article. Put all the argumentative elements in a scheme as explained during the most recent tutor group meeting. Use the authors' original phrases – if necessary use (…) to shorten long quotes. Put behind each fragment a letter that indicates the type of element and the page number: (M) motive, (O) objective, (MC) main conclusion, (I) implication, (S) support, (C) counterargument, and (R) refutation.
2. Try to find at least four implicit supports in the article. You may put these in a separate document – you do not have to put them in your scheme. For each implicit support, write down which text fragment is justified with it.

Presentation assignment:

1. Present the conclusion of the article of the week. Discuss the relation (according to the authors of the article) between O_2^-, NO, eNOS, iNOS, chronic heart failure, endothelial dysfunction, sGC, cGMP, and vasodilation. Discuss also the function of NADH, SNP, DETC, and SOD.

Week 7

Assigment for every student:

Read the article of the week.

1. Identify in the article the following elements and write them down: motive, objective, main conclusion, implication, important supports, counterarguments, and refutations. Please include for each element the page number.
2. Why did the authors use a) L-NMMA and b) SNP?
3. Why did the authors also study the "radial artery blood flow"?
4. Does the main conclusion follow from the presented data?
5. Do the authors mention all counterarguments? If not, what do you miss?
6. Are the refutations of sufficient quality? Why?

7. Does this study warrant the administration of vitamin C pills to patients with chronic heart failure? Why?

Presentation assignment:

1. Present your answers to questions four till seven.

Week 8

Assigment for every student:

Read the article of the week.

1. Identify in the article the following elements and write them down: motive, objective, main conclusion, implication, important supports, counterarguments, and refutations. Please include for each element the page number.
2. Why did the authors measure 8-epiprostaglandin $F_{2\alpha}$-levels ?
3. Does the main conclusion follow from the presented data?
4. Do the authors mention all counterarguments? If not, what do you miss?
5. Are the refutations of sufficient quality? Why?
6. Does this study warrant the administration of vitamin C pills to patients with chronic heart failure? Why?
7. How do you explain that the current article does not show an effect of vitamin C while the previous article did show an effect?
8. Looking at the data of the current and previous article: has vitamin C an effect on oxidative stress or not? Please explain your answer.

Presentation assignment:

1. Discuss your own preparation for the oral examination. What do you take with you? How are going to read the research article? How are you going to present your abstract? What will be your focus? What is the secret of a good quality abstract? Give your fellow students advice that may help them with passing the oral examination.

Dankwoord

In 1923 werd aan de Britse bergbeklimmer George Mallory gevraagd waarom hij de Mount Everest wilde beklimmen. Vanwege het geld? De roem? Uit vaderlandsliefde? "*Because it's there*" ("Omdat hij er staat"), antwoordde hij simpelweg. Ik heb altijd op dezelfde manier tegen mijn promotie-onderzoek aangekeken. Als iemand mij zou vragen waarom ik zes jaar geleden begonnen ben met het beklimmen van deze intellectuele berg, zou ik precies hetzelfde antwoorden als Mallory: "Omdat hij er staat." (Begrijp me niet verkeerd: ik wil promoveren niet vergelijken met het beklimmen van de hoogste berg ter wereld – promoveren is natuurlijk veel zwaarder.)

Uiteindelijk zou Mallory in 1924 een poging doen om de top van de Everest te bereiken. Hij en zijn klimpartner Andrew Irvine kwamen daarbij jammerlijk om het leven. (Het is tot op heden onduidelijk of Mallory tijdens deze fatale expeditie boven op het dak van de wereld heeft gestaan.) Ik heb, gelukkig, het er beter vanaf gebracht. Dat heb ik vooral te danken aan mijn collega's, familieleden en vrienden. Zij waren mijn onvermoeibare sherpa's: zonder hun steun zou het me nooit gelukt zijn om de soms ijle lucht en ijzige kou te trotseren.

Allereerst ben ik dank verschuldigd aan mijn twee begeleiders: Martin Goedhart en Miriam Ossevoort. Ik heb het altijd zeer gewaardeerd dat ze me de ruimte en het vertrouwen hebben gegeven om mijn eigen invulling te geven aan het onderzoek. Ook heb ik veel gehad aan hun commentaar op mijn teksten. Ik vond het vooral prijzenswaardig dat ze altijd benadrukten – en hier parafraseer ik Schopenhauer – dat je in de wetenschap gewone woorden moet gebruiken voor het zeggen van ongewone dingen (iets wat geen vanzelfsprekendheid is in ons vakgebied).

Hendrik Buikema en Rob Henning van het vak Biomedisch onderzoek hebben vanaf het begin op enthousiaste wijze met ons samengewerkt. Biomedisch onderzoek vormt het fundament van dit proefschrift; hun inzet was daarom cruciaal. De hoofdtutor Eline van Dijk, de overige tutoren en de studenten hebben eveneens een essentiële bijdrage geleverd aan dit onderzoek. Dankzij hun voorbeeldige medewerking verliep het verzamelen van data uitzonderlijk soepel.

Drie collega's hebben een belangrijke inhoudelijke bijdrage geleverd aan dit proefschrift: Maartje Kevenaar, Henk Pol en Marcel Koeneman. Maartje heeft waardevol voorwerk gedaan waarop ik verder kon bouwen toen ik begon met mijn onderzoek. Henk was in de eerste jaren een onofficiële mentor voor me. De motiverende gesprekken met hem hebben we enorm geholpen bij het richting geven aan dit onderzoek. Samen met Marcel heb ik onder andere het SAM-model ontwikkeld. Ik heb met bijzonder veel plezier met hem geknutseld en gefröbeld aan SAM, niet in de laatste plaats door zijn eruditie en scherpe intellect.

Ik ben mijn collega's van de vierde en vijfde verdieping eveneens dankbaar. (Ik zal ze niet allemaal bij naam noemen. Niet uit gebrek aan waardering, maar meer omdat ik veel te bang ben om iemand te vergeten!) Dankzij hun hartelijkheid en humor fietste en wandelde ik elke dag welgemoed naar Nijenborgh 4/de Bernoulliborg. Ik koester dierbare herinneringen aan de uitjes, koffiepauzes, etentjes, ganggesprekken en natuurlijk de lunches. Deze waren vaak hilarisch, soms verheffend, maar altijd de moeite waard.

Mijn nieuwe collega's in Amsterdam hebben me op een geweldige wijze ondersteund tijdens het afronden van mijn onderzoek. Ik weet niet of het anatomisch en taalkundig mogelijk is, maar voor mijn gevoel hebben ze me een ontelbare hoeveelheid harten onder de riem gestoken.

Elma Dijkstra en Karin van Lacum zijn mijn paranimfen. Ik voel me vereerd dat deze twee fantastische vrouwen – in meerdere opzichten mijn kariatiden – me vandaag willen seconderen.

Mijn vrienden heb ik de laatste jaren op soms criminele wijze verwaarloosd. Desalniettemin kon ik altijd op ze rekenen. Zoals Walt Whitman ooit zei: "Zonder twijfel heb ik mijn vijanden verdiend, maar ik geloof niet dat ik mijn vrienden verdiend heb."

Ook een groot woord van dank aan mijn familieleden en in het bijzonder mijn lieve ouders. Aan hen heb ik dit proefschrift opgedragen. Het is het minste wat ik voor ze kan doen.

En nu is het tijd om nieuwe bergen te gaan beklimmen. Bergen die wellicht minder hoog en stijl zijn dan deze. Maar voordat ik dat ga doen, wil ik eerst nog een tijdje genieten van het majestueuze uitzicht. Ik kom zo naar beneden...

Curriculum vitae

Edwin van Lacum werd op 19 augustus 1981 geboren in Groningen. In 2000 behaalde hij zijn VWO-diploma aan De Waezenburg in Leek. Hij ging vervolgens de bachelor Medische Biologie volgen aan de Rijksuniversiteit Groningen (RUG). In 2004 begon hij aan diezelfde universiteit met de master Wetenschapseducatie en -communicatie. Na zijn afstuderen in 2006 begon hij te werken als freelance tekstschrijver. Vanaf 2007 combineerde hij dit werk met zijn promotie-onderzoek aan het Instituut voor Didactiek en Onderwijsontwikkeling van de RUG. Op dit moment is hij verbonden aan de Universiteit van Amsterdam als docent Academische Basisvaardigheden.

Index

Over de gebruikte lettertypes

Het binnenwerk van dit proefschrift (op de tekst in de figuren na) is gezet uit de Sabon. Dit lettertype werd in de jaren zestig ontworpen door de Duitse typograaf Jan Tschichold (1902–1974) en staat bekend als één van de meest leesbare lettertypes ter wereld. Tschichold was in jaren veertig ook vormgever van de Penguin-paperbacks. Hij gebruikte daarvoor onder andere Gill Sans. De tekst in de figuren van dit proefschrift is gezet uit dit lettertype. Gill Sans werd in jaren twintig ontworpen door de Britse typograaf en beeldhouwer Eric Gill (1882–1940).

university of groningen

faculty of mathematics and natural sciences

It is beneficial for science students in higher education if they master the ability to read research articles – an important aspect of scientific literacy – at an early stage of their academic training. However, reading research articles is a demanding task for novice readers. Therefore, we have developed a teaching strategy that can be used to teach undergraduate life science students how to read research articles. A central part of the strategy was a heuristic we designed: the Scientific Argumentation Model. This heuristic may be used to read, analyze, and critically evaluate research articles. Our teaching strategy was implemented in a course module. We tested the efficacy of our strategy by analyzing group discussions and students' written answers to several assignments. The results described in this thesis show that our strategy is useful for teaching students how to read research articles.

ISBN 978-90-367-6249-6