NADJA BELOVA, INGO EILKS and TIMO FEIERABEND

THE EVALUATION OF ROLE-PLAYING IN THE CONTEXT OF TEACHING CLIMATE CHANGE

Received: 7 February 2013; Accepted: 31 May 2013

ABSTRACT. Role-plays are a common pedagogical tool in the Social Sciences. As an imitation of societal practices, role-plays are thought to support the development of argumentation and decision-making skills among learners. However, argumentation and decision making are also goals in science education in general and in socioscientific issues-oriented science teaching in particular. This paper discusses a grounded theory (GT) approach to evaluating students' performance within role-playing exercises. The context is climate change. Data come from 4 different role-playing scenarios covering climate change which were developed in parallel for Biology, Chemistry, Physics, and Politics education. Role-plays in 20 different learning groups were videotaped (5 per subject). An evaluation pattern was developed step by step according to GT. Finally, graphic representations of all the role-plays were derived. The representations enable a quick overview of the role-plays and allow the identification of four basic types of role-playing: role-plays that are (1) completely directed by a group of student moderators, or (2) by the teacher, (3) medium-quality role-plays with a certain amount of interactivity and free argumentation, and (4) real, spirited debates. Implications for the use of role-playing exercises in science education are derived, including the induction of such role-plays through the use of role cards and the influence exhibited by teacher behavior.

KEY WORDS: argumentation, climate change education, decision making, role-play, socioscientific issues

Introduction

In today's society, we are continually faced with problems and dilemmas requiring us to make decisions, many of which specifically deal with questions concerning the use of science and technology (Dawson & Venville, 2010). Therefore, argumentation and decision-making skills have become essential goals in science education due to their importance for science- and technology-related choices and for contributing to societal debates (Bell & Lederman, 2003; Erduran & Jiménez-Aleixandre, 2007). Although the entire framework still requires research and curriculum development (Duschl & Osborne, 2002), school curricula focusing on argumentation and decision making are widely acknowledged as an important part of science education (Erduran, Simon & Osborne, 2004; Sjöström, 2011). Skills in both domains are essential for allowing

International Journal of Science and Mathematics Education (2015) 13(Suppl 1): S165–S190

pupils to contribute to societal debates on socioscientific questions as responsible citizens (Holbrook & Rannikmäe, 2007; Roth & Lee, 2004).

The most effective topics for learning about societally oriented argumentation and decision-making processes are those which force learners to take a firm stand with respect to a specific societal controversy (Sadler, 2004). Eilks, Nielsen & Hofstein (2014) suggest that classroom topics with the greatest potential for promoting argumentation, evaluation, and decision-making skills must be relevant, authentic, controversial, and openly debatable in the classroom. Science education research and development suggest a more thorough implementation of such relevant, authentic, and controversial socioscientific issues in the science classroom (Sadler, 2004).

One such controversial topic is the issue of climate change (Feierabend & Eilks, 2010). Klafki (2000) states that climate change can be characterized as one of the most prominent key modern problems within today's society. Society's challenge by climate change requires judgment, decision-making, and argumentation skills from every citizen. Duschl & Osborne (2002) even view climate change as a prototype field when it comes to learning about societally oriented argumentation and decision making. The controversial character of climate change makes the ensuing science teaching practices challenging, motivating, and potentially useful for promoting general educational skills among students, i.e., in the field of argumentation and decision making (Sadler, 2004; Marks & Eilks, 2009). Curriculum reform and implementation in this direction is needed since the societal dimension of science education remains a neglected component of science teaching in many countries (Hofstein, Eilks & Bybee, 2011).

The background of this study is a curriculum development and implementation project called "The Climate Change before the Court" (in German "Der Klimawandel vor Gericht"; Feierabend & Eilks, 2011). In this project, climate change lesson plans for high school classes were cyclically developed in parallel for Biology, Chemistry, Physics, and Politics education. Development and cyclical optimizing of the lesson plans were performed using a partnership of university educators and in-service teachers. The approach follows the participatory action research model of science education as described by Eilks & Ralle (2002; see also e.g. Mamlok-Naaman & Eilks, 2012). Even though role-playing and open discussions are often used in school education in the Social Science school subjects, their application in science education remains spotty at best (Patronis, Potari & Spiliotopoulou, 1999; McSharry & Jones, 2000). That was the reason that within the curriculum innovation project "The

Climate Change before the Court," the decision was made to incorporate role-playing in all the different lesson plans. In each of the four subjects, different scenarios for these role-play exercises were developed, and the lesson plans were designed to be carried out in different grade levels ranging from grades 9 to 11 (age range 15 – 17). Due to the cyclical approach of the curriculum development by participatory action research, the lessons were repeatedly tested in different middle (visited by lower to medium achieving students), comprehensive (visited by integrated groups of students from all achievement levels), and grammar schools (visited by higher achieving students). Within the main phase of development, a huge amount of data was collected. Among the data, there were the role-playing exercises, which were videotaped in 20 different learning groups (five per subject).

This paper describes a systematic analysis of the students' performance within the role-plays. An evaluation grid for role-playing in general and climate change in particular was developed using grounded theory (GT) (Strauss & Corbin, 1990). The evaluation aims answering the research questions addressed by this paper:

- Which level of argumentation skills and societally oriented decision-making capabilities are evident in student role-playing exercises on climate change around the end of lower secondary science education (age 15 – 17)?
- Can different types of role-playing be identified with respect to their overall quality in promoting effective debates and consensus building processes?

Theoretical Framework

Role-playing (Hollingworth & Hoover, 1991) is used in a broad spectrum of teaching and training practices, ranging from school classrooms to staff and manager training programs at the corporate level (van Ments, 1999; Saunders, Percival & Vartiainen, 1996). Various teaching guides on student-active methods have already pointed out numerous advantages of such an approach (Eilks, Prins & Lazarowitz, 2013). Role-play acting can foster the imagination by allowing students to participate in new, unfamiliar situations. It can also boost empathy with and the tolerance of other people by forcing pupils to learn about and understand other peoples' opinions, including a look at one's own base assumptions and prejudices. Role-playing has the potential to activate all of the students in

a learning group and aid them in bettering their interpersonal skills (van Ments, 1999; Wilhelm, 2002; Killen, 2009). Generally, a pedagogy emphasizing the discussion of controversial issues avoids the automatic acceptance of outside views by the students and directly promotes more critical reflection on ideas encountered outside of the classroom (Oulton, Dillon & Grace, 2004). In school, role-playing still occurs most commonly in the humanities, i.e. History, Literature, Politics, and Sociology (van Ments, 1999).

There is currently no evidence for the frequency of the application of role-playing exercises in science classrooms. However, the rare occurrence of related articles in the science education literature (research and teacher journals) would seem to indicate that their use remains quite limited. Taking into account that societally oriented science education as such is insufficiently implemented worldwide (Hofstein et al., 2011), we can assume this to be even truer for the pedagogy of societally oriented role-playing exercises. Nevertheless, there have been some contributions in the literature which deal with the use of role-playing in the science classroom. Unfortunately, most suggested uses are found in practical papers and contain few or even no documented examples of research data. For example, McSharry & Jones (2000) suggest several ways to incorporate role-playing into science lessons. Their main emphasis stresses the playful, active components of the method, while the debating aspects play only a minor role. For example, they suggest that students "build" the electron structure of an element by organizing themselves in circles or that they demonstrate the physical states of matter by running around or standing in lines. Ødegaard (2003) offers a similar approach with ideas for "dramatizing" scientific concepts like electricity. She describes an approach with role cards and a structured frame appropriate for optimizing students' creativity, critical thinking, and the materialization of their understanding. McSharry & Jones (2000) or Ødegaard (2003) merely provides impulses for a more student-centered approach to science with the help of role-play without further research. In the case of Germany, the country this study stems from, the picture is similar. There have been many suggestions for the use of role-playing in the last 40 years, both for learning about purely scientific concepts and also the societal dimension of science. However, research into the underlying processes and resulting effects remain mainly at the level of single experience reports.

Only a few papers and articles offer research-based insight into potential methodologies for analyzing processes occurring during role-playing exercises and their effects. Duveen & Solomon (1994) used an

exercise based on Darwin's fundamental book "The Origin of Species." The setting is a court trial where Darwin is charged with blasphemy. The focus of the unit lies on the history of science and historical empathy. The main finding from the research is that students remember the main theses from the exercise strongly even months after the implementation. Simmoneaux (2001) presented probably the most thorough study of students' argumentation and decision-making processes in role-playing based on the topic of animal transgenesis. Her emphasis was on the comparison of role-playing and classic classroom debates and the domains of arguments used by the students. One of the results was that a greater change in opinion took place after role-playing than after a regular debate. Simmoneaux also defined a number of fields of references from which students' arguments emerged, such as economics, ecology, genetics, medicine, and ethics. She determined a difference in those fields for role-playing and debates. Simmoneaux also pointed out that there are currently no well-defined criteria for evaluating the quality of students' argumentation in such exercises. Marks, Bertram & Eilks (2008) and Marks & Eilks (2010) have also described applications of role-playing in a socioscientific issues-based curriculum. Their studies included teaching units on products connected to low-fat and low-carb diets and on the use of musk fragrances in shower gels. They evaluated the motivational character and the feasibility of the lesson plans, but less evidence was given of the students' skills in argumentation and decision making.

Apart from role-playing, there exists a large amount of research on students' skills in argumentation and decision making. Although most of this research is not geared toward the evaluation of role-playing exercises, it might provide some important guidance for the current study. Patronis et al. (1999) distinguished between students' arguments with respect to their role in the course of discussion. In their analysis of interviews, classroom situations and discussions on a socioscientific issue, Patronis et al. made distinctions between how arguments are referred back to earlier statements. They divided their data into arguments of defense or attack, dependent on a previous statement. A second suggestion regards differentiating between qualitative, semi-quantitative, and quantitative arguments. Quantitative arguments refer to numbers and calculations and thus are often connected to scientific evidence. When students use variable quantities or put numbers in relation to something, an argument is considered to be semi-quantitative. They found that the most often used arguments were the qualitative ones referring to single qualitative facts or opinions. They also found that it is easier for students to name arguments when they know about familiar situations, e.g., those stemming from their

personal lifeworld. This finding might explain the dominance of arguments which are less quantitative, less complex, and often opinion-based. This finding is in line with Fleming (1986), who found that most students tend to prefer arguments stemming from their social world when arguing about socioscientific issues. These arguments are more qualitative in nature, less complex, opinion-based, and lack connections to quantities and calculations. Simmoneaux (2002) came to a similar conclusion. Aside from the question of familiarity, a second reason for such an explanation might be found in Zeidler, Walker, Achett & Simmons (2002). This study found that students have problems understanding and dealing with terms like "theory," "scientific knowledge," "scientific data," or "opinion" and therefore they struggle with using and combining them to form more complex and better-supported arguments for socioscientific reasoning. This missing understanding about the role of evidence in science is furthermore explained by the fact that many students think that personal opinion is immune to change, even if another point of view is more convincingly supported by better evidence.

Tytler, Duggan & Gott (2001) have also researched the source of arguments used by students in socioscientific discussions. They identified three types of evidence which were utilized: informal evidence, evidence taken from the wider framework of the socioscientific issue, and scientific evidence. They described the use of scientific evidence as quite rare among most students when compared to the other two domains. This is in line with an older study by Solomon (1992), who asserts that students only rarely use specific knowledge from the science domain. Kolstø (2006) also differentiated between a more personal source of arguments and a more scientific one. He stated that there are always two sources for the emergence of any argument: a personal, ethical, societal side on the one hand, and science itself on the other. Kolstø found that some students use scientific knowledge in their decision making, even if the complex theories learned in school are rarely connected with the arguments. However, these findings might also be tied to the type of learner. Yang & Anderson (2003) identified three different types of students with reference to their preference of argumentation sources: scientifically oriented, socially oriented, and equally disposed. It seems that most students belong to the latter two types. Kolstø (2006) also identified five specific types of arguments: the relative risk argument, the precautionary argument, the decision impossible argument, the small risk argument, and the pros and cons argument. Mitchell (1996) also described different sorts of arguments—regular and critical. The latter type challenges existing theories and is more complex, but is also much more rarely applied.

When it comes to consolidating students' argumentation and decision making within appropriate models, several approaches have been suggested. Sadler & Zeidler (2005) characterized students' decision-making skills in discussions about genetic engineering dilemmas using the help of three modes: rationalistic, emotive, and intuitive–informal reasoning. Bell & Lederman (2003) derived their view of decision-making skills from the nature of science perspective. They concluded from an experts' survey that it is necessary to re-examine the goals of any nature of science instruction and to add more value-based instruction, including paying attention to intellectual/moral development. These components were suggested as necessary for learning about decision making in science education, which is connected to the real needs of future citizens. This is so, because social/political issues, ethical considerations, and personal values are also dominant facets of experts' own decisions on socioscientific issues, although the decision makers themselves stem from the science field.

Based on the literature discussed so far, it becomes clear that the field of analyzing argumentation is very complex. It is definitely more than simply counting arguments. From the discussion above, we see characterizing argumentation at least demands the analysis of chains of arguments, e.g., by reflecting the complexity of single arguments but also their interrelatedness to one another. However, it also needs to take the content into respect by analyzing the source the information stems from and the role it is suggested to take within debate, as has been suggested, e.g., by the works of Toulmin (Simon, 2008; Nielsen, 2011, 2012). This paper therefore tries to use a very open approach when analyzing the role-playing data gained from the project "The Climate Change before the Court." A grounded theory approach was chosen to reveal which of the above-mentioned ideas best fits the analysis of role-playing exercises based on climate change.

Data Background and Sample

The project "The Climate Change before the Court" (Eilks et al., 2011a) developed lesson plans for Biology, Chemistry, Physics, and Politics education using the participatory action research model of science education as described by Eilks & Ralle (2002). The four different lesson plans focused on domain-specific approaches to climate change. Each lesson plan lasted between 10 and 12 h. A common focus of all the lesson plans was the advancement of skills in argumentation, evaluation, and

decision making within the controversial socioscientific issue of climate change. All lesson plans had one thing in common: Each ended with a role-playing exercise in the last two to three classroom periods. In this project, an interpretation of role-playing was selected understanding the role paying as a cooperative exercise where groups of students make themselves familiar with the knowledge, views, and interests of groups from the individual living environment or the society at large to later act and represent the respective stakeholder group in a fictive debate in the classroom. The scenarios for the different subjects can be found in Table 1.

During the main phase of testing the lesson plans, the units were implemented in 20 different learning groups (five in each of the four subjects). Lessons took place in grades 9 – 11 learning groups (age range 15 – 17) coming from different middle, comprehensive, and grammar schools. From the information provided by the teachers, it is suggested that experience in role-playing was very limited among all the learning groups, especially in the science subjects.

The role-playing followed different scenarios, but were also organized with slight variation in mind. One difference was the source of the roles. In Biology, the students represented the roles of individual persons borrowed from their everyday life environments, such as portraying a local farmer or the operator of the school canteen (although each was represented by a group of students). In Chemistry, Physics, and Politics,

TABLE 1

Scenarios for the role-plays

Subject	*Problem*	*Scenario*
Biology	Should the new school canteen offer meals containing meat?	A conference of the school board (all the stakeholders in school like teachers, students, and parents) has to make a decision on that question.
Chemistry	Should the minimum age for a driving license be raised up to 21 to avoid traffic?	The youth committee of a parliament has to make a recommendation.
Physics	Should the transport or import of fruit by plane be controlled by a new EU directive?	A panel of the EU commission has to prepare an EU guideline.
Politics	Should the EU impose a ban of imports for bio ethanol from Brazil?	An ethical review committee has to make a recommendation.

roles were assigned according to interest groups and their agendas, such as the automobile industry or Greenpeace activists. The exercises also differed somewhat in their structures. Some of the units had a more rigid pattern. In Politics and Chemistry, the group began by summarizing the statements made by representatives of different interest groups, in order to clarify their positions. In these units, a committee (consisting of one of the role-playing groups) was assigned to make a final decision at the end of the exercise, while in the other subjects the open discussion started straight away and ended with a round of closing statements. In all the role-plays, one group of students was assigned to be the moderators. In chemistry and politics, the deciding committee also took over the role of moderators. These groups of moderators were responsible for leading the discussion by admitting the floor to someone or by calling their classmates to order if individuals or groups were not keeping respect to others sufficiently. One aspect all the exercises had in common was the assignment of individual roles through so-called role cards. The students received crib sheets to help them prepare for their roles. These cards could also be used during the active part of the exercises. However, the structure and length of the texts differed significantly in the four subjects. In Chemistry, only two to three sentences and a few links to Internet sources were provided. In Biology and Physics, a text of about one page in length was given, whereas in Politics the students received three to five pages of pre-structured information (Eilks et al., 2011b). An example of a short role card from the chemistry lessons is presented in Fig. 1.

While carrying out the exercises, structural aspects also had an effect on the length of the role-playing. Some of the very structured units were shorter and lasted only 20 – 30 min; more open ones were often longer and lasted up to 90 min. Altogether, 20 role-playing exercises were conducted within this study with a total length of nearly 14 h for the data collected. The average role-playing length was 40 min; four of the exercises lasted longer than 60 min. Each exercise was videotaped in an authentic classroom situation in the last periods of the lesson plan. To make sure that the role-plays were adequately captured, two cameras with extra microphones were placed on two opposite edges of the classroom. The role-plays were later transcribed.

Method

This study applies the GT approach according to Strauss & Corbin (1990) in order to evaluate the role-playing exercises. GT is an open approach of

Expert Group: Green Mobility

As **experts for Green Mobility** you stand for alternative means of transportation like electric cars, alternative fuels such as biodiesel, and public transport with buses and trains. **Your goal is the reduction of greenhouse gases and climate protection** with the help of alternative means of transportation and the limitation of individual transport. Individual transport means that everyone uses his own car or motorcycle. Therefore **a law** on the increase of the minimum age for a driving license up to 21 years is **a good idea** to achieve your goals, as it forces young people to use public transport and limits the overall traffic.

Figure 1. Example of a role card from the chemistry role-play. The role cards in the chemistry lesson plan were all that short and provided further information by giving the student a collection of selected links to the Internet. In all the other subjects, the role cards were much longer. They were made up of up to three pages of text and graphical information, some of them providing the students a full course of arguments and chain of potential argumentation

data handling and analysis with no pre-formulated hypotheses. The goal of GT according to Strauss and Corbin is to discover as many data codes as possible at the beginning and then formulate categories out of them (open coding), before relating these categories to one another (axial coding) and finally creating frame core categories (selective coding). The core categories are assumed to explain the essential characteristics of the data.

In our case, text transcripts of all 20 videos from the different role-playing exercises were used. As the first step, transcripts from five of the exercises were open coded. This step resulted in about 300 different codes that were cyclically refined through inductive and deductive grouping. This step resulted in 40 open-coding categories that were formulated. By axial and selective coding, three main categories emerged: (I) domain, (II) level of argument, and (III) reference:

I. Domain: Where do the arguments used by the students come from (science/everyday life/society/politics)?
II. Level of argument: How complex are the arguments?
III. Reference: Do the students make references to each others' statements? Does a conversation arise?

For quality control, this step of analysis was carried out by two independent coders. The agreement rate was high at more than 80 % agreement. The few cases of disagreement were negotiated by searching for inter-subjective agreement (Swanborn, 1996). This led to refinements

in the code descriptions. By applying the resulting three categories to the other 15 role-playing units, a high level of data saturation was achieved.

Based on the first five exercises, each of the core categories was then expanded into dimensions representing the different aspects or levels within them. The first category, "Domain", suggests that the students apply arguments from different sources such as everyday life, science, etc. This category was denoted a "D" followed by a number according to the different sources. Three sub-categories arose from the data: D0 miscellaneous with unclear origin, D1 everyday life/society/politics, and D2 science (Table 2) similar to the differentiation suggested by Kolstø (2006).

The second category, "Level", appeared to be the most complex. The gradation in this category suggests a spectrum of understanding the level of an argument, which reaches from simply repeating a particular claim all the way up to constructing complex arguments composed of different claims with accompanying justifications and/or reflective elements. This category was expanded into a six-step dimension paralleling the pattern suggested by Feierabend, Stuckey, Nienaber & Eilks (2012) for evaluating complexity of arguments within semi-structured group discussions on climate change. Table 3 explains and illustrates the different sub-categories (L0 to L5) that were found. To make the different levels as clear as possible, several exemplary arguments are provided for each sub-category.

The third category, "Reference", refers to the course the conversations took. Reference shows whether pupils' arguments refer back to comments made by a prior speaker or earlier statements which had been made (see also Patronis et al., 1999). This category also acknowledges different types of references which appear during role-playing. Sometimes students

TABLE 2

Category "Domain"

	Description	*Example*
D0	Miscellaneous with unclear origin	"Meat is not really important in life."
D1	Origin: everyday life/society/politics	"A lot of young people live far away from school."
D2	Origin: subject related (science)	"18 % of the greenhouse gas comes from cattle."

TABLE 3

Category "Level"

Symbol	*Level*	*Description*	*Examples*
L0	Not related	Students provide arguments which do not have a connection to the topic.	"Food turns into food for my wonderful pigs." "You can't work without a car, but you can go to university instead."
L1	1 argument	Students provide 1 relevant argument but do not provide any justification for it.	"The African people don't benefit from it; it's only good for Europe and America." "I can't imagine that being vegetarian can be healthy." "We have to get rid of the pigs and breed chickens instead."
L2	2 arguments	Students provide 2 or more relevant arguments without logical relation or sound justification.	"Students are left to themselves. And they lack nutrients." "The consumers have to decide by themselves if they want to buy the product or not. But we have to get rid of the eco labels."
L3	1 or 2 arguments, 1 justification	Students provide 1 relevant argument with well-founded justification. Or they give 2 or more arguments with at least 1 justification.	"Yes, but that's just one car and the parents drive anyway, they need to go to work anyway." "Due to the economic crisis and stuff it is very important that people buy as many cars as possible." "My son must be strong because he plays soccer. A balanced nutrition for him requires meat."
L4	2 and more connected arguments with justification	Students provide 2 or more relevant arguments connected in a logical chain, justified by facts and/or personal experience.	"We must also think of the jobs, because we have a lot of revenue through trade. You have to make people realize that the price rises because it is harmful to the environment."

			"You cannot carry all kinds of fresh fruit to Europe by car or by ship because some fruit would rot. But fresh food products should be much more expensive if they were imported by plane, people should pay for the transport!"
L5	1 and more connected arguments with justification and reflection	Students provide 1 or more relevant arguments, provide justifications for them, and draw sound conclusions from their argument's interconnectedness.	"If fewer cars are purchased, the economy has less money. But people between 18 and 21 have not much income and cannot afford the best sports cars. So the income of the economy is not particularly high. Consequently, 21 year olds have more money and can afford better cars straight away. So your argument is invalid." "If you want to grow your own fruit without depending on us you have to send us donations instead because our country will be ruined without the fruit industry. This industry gives us jobs and people have a good living, without that the whole continent will be in a miserable position soon."

referred to each other in a purely communicative fashion, for example using indicator sentences like "As the Greenpeace group has already mentioned …" as a reference point. In other cases, a reference referred to the content, but had no preceding communicative phrase. Four different sub-categories were found, ranging from R0 (no reference) to R3 (a communicative and contextual reference was made). Additional signs

were added for references made to earlier statements, starting a completely new line of argument, or for breaks in the discussion (Table 4).

All of the material was then rated according to the three categories. The decision was made to present each coded unit as a rectangle divided into three parts, one for each category (Fig. 1). The example in Fig. 2 shows a statement which refers to another one in a contextual and communicative way (R3), one which originated in everyday life/society/politics (D1), and one that is relevant but without justification (L1).

Presenting all of the role-playing units as a sequence of rectangles provided a very condensed course of the exercises as a whole. Connecting the different rectangles with arrows showing the statements' references and including notions for statements made by the moderator(s) or the teachers resulted in flowcharts emerging. Other symbols, for example for the use of role cards or for scientifically incorrect arguments, were used, too.

As an illustration, Fig. 3 shows an excerpt of such a flowchart. The "corollary arrow" shows that a statement directly results from another one; thus, it is an immediate reaction. Normal arrows illustrate references back to earlier statements. The first element on the right-hand side shows

TABLE 4

Category "Reference"

	Description
R0	No reference
R1	Only communicative reference (i.e. "Yes, I agree")
R2	Only contextual reference
R3	Contextual and communicative reference
	Discussion is directed to a new, previously mentioned topic
	Discussion is directed to a new, previously not mentioned topic
	Discussion stops, moderator has to boost

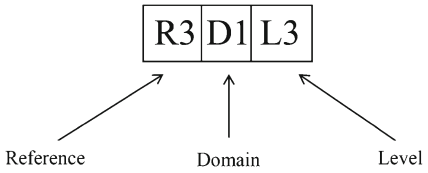

Figure 2. Basic module of a flowchart

that a new topic was introduced, whereas the two parallel lines indicate that the discussion stopped and the moderator had to intervene.

The flowcharts are all between one and three pages long. They allow for the analysis of single passages within the role-playing exercise as well as an overall consideration of the course and quality of the exercise as such in the sense that it was a functioning debate moving toward a consensus or final decision.

This approach proved useful for representing the data at good levels of quality. Application of the categories and dimensions to all of the 20 role-plays evidenced good data saturation. In order to analyze the argumentation skill level, the distributions of the different categories were placed together in tables and interpreted. To analyze the role-playing and see whether specific types are prevalent, including the overall level of quality, the connections within the flowcharts were analyzed in the sense of comparing them to smoothly flowing debates and consensus-building events.

Findings and Discussion

The diversity in the arguments provided by the students was very broad. Available arguments were borrowed from many different sources,

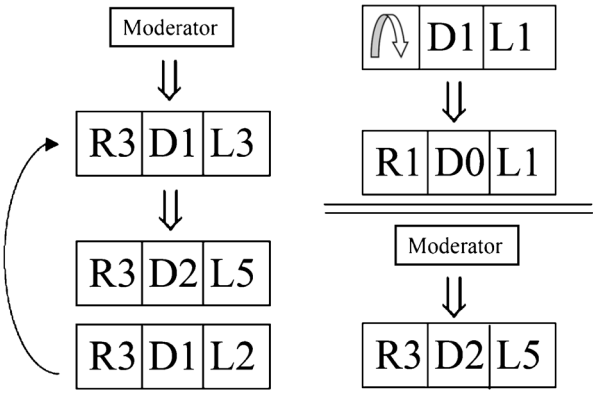

Figure 3. Excerpts from the flowcharts

including everyday life and science. Even though most of the role-playing took place in science classes, arguments stemming from everyday life and society prevailed considerably. This is in line with the findings of Fleming (1986), Solomon (1992), Tytler et al. (2001), and Kolstø (2006) in their studies of student discussions about socioscientific issues. In Biology, this sub-category represented almost 80 % of the arguments. This value was even higher than in Politics (72.4 %), where such arguments might have been expected to play quite a large role (Table 5). The percentage of arguments based on science was between 14.9 % (Biology) and 29.4 % (Chemistry), which means that only between one seventh and one third of the arguments, respectively, were related to science. These few scientific arguments also did not show a wide range of variety. Two domains in particular dominated the science-related arguments: the role of carbon dioxide in climate change and the anthropogenic greenhouse effect. Recurring keywords included rising sea levels, high carbon dioxide emissions due to air traffic and automobiles, renewable resources, and alternative fuels like biodiesel. Students use these science-related arguments, however, in doing so they use only those science-related arguments which are present in their everyday lives (Fleming, 1986). The greenhouse effect was mostly mentioned in the Chemistry exercises, possibly due to its central role in the lessons prior to the role-playing unit. Interestingly, this term was almost nonexistent in the Physics units, even though the topic was also emphasized in the Physics unit. The same lack was also evident in Biology. A central focus of the Biology lessons were carbon dioxide emissions for producing different foods. This might have resulted in solid arguments, but such factors were not mentioned at all during the role-plays. The same applied to topics like the energy content of different fuels or climate protection in everyday life. Many arguments were mentioned that related to the topics taught before

TABLE 5

Results for the category "Domain" for the four subjects

Domain	*Biology*	*Chemistry*	*Physics*	*Politics*	*Total*
D0	18 (7.7 %)	30 (16 %)	69 (12.9 %)	16 (6.9 %)	133 (11.2 %)
D1	182 (77.4 %)	102 (54.6 %)	375 (70.1 %)	168 (72.4 %)	827 (69.6 %)
D2	35 (14.9 %)	55 (29.4 %)	91 (17 %)	48 (20.7 %)	229 (19.2 %)
Total	235	187	535	232	1189

in the lesson plan; however, they were often insufficiently supported by theoretical explanations. Thus, for many arguments, it was difficult to say whether they were taken from the previous lessons, or whether the topics in the previous lessons just led the students to activate claims and arguments that were derived from other sources. A difference of course appeared in the chemistry group in comparison to the other subjects. Due to the fact that in chemistry the role-cards were very short (Fig. 1), the students were more concentrated on the previous lessons instead on just repeating the information provided on the role cards in the other subjects.

Quite often there was also inaccuracy in the use of science-related arguments by the students caused to the incorrect use of scientific language which frequently occurred. For example, the chemical terms "carbon dioxide" and "carbon" were often confused, as were the terms "resources," "fuels," or "greenhouse gases." These words were often used incorrectly, imprecisely, or superficially ("lots of resources go into the atmosphere"). Another frequent mistake was the confusion of the concepts weather and climate. There were many similar examples of the inaccurate usage of certain terms. An overall analysis of the science-related arguments showed that about 10 % of the total were used scientifically incorrect.

Regarding the complexity of the arguments employed, L0 was rarely to be found. L1 (one relevant argument without justification for it) was—except for Politics—the most frequent occurrence (Table 6). Patronis et al. (1999) also state that the most often-used arguments in students' socioscientific discussions tend to be those that are qualitative in nature and refer to single facts or opinions. L2 (several arguments with no

TABLE 6

Results for the category "Level" for the four subjects

Level	*Biology*	*Chemistry*	*Physics*	*Politics*	*Total*
L0	1 (0.4 %)	1 (0.5 %)	4 (0.8 %)	0	6 (0.5 %)
L1	03 (43.8 %)	90 (48.1 %)	266 (49.7 %)	75 (32.3 %)	534 (44.9 %)
L2	24 (10.2 %)	11 (5.9 %)	44 (8.2 %)	11 (4.7 %)	90 (7.6 %)
L3	65 (27.7 %)	53 (28.3 %)	127 (23.7 %)	33 (14.2 %)	278 (23.4 %)
L4	7 (3.0 %)	5 (2.7 %)	8 (1.5 %)	3 (1.3 %)	23 (1.9 %)
L5	35 (14.9 %)	27 (14.5 %)	86 (16.1 %)	110 (47.4 %)	258 (21.7 %)
Total	235	187	535	232	1189

justification) was hardly present in all subjects, reaching a maximum of about 10 % in Biology. L3 utterances (one argument with justification) constituted between 14.2 % (Politics) and 28.3 % (Chemistry) of all statements. L4 (a logical chain of arguments with justifications) makes up only 1 to 3 %, whereas the number of statements of the highest level L5 (justified and reflected) makes up about 15 % of the given arguments. In Politics education, this most complex sub-category made up about 50 % of the arguments. However, this is an artifact resulting from the structure of role-playing. This finding came up when analyzing the flowcharts of the Politics units. In Politics, the students were asked to provide an introductory statement, which was prepared in the previous lessons. These statements lasted up to 10 min each. The students in these learning groups also received the most extensive role cards of up to five pages of information. Therefore, they presented many pre-prepared, complex, and high-quality arguments during the opening of the exercise. However, in the later phases of the exercises, the presence of highly complex arguments no longer differed from that in the other learning groups in the other subjects.

Overall, the rote reproduction of arguments in sub-categories L1 and L2 made up slightly more than half of the arguments during the active role-playing phase. Arguments with the limited complexity represented by L3 composed about 30 % of the arguments. High-quality arguments with justification or reflection as represented by types L4 and L5 formed only about 15 % of the total. Thus, only every sixth or seventh argument belonged to one of these two types. This parallels the findings from Mitchell (1996), who described the rarity of critical arguments in students' socioscientific discussions when challenging existing claims. When comparing students from different types of schools, the quality of arguments was much better in the grammar school classes (schools with higher-achieving students) as compared to comprehensive and middle school classes. Students in grades 10 and 11 performed better than those from grade 9.

The category "Reference" is not so easily interpreted in terms of quality when compared the other two dimensions. The value can be either positive or negative during the course of role-playing, if a student breaks the flow of conversation by introducing a new topic into the discussion. Nevertheless, this category is essential for answering our second research question, which focuses on potential types of role-playing. The question of whether students refer back to prior speakers and utterances either in a communicative or textual way is also important. It allows us to consider role-playing as a process of exchange and shows positive indications that

the exercise is working as a debate. In contrast, coding such as "No reference" and "Discussion stops" demonstrate that the character of the discussion is far from a debate and/or sporadic in nature. In the findings, we see that code R3 dominates all of the subjects, especially Politics. On average, more than half of the references were of this type. Table 7 overall the amount of the codes R1 to R3 (direct communicative and/or contextual reference) amounts to more than 80 % of the material in all the subjects. Codes recording no references or a stop in conversation appeared only very rarely. We can conclude that the course of the discussion in most cases flowed smoothly and that the contributions made by different participants were networked by references to each other.

Analyzing the flowcharts allowed us to group the role-playing into different types. Categorization was performed using the flowcharts with respect to the liveliness of the discussions. This was interpreted as chains of arguments given by different speakers, the structuredness of debate in regard to the amount of references to previous statements, and the absence of breaks or interrupting events caused either by the teacher or the students. Another criterion defining functional role-playing consisted of analyzing the final phase to see whether the resulting, final conclusions

TABLE 7

Results for the category "Reference" for the four subjects

Reference	*Biology*	*Chemistry*	*Physics*	*Politics*	*Total*
	1 (0.5 %)	0	7 (1.3 %)	0	8 (0.7 %)
R0	2 (1 %)	1 (0.6 %)	2 (0.4 %)	0	5 (0.5 %)
	17 (8.8 %)	13 (7.7 %)	37 (7.1 %)	4 (2 %)	71 (6.5 %)
	18 (9.3 %)	9 (5.3 %)	52 (9.9 %)	19 (9.5 %)	98 (9 %)
R1	6 (3.1 %)	16 (9.5 %)	12 (2.3 %)	6 (3 %)	40 (3.7 %)
R2	52 (26.8 %)	18 (10.6 %)	95 (18.2 %)	17 (8.4 %)	182 (16.7 %)
R3	98 (50.5 %)	112 (66.3 %)	318 (60.8 %)	155 (77.1 %)	683 (62.8 %)
Total	194	169	523	201	1087

referred back to previous arguments, thereby yielding evidence of a search for compromise and consensus. Four major types were found:

Type 1 is characterized by the fact that the course of the discourse is completely directed by moderators from the learning group. In this sort of role-playing, the different actors present their point of view to a "council," then are subsequently asked follow-up questions. The different representatives in the roles rarely communicate with each other at any point. Open discussion hardly takes place. A conclusive decision is solely reached by the council; a search for compromise and consensus is not explicit in this type of exercise. This group consisted of four Chemistry and one Politics unit. Most of the Chemistry groups were grade 9 middle school classes (age range 14 – 15).

The quality of discussion in type 2 is very similar to type 1. However, here the teacher's dominance and strong guidance hinders open debate, the search for compromise, and consensus building. This type was found in two of the Biology learning groups.

Type 3 can be considered to be a role-play of medium level quality. Role-playing proceeds in a "civilized" and calm fashion. There are chains of argumentational exchange, but they do not lead over into controversial debates. However, an exchange of ideas takes place, and initial indicators are present that within the final statements earlier arguments are quoted and consensus building is approached. This was the case in one Physics, three Biology, and one Politics learning groups.

In type 4, real debates and spirited discussions took place. These exercises can be considered very successful examples. The chains of evidence are long and precise and moderators do not play a dominant role, but rather a reserved and objective one. The students really "live out" their roles and act convincingly. They try to give their role an important voice for the final consensus-making phase. Students often use their own arguments, refer them back to others, and come to reasonable compromises. This was the case in four Physics, one Chemistry, and one Politics learning group. All of these groups came from different grammar schools and were from grade 10 or 11 (age range 16 – 17).

During the flowchart analysis, teacher behavior also came more clearly into focus. Each teacher acted very differently. Some of them did not pay much attention to the role-playing and performed other work such as correcting tests. Other teachers acted as moderators during the exercises. Several of them even actively participated in the discussion, trying to bring in their own ideas or arguments, which seemed to them to be more

appropriate than the ones presented by their students. One of the teachers in a type 2 exercise constantly tried to provoke the students ("Are grapes from Chile more ecological than a bread roll with ham? I don't think so!"). Another teacher expected students to raise their hands whenever they wanted to make a statement and called them by their real names. All this clearly limited the motivation of the students. A free-wheeling discussion did not develop. Teacher behavior in these cases limited the opportunities for students to openly discuss and freely experiment with different arguments in a comfortable atmosphere (Nielsen, 2009).

Conclusions and Implications

This paper discussed a grounded theory (Strauss & Corbin, 1990) approach for evaluating role-playing exercises within a framework of socioscientific issues-based teaching. An evaluation pattern was derived from 20 different lessons for lower secondary pupils focusing climate change. The students came from different types of schools, and the role-playing exercises took place in four different school subjects: Biology, Chemistry, Physics, and Politics.

The evaluation grid consisted of three main categories (Level, Domain, and Reference) which were each operated within a dimension. The grid made the analysis differentiated and precise and provided a good data saturation level for analysis. Using the evaluation grid, a broad representation of entire role-playing units in the form of flowcharts becomes possible. These flowcharts allow for overall consideration of the exercises with respect to whether debate really takes place and whether searches for compromises and consensus actually occur. Four different types of role-playing performances were identified, which represent different levels of quality concerning performance.

Within the role-playing exercises, we found that students were able to create a forum for discussion. However, the level of argumentation and discussion was quite diverse. Only one third of the arguments were complex and reflective in nature. Only about one in three of the role-playing exercises was considered to be of high quality with respect to the character of debate and a focus on consensus building. Although this was not a quantitative and comparative study, we also found indications that better quality is more evident if the pupils come from grammar school compared to those from middle and comprehensive schools. Additionally, discussion quality increases with learner age. Overall, only about 20 % of the presented arguments were related directly to science, even though

most of the role-playing exercises were carried out in science classrooms. The level of science-related arguments was not very high; the reasoning tended to mirror societal debates in real life. If looking into everyday forums as TV talk shows, the amount of science-related arguments is also usually low when discussion circles around socioscientific issues such as climate change. Even when scientists take part in such discussions, many of their arguments stem from the personal and societal domain (Bell & Lederman, 2003). It is also interesting to note that there was no significant difference in the overall amount of science-related arguments in groups discussing climate change in science classrooms when compared to Politics. As well, in everyday life or public media-based discussions, most arguments are not very complex and reflective. Also this finding was mirrored in this study. Even in public discussions arguments often are just repetitions of previously heard or memorized statements. Therefore, the low levels of complexity and non-reflectiveness should not be viewed too critically. However, this finding might be used in the classroom. The results from this analysis might allow for reflection with the students where their arguments stem from. Such a reflection on role-playing exercises might lead to a more well-reasoned view of the potential of science as a source for informed decisions in the classroom, but also to an admission of the limitations of science for contributing to open debates on controversial, socioscientific issues.

Overall, this study reveals that role-playing has the potential to mimic societal debating and consensus-building processes in the science classroom. However, such methods are difficult to carry out for many students and teachers (in the pedagogical sense). While older, more experienced learning groups often behave quite well, there exist other settings where no real "play acting" arises. This can result from insufficiently developed communication and argumentation skills among younger or lower-achieving students. However, it can also be a result of an overly dominate role chosen by the moderator(s), whether this comes from a group of students or the teacher. If the teacher is too actively involved into the exercise, both the quality of the debate and the processes leading to decision-making and consensus-building decline (Nielsen, 2009).

We also found that careful reflection is necessary on how students are required to familiarize themselves with their personal roles. Very open, brief role cards are quite demanding mentally, but allow learners to think out their own arguments. Long, detailed role cards lead to more qualitative arguments, but sometimes hinder open, free-flowing discussions. Students' arguments would quite probably be less complex without

detailed role cards, but would more accurately represent their own thinking and not simply a regurgitation of pre-sorted information. Even if polemical or inconsiderate arguments arise, these can often be used to bring about critical dispute in the classroom. All these points demand reflection on when and how teachers introduce role-playing into classroom situations. For example, the question of how to introduce role cards must also be addressed. Whenever the topic and the roles are close to the learners' everyday environment, less details are necessary. If the issue or roles are more abstract, more guidance might be appropriate.

One last implication concerns the teacher's role. The current study showed that role-playing is a promising pedagogy, yet teachers often have difficulty in applying it. The preparation, introduction, and application of role-playing require a large amount of specific knowledge and particular teaching skills. While role-playing exercises are quite common in the humanities, they seem to be much less common in science education (Patronis et al., 1999; McSharry & Jones, 2000). Increased implementation of such exercises will therefore demand innovation in science teacher education. Role-playing-related strategies, examples, and research evidence must be specifically covered in science teacher education programs. A more through learning of how to introduce and use role-plays in science education might have positive influence not just on a broader implementation, but also on the effectiveness in its use.

REFERENCES

Bell, R. L. & Lederman, N. G. (2003). Understandings of the nature of science and decision making on science and technology based issues. *Science Education, 87*, 352–377.

Dawson, V. M. & Venville, G. (2010). Teaching strategies for developing students' argumentation skills about socioscientific issues in high school genetics. *Research in Science Education, 40*, 133–148.

Duschl, R. A. & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education, 38*, 39–72.

Duveen, J. & Solomon, J. (1994). The great evolution trial: Use of role play in the classroom. *Journal of Research in Science Teaching, 31*, 575–582.

Eilks, I., Feierabend, T., Hößle, C., Höttecke, D., Menthe, J., Mrochen, M. & Oelgeklaus, H. (2011a). Bewerten Lernen und der Klimawandel in vier Fächern—Einblicke in das Projekt "Der Klimawandel vor Gericht" (Teil 1). *Der Mathematische und Naturwissenschaftliche Unterricht, 64*, 7–11–71–76.

Eilks, I., Feierabend, T., Hößle, C., Höttecke, D., Menthe, J., Mrochen, M. & Oelgeklaus, H. (2011b). *Der Klimawandel vor Gericht*. Halbergmoos: Aulis.

Eilks, I., Prins, G. T. & Lazarowitz, R. (2013). How to organize the classroom in a student-active mode. In I. Eilks & A. Hofstein (Eds.), *Teaching chemistry—a studybook* (pp. 183–212). Rotterdam, the Netherlands: Sense.

Eilks, I., Nielsen, J. A., & Hofstein, A. (2014). Learning about the role of science in public debate as an essential component of scientific literacy. In C. Bruguière, P. Clément, A. Tiberghien (Eds.), *9th ESERA Conference Contributions: Topics and trends in current science education*. Dordrecht, the Netherlands: Springer (in press)

Eilks, I. & Ralle, B. (2002). Participatory action research within chemical education. In B. Ralle & I. Eilks (2002) (Eds.), *Research in chemical education—what does it mean?* (pp. 87–98). Aachen, Germany: Shaker.

Erduran, S. & Jiménez-Aleixandre, M. (Eds.). (2007). *Argumentation in science education*. New York: Springer.

Erduran, S., Simon, S. & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education, 88*, 915–933.

Feierabend, T. & Eilks, I. (2010). Raising students' perception of the relevance of science teaching and promoting communication and evaluation capabilities using authentic and controversial socio-scientific issues in the framework of climate change. *Science Education International, 21*, 176–196.

Feierabend, T. & Eilks, I. (2011). Innovating science teaching by participatory action research—reflections from an interdisciplinary project on curriculum development in the field of climate change. *Centre for Educational Policy Studies Journal, 1*, 93–112.

Feierabend, T., Stuckey, M., Nienaber, S. & Eilks, I. (2012). Two approaches for analyzing students' competence of 'evaluation' in group discussions about climate change. *International Journal of Environmental and Science Education, 7*, 581–598.

Fleming, R. (1986). Adolescent reasoning in socio-scientific issues part II: Nonsocial cognition. *Journal of Research in Science Teaching, 23*, 689–698.

Hofstein, A., Eilks, I. & Bybee, R. (2011). Societal issues and their importance for contemporary science education: A pedagogical justification and the state of the art in Israel, Germany and the USA. *International Journal of Science and Mathematics Education, 9*, 1459–1483.

Holbrook, J. & Rannikmäe, M. (2007). The nature of science education for enhancing scientific literacy. *International Journal of Science Education, 29*, 1347–1362.

Hollingworth, P. M. & Hoover, K. H. (1991). *Elementary teaching methods*. Boston: Allin and Bacon.

Killen, R. (2009). *Effective teaching strategies*. Melbourne, Australia: Cengage Learning Australia.

Kolstø, S. D. (2006). Patterns in students' argumentation confronted with a risk-focused socio-scientific issue. *International Journal of Science Education, 28*, 1689–1716.

Kslafki, W. (2000). The significance of classical theories of Bildung for a contemporary concept of Allgemeinbildung. In I. Westbury, S. Hopmann & K. Riquarts (Eds.), *Teaching as reflective practice. The German Didaktik Tradition* (pp. 85–107). Mahwah, NJ: Lawrence Erlbaum.

Mamlok-Naaman, R. & Eilks, I. (2012). Action research to promote chemistry teachers' professional development—cases and experiences from Israel and Germany. *International Journal of Mathematics and Science Education, 10*, 581–610.

Marks, R., Bertram, S. & Eilks, I. (2008). Learning chemistry and beyond with a lesson plan on potato crisps, which follows a socio-critical and problem-oriented approach to chemistry lessons—a case study. *Chemistry Education Research and Practice, 9*, 267–276.

Marks, R. & Eilks, I. (2009). Promoting scientific literacy using a socio-critical and problem-oriented approach to chemistry teaching: Concept, examples, experiences. *International Journal of Environmental and Science Education, 4*, 131–145.

Marks, R. & Eilks, I. (2010). Research-based development of a lesson plan on shower gels and musk fragrances following a socio-critical and problem-oriented approach to chemistry teaching. *Chemistry Education Research and Practice, 11*, 129–141.

McSharry, G. & Jones, S. (2000). Role play in science teaching and learning. *School Science Review, 82*(298), 73–82.

Mitchell, S. (1996). *Improving the quality of argument in higher education interim report*. London: Middlesex University.

Nielsen, J. A. (2009). Structuring students' critical discussions through processes of decision-making on socio-scientific controversies. *Revista de Estudos Universitários, 35*, 139–165.

Nielsen, J. A. (2011). Dialectical features of students' argumentation: A critical review of argumentation studies in science education. *Research in Science Education*, 43, 371–393.

Nielsen, J. A. (2012). Science in discussions: An analysis of the use of science content in socioscientific discussions. *Science Education, 96*(3), 428–456.

Ødegaard, M. (2003). Dramatic science. A critical review of drama in science education. *Studies in Science Education, 39*, 75–102.

Oulton, C., Dillon, J. & Grace, M. (2004). Reconceptualizing the teaching of controversial issues. *International Journal of Science Education, 26*, 411–423.

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

Roth, W.-M. & Lee, S. (2004). Science education as/for participation in the community. *Science Education, 88*, 263–291.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching, 41*, 513–536.

Sadler, T. D. & Zeidler, D. L. (2005). Patterns of informal reasoning in the context of socioscientific decision-making. *Journal of Research in Science Teaching, 42*, 112–138.

Saunders, D., Percival, F. & Vartiainen, M. (1996). *The simulation and gaming yearbook. Vol. 4*. London: Kogan Page.

Simmoneaux, M. (2001). Role play or debate to promote students' argumentation and justification on an issue in animal transgenesis. *International Journal of Science Education, 23*, 903–927.

Simmoneaux, M. (2002). Analysis of classroom debating strategies in the field of biotechnology. *Journal of Biological Education, 37*, 9–12.

Simon, S. (2008). Using Toulmin's argument pattern in the evaluation of argumentation in school science. *International Journal of Research & Method in Education, 31*, 277–289.

Sjöström, J. (2011). Towards Bildung-oriented chemistry education. *Science & Education*, *22*, 1873–1890. doi:10.1007/s11191-011-9401-0.

Solomon, J. (1992). The classroom discussion of science-based social issues presented on television: Knowledge, attitudes and values. *International Journal of Science Education, 14*, 431–444.

Strauss, A. & Corbin, J. (1990). *Basics of qualitative research*. Thousand Oaks, CA: Sage.

Swanborn, P. G. (1996). A common base for quality control criteria in quantitative and qualitative research. *Quality and Quantity, 30*, 19–35.

Tytler, R., Duggan, S. & Gott, R. (2001). Dimensions of evidence, the public understanding of science and science education. *International Journal of Science Education, 23*, 815–832.

Van Ments, M. (1999). *The effective use of role play*. London: Kogan Page.

Wilhelm, J. (2002). *Action strategies for deepening comprehension*. New York: Scholastic.

Yang, F.-Y. & Anderson, O. R. (2003). Senior high school students preference and reasoning modes about nuclear energy use. *International Journal of Science Education, 25*, 221–244.

Zeidler, D. L., Walker, K. A., Achett, W. A. & Simmons, M. L. (2002). Tangled up in views: Beliefs in the nature of science and responses to socioscientific dilemmas. *Science Education, 86*, 343–367.

Department of Biology and Chemistry, Institute of Science Education
(IDN)—Didactics of Chemistry
University of Bremen
Leobener Straße NW2, 28334 Bremen, Germany
E-mail: n.belova@uni-bremen.de