



UvA-DARE (Digital Academic Repository)

Effects of dialogic learning on value-loaded critical thinking

Frijters, S.; ten Dam, G.; Rijlaarsdam, G.

DOI

[10.1016/j.learninstruc.2006.11.001](https://doi.org/10.1016/j.learninstruc.2006.11.001)

Publication date

2008

Document Version

Final published version

Published in

Learning and Instruction

[Link to publication](#)

Citation for published version (APA):

Frijters, S., ten Dam, G., & Rijlaarsdam, G. (2008). Effects of dialogic learning on value-loaded critical thinking. *Learning and Instruction*, 18(1), 66-82.
<https://doi.org/10.1016/j.learninstruc.2006.11.001>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

Effects of dialogic learning on value-loaded critical thinking

Stan Frijters, Geert ten Dam, Gert Rijlaarsdam*

Graduate School of Teaching and Learning, University of Amsterdam, Spinozastraat 55, 1018 HJ Amsterdam, The Netherlands

Received 20 January 2006; revised 22 October 2006; accepted 27 November 2006

Abstract

Education should endeavour to include the competences that students need to participate in a democratic society in a responsible and adequate way. The crucial question is how this can be realized in a curriculum aimed at domain-specific knowledge and skills and organized in school subjects. In this article we present a study on teaching value-loaded critical thinking in pre-vocational secondary education implemented in the school subject of Biology. The sample comprised 297 students. We developed two lesson series for teaching value-loaded critical thinking: a Dialogic lesson series and a Non-Dialogic one. The results indicated that the dialogic learning condition, compared to the non-dialogic, resulted in a more positive effect on the critical-thinking competences of the students, both in terms of generative fluency of reasoning and quality of value orientation. The higher scores on critical-thinking variables were not at the expense of the subject-matter knowledge of Biology taught in these lessons.

© 2006 Published by Elsevier Ltd.

Keywords: Dialogic learning; Attitudes; Values; Reasoning; Biology

1. Introduction

In Western Europe there is a general tendency to emphasize the social task of education. Education should endeavour to include the competences that students need to participate in a responsible and adequate manner in a democratic society (e.g., Rychen & Salganik, 2003). More specifically, the prosocial and moral development of students as a responsibility of schools has increasingly been recognized (Solomon, Watson, & Battistich, 2001).

This development towards a social reconstructionist curriculum (McNeil, 1996) is indicated by the promotion of citizenship education in several countries in the last few decades. How this education takes form, however, varies. In the UK, citizenship education was introduced as a compulsory part of the renewed national curriculum in the 1990s (Kerr, 1999) and as a separate subject in primary and secondary education since 2001. The same happened in France with *éducation civique*. According to the Dutch Education Council fostering citizenship, i.e., stimulating the willingness and ability to become an active member of a community, should also be part of the regular subjects (languages, history, biology, etc.) (Onderwijsraad, 2003). In the Netherlands, the bill ‘Active Citizenship and Social Integration’

* Corresponding author. Tel.: +31 20 5251288; fax: +31 20 5251290.

E-mail address: G.C.W.Rijlaarsdam@uva.nl (G. Rijlaarsdam).

was passed in 2005 obliging schools to integrate citizenship in the curriculum of both primary and secondary education.

Integrating citizenship education in the regular curriculum implies that students must learn to act and behave responsibly and adequately in society with the support of domain-specific knowledge and skills. Such learning goals in particular refer to learning to think critically, to learning values and reflecting upon them, and to communicative skills. The focus on linking academic subjects with social issues is receiving increasing support in curriculum studies in general and in science in particular (see, e.g., Yore, Bisanz, & Hand, 2003). At the same time, at the level of educational practice, inspiring and meaningful examples are relatively sparse. Moreover, empirical research into the effectiveness of the proposed instructional designs appears to be non-existent (what do students actually learn?) (Schuitema, ten Dam, & Veugelers, *in press*).

In this article, we focus on critical thinking as a crucial component of the competences that citizens need in order to participate in a modern democratic society. We examined whether dialogic learning is a good candidate for an effective learning environment, i.e., an environment in which the emphasis is on interaction, mutual understanding and collaborative meaning making.

1.1. Critical thinking

For identifying effective learning environments, we relied on two perspectives on critical thinking (see Ten Dam & Volman, 2004): a psychological and a critical pedagogical perspective. Psychologists conceptualize critical thinking as a higher-order thinking skill and focus on the appropriate learning and instruction processes. Critical pedagogy emphasizes critical and democratic citizenship and the importance of value development.

Although most authors agree that critical thinking involves both skills and dispositions (see Burbules & Berk, 1999), in empirical research the emphasis is on thinking *skills* that are conceptualized as higher-order thinking skills. Halpern (2003, p. 6), for example, defines critical thinking as “cognitive skills and strategies that increase the likelihood of a desired outcome... thinking that is purposeful, reasoned, and goal-directed — the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions.”

1.1.1. Psychological research on critical thinking

Instructional designs aimed at critical thinking stress the importance of interaction between students. Paul (1992) argued that students learn best when their thinking involves an extended exchange of points of view or frames of reference. More specifically, various forms of cooperative learning are proposed. For example, focused discussions, student-led seminars, problem-based learning (Dennick & Exley, 1998), and the method of ‘academic controversy’ (Johnson & Johnson, 1993). Empirical research on the effectiveness of cooperative learning for enhancing critical-thinking skills, however, is sparse. Garside (1996) reported an experimental study (pretest–posttest design) on the question whether group discussion facilitates the development of critical-thinking skills more than traditional methods of instruction such as lectures. No significant difference was observed between the two instructional methods in developing critical-thinking skills. Garside attributed this result to the students’ lack of experience of group discussions: communicative skills were taken for granted instead of being trained.

A second instructional strategy for enhancing critical thinking is to make use of meaningful, authentic, domain-specific subject matters. Using real-life problems is supposed to motivate students and to stimulate their active involvement; moreover, real-life problems are precisely the kind of ill-defined, complex problems for which critical thinking is needed (Brown, 1997). Studies that correlated learning environments with students’ critical-thinking skills confirmed the importance of the active involvement of students in the learning process (Terenzini, Springer, Pascarella, & Nora, 1995; Tsui, 1999). If significant effects of such course characteristics are *not* found, this is explained by the limited instruction time, the small sample and/or the use of too general measures. With regard to measures, essay tests and ‘authentic assessments’ are considered to improve the validity of assessing critical-thinking competences in a particular domain (Tynjälä, 1998).

Moreover, research on teachers’ beliefs about ways that stimulate the critical-thinking skills of students shows that teachers are more likely to support the use of higher-order critical thinking activities with high-ability learners than with low-ability learners (Torff, 2005; Zohar, Degani, & Vaakin, 2001). This suggests an interaction between cognitive ability and critical thinking instruction. Empirical research, however, is virtually non-existent.

Another pervasive issue in research on critical thinking pertains to the generalizability—specificity debate (see Siegel, 1992). The instructional variant of this debate centres around the question of whether critical-thinking skills should be taught in the context of specific subject matter or “out of context” in programs especially devised to improve critical thinking. Studies show that special “critical-thinking programs” do not usually result in long-lasting effects (see Tsui, 1999). But at the same time, several authors point out that some general principles of critical thinking transcend specific subjects (e.g., Klaczynski, 2001; Stanovich & West, 2000; Tsui, 1999). Nowadays, in this line of thinking, the focus in psychological research is on crossdomain transfer of (domain-general) critical-thinking skills (Halpern, 2003; Stark, Mandl, Gruber, & Renkl, 1999; Toplak & Stanovich, 2002). Thus, critical thinking must be taught in the context of specific meaningful subject matter in such a way that transfer to other domains is possible (cf. Brown, 1997).

1.1.2. Critical pedagogical research on critical thinking

Critical pedagogues argue that critical thinking focuses too much on internal consistencies while paying too little attention to the political nature of arguments and reasonings: for example, doing justice or injustice to others (Giroux, 1994; see also Veugelers, 2000). Other scholars have emphasized that by focusing on the ability to think logically, critical thinking neglects the importance of a caring attitude, empathy and commitment (Burbules & Berk, 1999; Noddings, 1992). Critical thinking should combine logical reasoning with value development to enable citizens to make their own contribution to society in a critical manner, with sensitive awareness. We use the concept ‘value-loaded critical thinking’ to make apparent the inherent normative character of critical thinking. The subject matter about which students are expected to reason is always ridden with implicit values. We agree with Burbules and Berk (1999) who point out that the strict distinction between facts and values in the critical-thinking tradition makes the consequences of certain knowledge on institutional and societal levels invisible. This plea for conceptualizing critical thinking as value-loaded reasoning, however, has so far not been supported by empirical research on learning environments that aim at stimulating critical thinking.

1.2. Dialogic learning

Critical thinking and dialogue are often linked. Paul (1992, 1994) considers fostering dialogue to be part of the method of critical thinking, because dialogue makes it possible to take the perspective of others into account, which is necessary for ‘the assessment of truth claims’. Instructional formats in which cooperative learning and dialogue feature are expected to promote the students’ active learning and higher-order thinking skills simultaneously (Renshaw, 2004; Salomon & Perkins, 1998).

From a cognitive perspective it is emphasized that social interaction affects cognitive elaboration processes (Dekker, Elshout-Mohr, & Wood, 2004; O’Donnell & King, 1999). Language functions such as explaining, reasoning, asking questions, stimulate thinking and the development of knowledge (cf. “exploratory talk”, Mercer, 2000; Wegerif, Mercer, & Dawes, 1999). From a social constructivist approach, the focus is on “the process of becoming a member of a certain community” (Wells, 2000). Learning is seen as a ‘dialogue’, a form of collaborative meaning making (co-construction) (Ten Dam, Volman, & Wardekker, 2004).

In the present study we borrowed elements from a cognitive as well as a social constructivist perspective on ‘dialogic learning’. The importance of ‘elaboration’ in enhancing critical-thinking skills is taken from the cognitive perspective. Two elements are taken from a social constructivist perspective. First, the learning content should be meaningful in the sense that the subject matter must relate to the real world, to students’ own position and that of others, and to students’ opportunities to influence this position. Second, engagement in dialogues in the classroom should encourage students to gain practical experience in “skills and dispositions” such as “being able to relate a question to one’s own standards and values”, “being able to relate a question to general principles such as social justice, equality, respect and consideration”, “being open towards and considering other people”, “daring to express a different opinion”. Interaction in small groups is better suited to this than discussions involving the whole class (Haworth, 1999).

Dialogic learning is a potentially adequate instructional strategy for stimulating value-loaded critical thinking. It stimulates a student to process beyond the level of ‘knowledge telling’ (Bereiter & Scardamalia, 1987): the questions that the other participants ask stimulate further thinking and add other (moral) perspectives to the issue at hand.

1.3. Science education

In the present study Biology was chosen as a domain in which students can become involved with meaningful, authentic subject matters. According to [Sadler and Zeidler \(2005\)](#) science education should contribute, in part, to citizenship education.

[Driver, Newton, and Osborne \(2000\)](#) make a strong case for seeing argument and argumentative practice as a core activity of scientists. Students should learn to think critically about the knowledge claims of science and the criteria used to establish scientific theories. Their focus is primarily on establishing better science education, i.e., helping young people to engage with the social practice of scientists. At the same time [Driver et al. \(2000\)](#) emphasize the importance of including so-called “socioscientific issues” in the curriculum.

Empirical studies on ‘argumentative’ science education focus on students’ higher-order thinking skills, namely identifying assumptions and valid information sources and drawing valid conclusions (e.g., [Kuhn & Udell, 2003](#); [Osborne, Erduran, & Simon, 2004](#); [Zohar & Nemet, 2002](#)). In almost all of these studies interaction with and between students is investigated as a means to improve argumentation. In this field, the attention paid to *value-loaded* reasoning is relatively sparse again. One of the few exceptions is the study on informal reasoning by [Sadler and Zeidler \(2005\)](#). The authors focus explicitly on the moral aspects of decision making. In their study it is demonstrated that in science classrooms where decision making on socioscientific issues is asked, students can be stimulated to explicitly explore their own principles, emotions, and intuitions pertinent to science and its social applications. As [Sadler and Zeidler \(2005, p. 130\)](#) state “This can only enhance their roles as citizen participants in democratic societies largely influenced by science and technology”.

One needs, however, a more detailed analysis of the required skills and desired learning activities for designing effective learning environments. [Driver et al. \(2000\)](#) provide such an analysis. They have included values in their list of skills for scientific literacy. Students should be able to distinguish between questions that have a scientific base and questions that relate to other types of knowledge (e.g., ethical, economic, legal questions), and they should also be able to recognize personal and social values and perspectives that impact decision making in science. In the present study we designed a learning environment that implements value-loaded thinking, and tested whether a dialogic format is more effective than a non-dialogic.

1.4. Aims – hypotheses

In the present study we implemented a dialogic approach to teaching and learning aimed at fostering critical thinking in the context of domain-specific subject matter. The research question was: What is the effect of dialogic learning activities, as compared to non-dialogic, on value-loaded critical thinking with regard to environmental issues within the subject of Biology and on the learning results of students? In the study we implemented a pretest–posttest experimental design.

The basic principles for designing the experimental and the control conditions were the following. First, learning to think critically should involve reasoning as well as value orientation. Second, learning to think critically should take place in the context of meaningful, rich, domain-specific subject matters. Both principles were followed in the design of the experimental and the control condition. Third, learning to think critically should take place in dialogue with other learners; this principle was applied only in the Dialogic condition.

The following hypotheses guided the study:

1. The dialogic approach will have a more positive effect than the non-dialogic (a) on students’ argumentations, that is fluency of reasoning, (b) on their value orientation, and (c) on their attitudes towards dialogic learning, without detracting from content knowledge (Biology). That is, the two learning conditions will similarly affect the content knowledge.
2. The students with high academic aptitude will profit most from the dialogic approach, because value-loaded critical thinking as well as the instruction format of dialogic learning is new for most students.

We also explored other aptitude-treatment interactions, such as condition by attitude to dialogic learning and general reasoning skills, but no prediction was made due to lack of previous research.

2. Method

2.1. Participants

A total of 297 students (54.9% female) from 14 Grade 8 classes from the pre-vocational track of three secondary schools in Amsterdam participated in the study. From an academic point of view the pre-vocational track is the lowest of the four tracks in the secondary educational school system in the Netherlands. The three participating schools have an ethnically heterogeneous population and provide secondary education at different levels, ranging from pre-vocational to pre-university education, the highest track in the system. Students with learning and/or behavioural disabilities were excluded from the study.

School A is a small secondary school for Individual Artistic Education with approximately 300 students. The majority of students are from a higher than average socio-economic background. School B is a moderately large school with approximately 1300 students. Very few students at this school come from a higher than average socio-economic background. School C has a similar size with approximately 1350 students. Most of the parents have a relatively low socio-economic status (SES); some parents with a higher than average SES make a principled choice to let their children attend this school because of the diversity policy.

The lessons were part of the regular Biology curriculum and were taught by the students' regular Biology teachers. Nine teachers were involved in the study voluntarily. They were all qualified teachers, combining a higher education degree in Biology and teaching. All but two teachers had more than 3 years experience in teaching Biology, varying from 3 years to 25 years of experience. Two were less experienced: they were in their second and third year of teaching.

2.2. Designing the lesson series

Focusing on value-loaded critical thinking in academic subjects is a rather rare phenomenon in Dutch secondary education. Therefore, we could not rely on existing text book materials for designing the dialogic and non-dialogic approach.

The designing process could be labeled as incremental and interactive. First the content of the lesson series was designed, resulting in a basic script (see below). In this process, the designer (the first author) discussed his texts and decisions individually with two teachers (8 and 25 years of teaching experience), and with the co-authors. Then the two teachers were asked for advice on instructional formats for dialogic learning situations; step by step these formats were realized, discussed, revised etc. until the whole Dialogic lesson series had been designed. The result was intensively reviewed by the two teachers and the co-authors; several revisions followed. One teacher tried out the final revision; almost all lessons were observed by the designer, who also interviewed the teacher and some students, and analyzed the students' workbooks. This trial resulted in subsequent revisions.

Finally, two lesson series, each including four units, were available. They were similar in all but one respect: students in the Dialogic lesson series developed values and critical-thinking skills dialogically by communicating and cooperating with at least one other student, with dialogic input, whereas students in the Non-Dialogic lesson series worked individually, the usual teaching format in the Netherlands. Teacher-led whole-class discussions, mostly at the end of a unit, were similar in both conditions.

2.2.1. The basic script

The starting point was the basic script. This script aimed at Biology lessons, focusing on environmental education, as part of the regular curriculum. The generation of the content focused on the national attainment objective as given below:

Students are able to name the most important causes and effects of damage to nature and the environment by certain forms of refuse, by traffic, by agriculture, by recreation and by energy use and can give examples of measures that can contribute to sustainable development of the relationship between human beings and the environment, including measures by students in their own environment; they are able to explain that these influences transcend country boundaries.

A second content design principle for the basic script came from the objectives for value-loaded critical thinking. We distinguished five communicative skills: (1) Being able to share personal opinions with fellow students; (2) Being able to form one's own opinion in a dialogue, utilizing the input of fellow students; (3) Being able to contribute in a dialogue to the opinion forming of fellow students (co-construction of opinions (standards and values)) and to contribute to clarification of the opinions (standards and values) of fellow students; (4) Being able to validate/appraise one's own opinion from the perspective of justice and respect; (5) Being able to validate/appraise the opinions of others from the perspective of justice and respect.

A third design principle was the cumulative order of these five aims. As a result, the distribution of learning tasks followed the cumulative order (see Table 1). In Unit 1, the tasks focused on sharing ideas, constructing one's own opinion, and contributing to others' opinions. Two tasks focused on two skills (Task 3, Task 10), the other tasks focused on one skill. The whole of Unit 2 was dedicated to the third skill, contributing to the others' opinion; Units 3 and 4 focused on the last two skills, namely, legitimating one's own opinion and legitimating other's opinion.

From the basic script, we generated two lesson series, keeping the learning content and order of learning tasks for both conditions the same. The target tasks (see Table 1) were written in two formats: A dialogic and a non-dialogic or individual working format. The basic structure for a dialogic task is shown in Appendix A. Each skill was introduced by a short written simulated dialogue to provide a model. Students then had the opportunity to practice the skill. They first analyzed the opinions and arguments, and then formulated their own opinion, working in pairs or quartets. In the Non-Dialogic condition the assignments included the same information, but the input content was presented in a coherent text, as is usual in student textbooks. Then students answered debate questions, formulated a standpoint and at least one argument.

Both the Dialogic and the Non-Dialogic lesson series consisted of a student book and a teacher's manual. Student books had four units and consisted of approximately 50 tasks. Of these tasks 18 (36%; see Table 1) were dialogic. The whole course took about 8–11 periods of 45 min. The two textbooks were self-instructional, guided by pictograms and standardized guidance and instructions. They both had a glossary aimed at students who had problems with difficult words (domain-specific jargon) and the language of instruction.

Compared with the regular Biology curriculum in the Netherlands, both lesson series were experimental in learning content. For research reasons, however, we appoint the Dialogic condition as the experimental condition and the Non-Dialogic condition as the control condition.

2.3. Measures

2.3.1. Background variables

To test whether the students in the two conditions, differed in terms of abilities that were supposed to be related to the dependent variables, two measures were used.

2.3.1.1. General reasoning skills. Two relevant sections from the so-called Cross-Curricular Skills Test were adapted (Meijer, Elshout-Mohr, & Van Hout-Wolters, 2001), namely, "forming opinions", and "distinguishing opinions from facts". The total number of items was 14. Reliability proved to be satisfactory for research purposes ($\alpha = 0.68$).

2.3.1.2. Academic aptitude. Data were collected from the school administrations. They were test scores for language and mathematics as measured by the Primary Education Final Test (a standard test in the Netherlands), which is

Table 1
Distribution of learning tasks over the Dialogic lesson series

Skill	Unit 1	Unit 2	Unit 3	Unit 4
Sharing ideas	3, 8, 9, 15			
Constructing own opinion	10, 17			
Contributing to other's opinion	3, 4, 10	6, 9, 10, 11		
Legitimating own opinion			5, 9	3
Legitimating other's opinion			6, 7	2, 3, 4, 5

Note: Numbers in the cells refer to the task within the learning unit.

administered in Grade 6, the last year of primary education. Schools use these test scores to place students in the appropriate secondary-education track.

2.3.2. Dependent variables

2.3.2.1. Biology subject-matter knowledge. The test consisted of 10 questions. It covered knowledge of environmental education issues (e.g., the biologic concept of “recycling”), applying knowledge to environmental issues (e.g., recycling), and applying ecological concepts to an environmentally ‘authentic’ context (e.g., a food chain in a marine environment). Question format varied from multiple choice (6), short-answer questions (2), to reasoning questions (2). Example of multiple-choice item is the following: “Place five elements of a food chain in the correct order: (a) herbivore, (b) shit of the carnivore, (c) bacterium and moulds, (d) plants, and (e) carnivore.” Cronbach’s alpha was 0.65.

2.3.2.2. Attitudes to dialogic learning. We developed, piloted, and revised an inventory of attitude towards three concepts underlying dialogic learning: (1) Sharing ideas, (2) Co-construction of ideas, and (3) Legitimizing the ideas. Representative items are, respectively, “I think that the others are interested in what I think”, “By talking about everyone’s opinions with each other, we can make progress together”, and “When I hear someone else’s opinion I always ask myself whether it is fair”. Responses were given on a five-point Likert scale, varying from “1 = totally agree” to “5 = totally disagree”. The reliability of the inventory for each of the three concepts and of the inventory as a whole is given in Table 2. The intercorrelations between the three concepts were significant but low, from 0.16 to 0.38.

2.3.2.3. Generative fluency of reasoning (GFR) and quality of value orientation (QVO). To measure value-loaded critical thinking we constructed a curriculum-independent measure, that is, a written essay test. The principles of the Pragmatic–Dialectic School of Argumentation (Van Eemeren et al., 1996) were applied. This approach to argumentation sees it as a dialogic, reciprocal discussion on an issue, in which all participants have an equal responsibility to solve the difference of opinion. The protagonist initiates the discussion by formulating his/her standpoint on the issue at hand and provides at least one argument. The antagonist must critically respond with a counter argument, or express his/her doubts concerning the data that underpin the protagonist’s claim. The issue of value orientation (the importance of justice and respect) was also added to warrant or doubt the quality of an argument.

To measure the fluency of argumentative behaviour in these two roles, two short essay tests were developed: a protagonist test and an antagonist test. To increase the level of generalization, two tasks per role were designed. In the antagonist role, participants were confronted with a written position on a certain issue under discussion from a fictitious peer. The written position was pre-selected from a set of classroom papers and was piloted. The issues were “preventing pollution by recycling” and “export of waste to developing countries”. The participants were asked to determine their own position and to give reasons for their choice: Did they agree with the peer’s positions or not, and why? In the protagonist role, participants were confronted with two environmental issues, set in a fictitious but meaningful framework of a specific case. The issues were “small domestic appliances: throw away or recycle?”

Table 2
Reliability indices of the study’s measures

Measure	No. of items	Reliability
Attitude to dialogic learning (pretest)		
All items	25	0.77
Sharing	8	0.71
Co-construction	9	0.68
Legitimizing	8	0.73
Attitude to dialogic learning (posttest)		
All items	25	0.79
Sharing	8	0.74
Co-construction	9	0.57
Legitimizing	8	0.75

and “the production of paper and its consequences for forests”. Participants were asked to elaborate on their opinions regarding each of the issues, to generate a ‘solution’ and to support this solution with reasons and arguments. In the protagonist role, the students had to generate their own standpoint and at least one argument to open the discussion.

The written responses were coded to indicate two variables: generative fluency of reasoning (GFR) and quality of value orientation (QVO). The GFR refers to the argumentative productivity of either the protagonist or the antagonist. The GFR scores are based on the number of points raised or questioned by the participant when responding to the issue (protagonist) or the position taken (antagonist) and the number of arguments in support of the positions taken by the participant, the number of coordinated and subordinated arguments generated by participants in support of the measures that they proposed in the two protagonist tasks. In total we distinguished 14 categories of argumentative types, and counted how often a type was realized in a text (variation was large, ranging from 0 to about 40). Interrater agreement for GFR was 0.71 and the internal consistency was $\alpha = 0.80$.

QVO refers to the values to which the protagonist or antagonist refers when exchanging arguments. All student texts were exploratorily coded by two raters for the value orientation in the written response. In total 21 different categories of value orientation were created to code these orientations in the antagonist role. Of them 11 categories indicated a social value orientation. Examples of the orientations are “respect for nature” and “waste and future generations”. Of the 19 categories generated by coders for the two protagonist tasks, 10 indicated a social value orientation. Examples are “conservation of the environment” and “waste prevention”. The interrater agreement of the categorization of students’ texts was 0.63 and 0.81 for the antagonist and protagonist, respectively.

The score for QVO was how often a student referred to social values. The correlation between the two resultant QVO scores was significant but low (0.12); therefore, the scores were kept separate as indicative of two different constructs: QVO-protagonist (QVO-P) and QVO-antagonist (QVO-A). The correlation of GFR with QVO-A and QVO-P was $r = 0.45$ and $r = 0.55$, respectively. The moderate relationship indicated that GFR and QVO could be considered to be two separate constructs.

2.3.2.4. Implementation measures. A detailed manual for teachers was written so that the lesson material could be properly implemented. The manual drew attention to both the educational vision on which the material was based and the instruction strategies in the classroom for each lesson and task. The quality of implementation was further enhanced by training the teachers individually (2 h per teacher) and checked by informal visits and formal observations.

During formal observations research assistants working in pairs observed key episodes from the lesson series. Each observer pair, randomly assigned to groups, chose randomly six students from different pairs or seating groups to observe. A time-sampling scheme was applied. As soon as the key task started, the research assistants began observing Student 1 for 15 s, noted the most dominant learning behaviour, observed Student 2 for 15 s and so on. After Student 6, there was a short break of 30 s. The period of observation ended when the key task was finished in the class. We distinguished the following activities: Dialogic work; Individual work; Doing nothing because task has been finished; Non-task related behavior, and a Rest category. As observers worked in pairs providing one score, no interrater indices could be calculated.

2.4. Design – procedure

Four teachers (from three schools) participated with two classes in the study; five other teachers from the same schools participated with one class. For the teachers with two classes, one class was randomly selected as experimental and the other as control class. Complete classes from the other five teachers were randomly assigned to one of the two conditions. The result is a mixed research design: the teacher effect is controlled in the four cases where one teacher taught two classes, one in each condition; a possible teacher effect might play a role in interpreting the experimental effects in the five cases where a teacher taught just one class. In data analyses we added a Design factor with two values ‘balanced/unbalanced’ to check the experimental design effect.

The research assistants administered the pretest measures before the lessons started and posttest measures 1–2 weeks after the last lesson finished. For attitudes, a pretest–posttest design was realized; for the other dependent variables (GFR, QVO, subject-matter knowledge) a posttest only design was implemented, with relevant proxy variables as possible covariate.

The duration of the lesson series in the classroom varied from eight to 11 lessons and was similar across conditions.

3. Results

3.1. Testing initial differences

Univariate analyses of variance did not reveal initial differences between the two learning conditions as regards general reasoning skills, $F(1, 256) = 0.021$, $p = 0.885$, $\eta^2 = 0.000$, general academic aptitude, $F(1, 251) = 0.359$, $p = 0.55$, $\eta^2 = 0.001$, and attitude to dialogic learning, $F(1, 250) = 0.043$, $p = 0.835$, $\eta^2 = 0.000$ (see Table 3). For the three concepts addressed in the attitude to dialogic learning, a 2(conditions) \times 3(dependent variables) multivariate analysis showed no difference between conditions, Pillai's trace = 0.002, $F(3, 248) = 0.16$, $p = 0.816$, $\eta^2 = 0.002$. The distribution of male/female was not significantly different over conditions, $\chi^2 = 1.958$, $p = 0.162$. Therefore, the two learning conditions did not differ on all relevant variables.

3.2. Validity of implementation

To check the quality of implementation we observed key episodes, where learning activities in the students' workbooks should result in different learning behaviours for the two conditions. In the Dialogic condition more time should be spent on dialogue than in the Non-Dialogic condition, while in the Non-Dialogic condition more time should be spent on individual work. No differences should be revealed between the non-task-related behaviours, as this indicates the general time-on-task behaviour. Further, no differences should be observed between the time students waited after having finished a task before they could proceed with another task. This proportion of time indicates how efficiently the condition was implemented. Table 4 shows the proportion of time devoted in the two conditions to dialogue and individual learning, to waiting time and to non-task-related behaviour. In the Dialogic condition significantly more time was devoted to Dialogic work (effect size about 1 standard deviation), while in the Non-Dialogic condition the dominant activity was Individual work (effect size also about 1 standard deviation). These results indicate that in both conditions the intended learning activities took place. It is interesting that not much time was lost in either condition after students finished their tasks, as in both conditions students could proceed without a loss of time (7% and 3% of time in the Dialogic and Non-Dialogic condition, respectively). About 70% of the time was devoted to learning activities in both conditions, which is quite a lot in lower secondary classes in the Netherlands, especially in the pre-vocational education track.

3.3. Effects of the independent variables

A multivariate analysis of outlier data resulted in deletion of 21 score sets. Of the participants 20 were not included in the data analysis because of incomplete pretest or posttest data.

To test Hypothesis 1 the effect of the independent variable, i.e., condition, was estimated by applying (multivariate) variance of analysis, with the relevant pretest measures as covariate.

When exploring interaction effects (including Hypothesis 2), a two-step procedure was followed. First, ANCOVAs with a condition-specific defined covariate (e.g., General reasoning skill in Dialogic condition and general reasoning skill in Non-Dialogic condition) were run. If one or the other covariate proved to contribute significantly, regression analyses were run to estimate the beta weight. This weight was then used in figures to illustrate the interactions between student variables and the independent variable.

Table 3
Mean (and standard deviations) of pretest measures for the two experimental conditions

Measure	Dialogic condition	Non-Dialogic condition
General reasoning skills	0.55 (0.15)	0.55 (0.15)
General academic aptitude	531 (7.7)	530 (7.2)
Attitude to dialogic learning	2.3 (0.37)	2.3 (0.40)
Sharing	2.19 (0.56)	2.16 (0.53)
Co-construction	2.39 (0.42)	2.42 (0.52)
Legitimizing	2.32 (0.53)	2.36 (0.63)

Table 4

Proportion of learning time devoted to learning activities during key tasks in the two experimental conditions: mean (and standard deviations) and number of lessons observed

	Dialogic condition ($n = 30$)	Non-Dialogic condition ($n = 16$)
Individual work**	0.26 (0.17)	0.44 (0.24)
Dialogue*	0.37 (0.18)	0.19 (0.20)
Doing nothing, tasks finished***	0.07 (0.12)	0.03 (0.05)
Not task related***	0.30 (0.17)	0.34 (0.13)

Note: * $F(1, 46) = 9.310$, $p = 0.004$, $\eta^2 = 0.157$, no effect of design, no design by condition effect. ** $F(1,46) = 6.587$, $p = 0.014$, $\eta^2 = 0.136$, no effect of design, no design by condition effect. ***No condition effect, no design effect, no condition by design interaction.

3.3.1. Subject-matter knowledge

The two learning conditions were expected to reach the same level of subject-matter knowledge. A 2(condition) \times 2(design) ANCOVA with general academic aptitude and general reasoning skills as covariates revealed no statistically significant main effect of condition, $F(1, 222) = 0.000$, $p = 0.989$, $\eta^2 = 0.000$ (see Table 5). No main effects of design (balanced vs. unbalanced) and no interaction between condition and design were observed. This implies that the level of subject-matter knowledge was not affected by the learning condition and the research design.

3.3.2. Attitudes to dialogic learning

First we analyzed the effect on the attitude to dialogic learning, as a whole, using the compound score of the three concepts. This attitude score proved to be relatively stable over time. The correlation between pretest and posttest measurements over a period of about 6 weeks was moderately high and about the same in both conditions (0.74 in the Dialogic condition, 0.67 in the Non-Dialogic condition). ANCOVA with the respective pretest measure as covariate, $F(1, 235) = 0.347$, $p = 0.57$, $\eta^2 = 0.001$, showed no effect of condition on the attitude to dialogic learning. A 2(condition) \times 2(design) \times 2(time) repeated measures MANOVA showed a significant main effect of time, Pillai's trace = 0.018, $F(1, 232) = 4.32$, $p = 0.039$, $\eta^2 = 0.018$, but no interaction between condition and time, Pillai's trace = 0.001, $F(1, 232) = 0.321$, $p = 0.572$, $\eta^2 = 0.001$. In both learning conditions the attitude to dialogic learning decreased with time. No main effect of design or interaction between condition and design was observed.

A 2(condition) \times 2(design) MANCOVA for the three concepts of the attitude to dialogic learning and the three pretest measurements of these concepts as covariates showed an effect of condition, Pillai's trace = 0.038, $F(3, 227) = 0.298$, $p = 0.032$, $\eta^2 = 0.038$, a significant effect of design, Pillai's trace = 0.034, $F(3, 227) = 2.68$, $p = 0.048$, $\eta^2 = 0.034$, and no interaction between the independent variables. Subsequent univariate analyses revealed a main effect of condition only for Attitude to Sharing, $F(1, 235) = 4.78$, $p = 0.03$, $\eta^2 = 0.02$, and design, $F(1, 235) = 7.11$, $p = 0.008$, $\eta^2 = 0.03$. No interaction between condition and design was observed. In the unbalanced design – the one-class teacher group – and in the Non-Dialogic condition, scores on the posttest Attitude to Sharing were higher.

3.3.3. Generative fluency of reasoning and quality of value orientation

The 2(condition) \times 2(design) ANCOVA for GFR showed a significant main effect of condition, $F(1, 253) = 16.41$, $p = 0.000$, $\eta^2 = 0.06$: the Dialogic condition outperformed the Non-Dialogic condition. Condition also interacted with

Table 5

Mean (and standard deviations) for the two experimental conditions for subject-matter knowledge and attitudes to dialogic learning

Measure	Dialogic condition	Non-dialogic condition
Subject-matter knowledge test	0.60 (0.19)	0.59 (0.18)
Attitude to dialogic learning		
All items	2.25 (0.34)	2.27 (0.37)
Sharing	2.13 (0.54)	2.17 (0.60)
Co-construction	2.37 (0.42)	2.32 (0.44)
Legitimizing	2.26 (0.51)	2.32 (0.55)

design, $F(1, 253) = 7.072$, $p = 0.014$, $\eta^2 = 0.024$, indicating that the condition effect was larger in the one-teacher than in the two-teacher classes.

For QVO a $2(\text{condition}) \times 2(\text{design})$ multivariate analysis with QVO-A and QVO-P as dependent variables showed a condition effect, Pillai's trace = 0.05, $F(2, 249) = 6.8$, $p = 0.001$, $\eta^2 = 0.05$. Subsequent univariate analyses with the two QVO variables (QVO-A and QVO-P) revealed that the effect was observed in the antagonist tasks and was absent in the protagonist tasks, $F(1, 252) = 13.67$, $p = 0.000$, $\eta^2 = 0.052$ and $F(1, 252) = 0.175$, $p = 0.676$, $\eta^2 = 0.001$, respectively. The Dialogic condition outperformed the Non-Dialogic condition when students had to react on a given case: students referred more often to socially oriented values in their responses. This effect was absent when students initiated the argument, which might be a more demanding task.

For QVO-A in a $2(\text{condition}) \times 2(\text{design})$ univariate analysis a significant main effect of design was observed, $F(1, 252) = 5.47$, $p = 0.020$, $\eta^2 = 0.021$, indicating that the one-teacher classes scored significantly higher on QVO-A than the two-teacher classes. The interaction between the condition and design, tended to be significant, $F(1, 253) = 3.66$, $p = 0.057$, $\eta^2 = 0.014$, suggesting a larger condition effect in the one-teacher classes than in the two-teacher classes.

3.4. Testing and exploring interaction effects

We had hypothesized an interaction between aptitude and the learning condition, that is, an interaction between attitude and the learning condition. Students with a more positive attitude to dialogic learning would profit more from that condition than students with a less positive attitude, whereas this characteristic would not play a role in the non-dialogic learning. As regards the general academic aptitude we expected that the higher the aptitude, the more the student profits from the dialogic learning as a result of being involved in a double learning agenda: students had to acquire the skills to work dialogically, which is new for them, and to acquire subject-matter knowledge. The same explanation may hold for general reasoning skill.

3.4.1. Effect of general reasoning skill

A main effect of the general reasoning skill was observed on generative fluency of reasoning in the Dialogic condition, $F(1, 242) = 8.321$, $p = 0.004$, $\eta^2 = 0.033$, $\beta = 6.326$, $t = 2.288$, $p = 0.004$, but not in the Non-Dialogic condition. This indicates that students with a higher level of general reasoning skill profited more from the Dialogic condition than students with relatively low scores on the General Reasoning test, while in the Non-Dialogic condition no effect of general reasoning skill was observed (see Fig. 1A).

No significant effect of general reasoning skill was observed on QVO, although trends pointed to potential effects (on antagonist tasks in Dialogic condition, $F(1, 242) = 3.467$, $p = 0.064$, $\eta^2 = 0.014$).

General reasoning skills contributed negatively to the effect of condition on attitude to dialogic learning in the Non-Dialogic condition, $F(1, 239) = 4.069$, $p = 0.045$, $\eta^2 = 0.017$, $\beta = -0.384$; the better the students scored on general reasoning skill tests, the less they profited from the Non-Dialogic condition as regards the attitude to dialogic learning (see Fig. 1B).

3.4.2. Effect of pretest attitude to dialogic learning

Pretest attitude scores in both learning conditions affected the resulting attitude score in the posttest (see Fig. 1C). In the Dialogic condition the effect was larger, $F(1, 235) = 145.078$, $p = 0.000$, $\eta^2 = 0.385$, $\beta = 2.352$, than in the Non-Dialogic condition, $F(1, 235) = 87.614$, $p = 0.000$, $\eta^2 = 0.274$, $\beta = 1.903$.

3.5. Effects of implementation

Although the implementation measures — collected via classroom observations (time sampled time-on-task observations, randomly chosen learners) — are not independent, for exploratory reasons the classroom observation data on the students were disaggregated, so that we connect them to the individually measured dependent variables. In this way we would build an impression of the effect of the various learning activities on the dependent variables within the two experimental conditions — not more than that.

For individual learning activities, no differential effect of the relative number of these activities in the two conditions was observed on GFR.

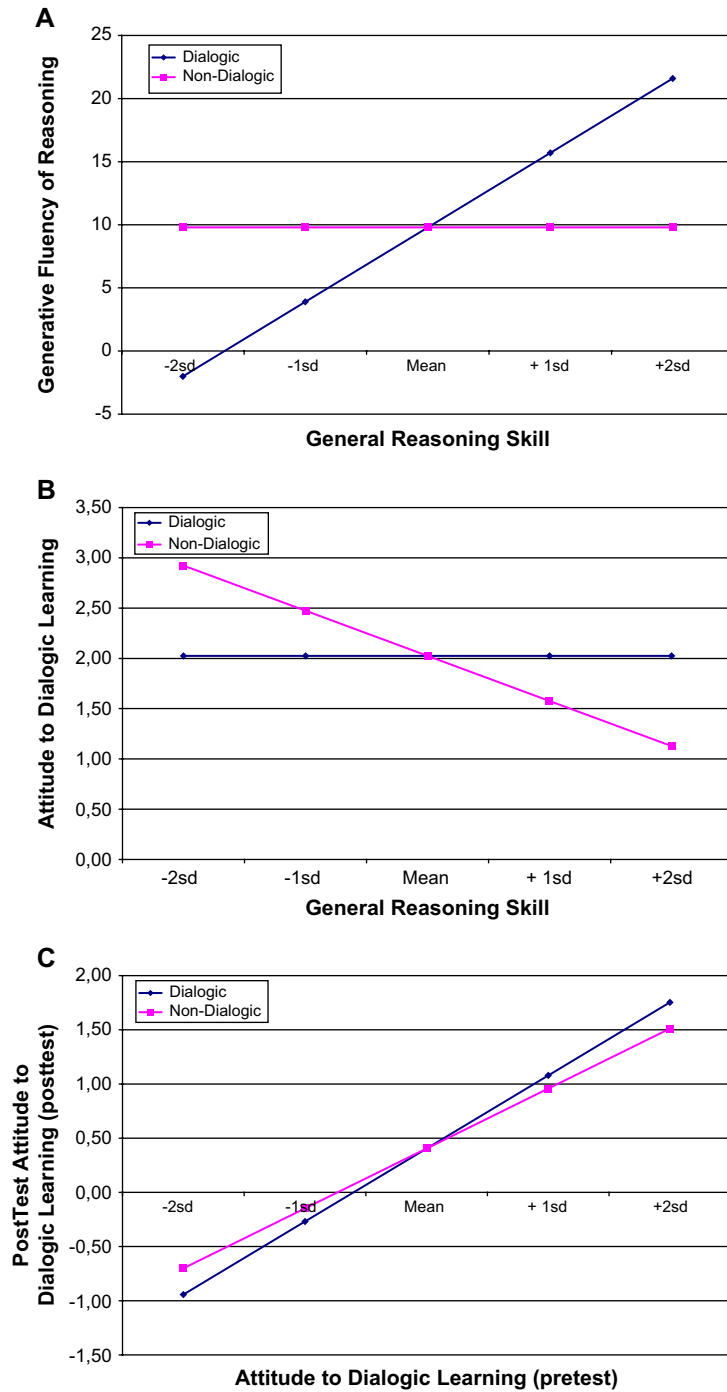


Fig. 1. (A) Interaction between learning conditions and student characteristics. (B) Contribution of student’s level of general reasoning skill on attitudes to dialogic learning. (C) Contribution of student characteristics to conditions: Attitude to DIALOG.

For dialogic learning activities, an effect of implementation was observed in the Dialogic condition on GFR, $F(1, 252) = 5.651, p = 0.018, \eta^2 = 0.022, \beta = 0.381$, and on QVO-P, $F(1, 251) = 13.188, p = 0.000, \eta^2 = 0.050, \beta = 0.587$. No effect was observed in the Non-Dialogic condition (see Fig. 2A and B). This suggests that in a classroom with a relatively high level of dialogic behaviour, students scored better on GFR and on QVO-A than in a classroom

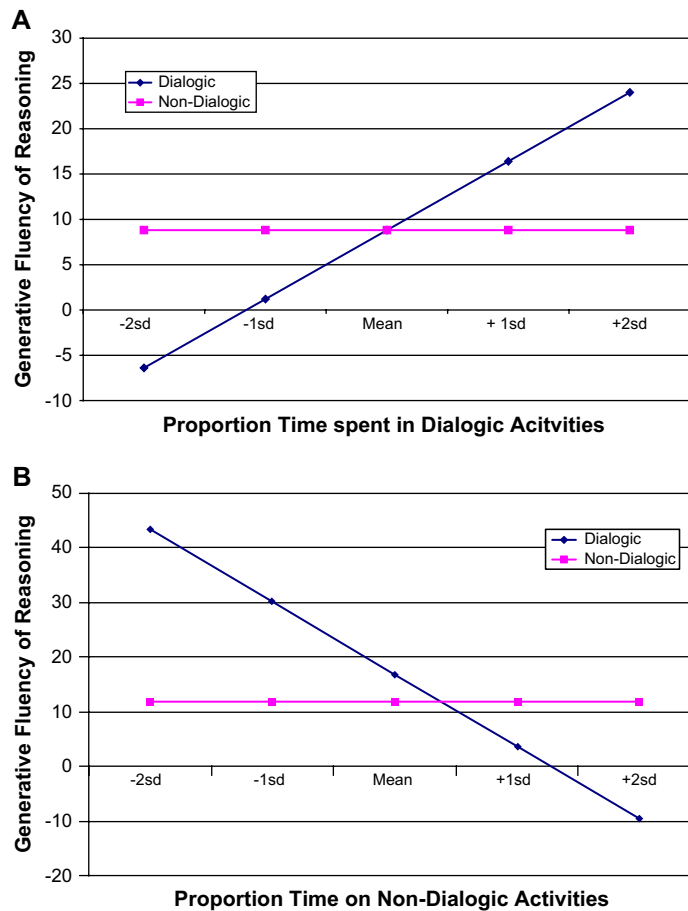


Fig. 2. (A) Effect of time spent on dialogue activities within conditions on generative fluency of reasoning. (B) Time-on-task effects within conditions.

with a relatively low level of dialogic behaviour. This effect was not related to the Design factor (balanced vs. unbalanced).

4. Conclusions

In our view teaching critical thinking should focus on the competences of students to argue in a logical and substantial way, as well as on the competences of collaborative meaning making and legitimating opinions with the support of values such as ‘respect’ and ‘justice’, that is, reasoning within a normative context. We hypothesized that a dialogic approach to learning and teaching would have a more positive effect on students’ critical-thinking competences than a non-dialogic approach, both in terms of generative fluency of reasoning and of quality of value orientation. At the same time we expected that these learning environments had no differential effects on the subject-matter knowledge of Biology that had to be acquired. These hypotheses were confirmed. First and foremost, students in the Dialogic condition scored higher on generative fluency of reasoning than students in the Non-Dialogic condition. They produced more positions when responding to the position formulated by a peer in the stimulus text, more supportive arguments, and more coordinated and subordinated arguments to support the measures they proposed in the two protagonist’s tasks.

Second, there were significantly more value orientations in the reasoning of students in the Dialogic condition. When taking a position on an issue, these students referred more often than students in the Non-Dialogic condition to social justice as a framework of value orientations. However, this effect was only observed in the antagonist tasks,

not in the protagonist tasks. As we argued, this may be due to the more difficult nature of the protagonist tasks (Kellogg, 1999).

We had also hypothesized that the Dialogic condition would result in more positive attitudes to dialogic learning. This effect was not observed when applying analyses of variance. Instead, the attitude to dialogic learning decreased significantly in both conditions at the posttest. Thus, we may conclude that the two instructional methods did not affect students' attitudes to dialogic learning in a differential way. It seems that the attitude to dialogic learning is a rather stable construct, as shown by the rather high correlations over time. However, when we allowed different regression weights of pretest attitudes for each condition, we observed that the Dialogic condition scored significantly higher on the posttest than the Non-Dialogic condition. This implies that in the Dialogic condition the decrease over time was less than in the Non-Dialogic condition. Yet it is remarkable that attitude to dialogic learning in both conditions decreased.

Finally, we had expected that the differences in instructional design between the conditions would not affect the learning outcomes regarding the subject matter. Indeed, no effects of the conditions on subject knowledge were observed. This implies that the attention paid to dialogic activities in the Dialogic condition resulted in a better performance on value orientation and reasoning, without a counter effect on subject-matter knowledge. From secondary analyses it appeared that the intended learning activity contributed to the subject-matter knowledge score. In the Dialogic condition, the effect of individual work was significant, as was the effect of dialogic work. This implies that in the Dialogic condition, the more individual work was done, the lower was the effect on subject-matter knowledge; on the contrary, the more Dialogic work was implemented, the better were the results on subject-matter knowledge. In the Non-Dialogic condition, no such relations between learning activities and learning results were observed.

Research often reveals a differential impact of the characteristics of a specific instructional design on the learning processes and the learning results of different groups of students. Our study also showed interactions between student characteristics and learning condition: the better the students' general reasoning skills, the more they profited from the Dialogic condition on the generative fluency of reasoning, whereas in the Non-Dialogic condition, no relation between this student feature and condition effect was observed. This finding suggests that for students with a relatively low level of general reasoning skill, a Non-Dialogic condition is a better option than a Dialogic condition. For students with a relatively high level of reasoning skill, a Dialogic condition is more beneficial than a Non-Dialogic condition. This finding may relate to features of the experimental condition. The lesson series were new for students in two respects: in content — value-loaded critical thinking — and the instruction format — dialogue learning. As a consequence, the experimental condition did a relatively strong appeal on the cognitive capacity of students favouring the more apt students.

Construction of effective curricula is one step, implementing these curricula is a second step. We have been able to determine that the quality of implementation of the experimental lesson series developed for critical thinking was satisfactory. Observations during the lesson series showed that significantly more time is spent on dialogic learning activities in the Dialogic condition than in the control condition.

Finally we would like to emphasize that both conditions were experimental from an educational point of view and a practical point of view. Although focusing on value-loaded critical thinking in academic subjects as an aspect of citizenship education is required in the national attainment objectives, it is rarely implemented in Dutch textbooks. In this study, we constructed two conditions focusing on position taking and opinion building. The conditions only differed in the main learning activity implemented: Dialogic vs. Non-Dialogic (individual work). From evaluative comments in debriefing sessions it became clear that the teachers experienced the Non-Dialogic condition to be just as 'experimental' as the Dialogic condition, which indicates that critical thinking, especially value-loaded critical thinking, is practiced infrequently in the Biology curriculum.

For science education research, the present study shows how the constructivist view can be given form in a science curriculum: science teaching can contribute to citizenship education if the following instructional elements are incorporated in the lessons and are not at the expense of the content knowledge. The learning content should refer to authentic problems in the world, to students' own positions and those of others, and to students' opportunities to influence these positions. Moreover, students should be able to involve themselves personally in the process of collaborative meaning making (co-construction). Last but not least, students should be encouraged to participate in value-loaded discourse on socioscientific issues. The present study may lead to a better understanding how science teaching can help students to participate in a democratic society (Sadler & Zeidler, 2005) and reflect on social values when decisions have to be made (Driver et al., 2000).

Appendix A. To give an impression of the lesson material used in the two different conditions, an example is given below of an assignment implemented in two different ways

Example 1: Dialogic condition

In the Dialogic condition the assignment starts with a simulated dialogue between two students about recycling computers in developing countries. In consequent tasks students working in pairs analyze the opinions of peers in the simulated dialogue. Then they formulate their own, *joint*, opinion.

Soeraya and Glenn are watching TV. They see how old computers from Europe are dismantled in India. The different materials are sorted – plastic with plastic, etc. The company employs children too.

Glenn says: “Now that’s a good example of recycling! Our old computers are then re-used.”

Soeraya: Why do they first take those computers all the way to India? Surely, that must cost *lots* of money?

Glenn: No, that’s good. Here a worker costs 40 euros a day, in India only 1 euro. That saves a lot of money. Yeah, good thinking, eh? It’s a really good way of recycling, otherwise we’d have to burn the things or something. Recycling in the Netherlands costs far too much money.

Soeraya: But you can see that they’re children doing the work in India.

Glenn: But those people are really poor, now the children are earning a bit extra too. And the computers are recycled for us. So we’re helping India a bit.

Soeraya: But children, Glenn! It’s terrible to pay children a pittance to take our computers apart. It’s not *fair* to recoup our recycling costs on those poor children.

Glenn: H’m, no? Why not?

Soeraya: I don’t think it’s fair. You wouldn’t dream of setting children of six years old to work in the Netherlands

Glenn: Not fair? But those children wouldn’t earn anything otherwise.

Soeraya: If you want those children to earn something, then you have to make sure that their parents have a good job. They need to teach them how to make computers. And build computer factories there. Then the children could just go to school. Everybody should clear up their own rubbish!

Glenn: Yeah, I guess you’re right. If they built computers there, they wouldn’t be so expensive! In India everything costs less.

Soeraya: That’s right! And then we could look Indian children in the eye. Show a bit of *respect*, that wouldn’t do any harm.

Glenn hadn’t thought about it like this before. He was just so pleased that the computers were being recycled. Things are sometimes more complicated than we think. Soeraya explained to Glenn why she thinks it’s not fair to recycle our computers in India.

Assignment

1. Recycling computers costs money. What are Glenn and Soeraya’s opinions and arguments and what are yours?

Glenn	Opinions and arguments
Soeraya

	Our opinions and arguments about the recycling of computers
I’m working with

2. Now you know that recycling, for example computers, can also have disadvantages.
 We think that a disadvantage of recycling computers is:, because
 In our opinion the best solution for old computers is:

Example 2: Non-Dialogic condition

In the Non-Dialogic condition the assignment includes the same information as in the Dialogic condition. The content is not presented in a simulated dialogue but in a coherent text, as is usual in textbooks. Students analyze the content of the text individually and then formulate their own opinion.

A good example of recycling?

Every year lots of computers are thrown away. Some are broken, sometimes the model is out-of-date. What happens to these thousands of old computers? They're collected and dismantled. The different materials are sorted – metal with metal, plastic with plastic, etc. These materials are made really small and melted down, ready to use again. The materials from these old computers are re-used to make new computers. An example of recycling. But did you know that this recycling of computers is done in India! Our computers are taken apart in India. In Europe a worker costs 40 euros a day and in India only 1 euro. That saves a lot of money. Recycling old computers in India is therefore much cheaper than here in Europe. That's a great advantage.

But there are also disadvantages to recycling computers in India. Transporting the computers all the way to India does of course use a lot of energy.

Children recycle our computers. More than 200 million (200,000,000) children in the world have to work every day. We call this child labor. In India at least 50 million children work, not just doing a paper round but more than 10 hours a day, every day. So they cannot go to school. If these children don't work, their families have far too little money to live on. Thus they have no choice! You could ask yourself whether it's *fair* that children take our computers apart. Shouldn't we take care of our own old things?

Assignment

1. What is the best solution for all our old computers in your opinion?
 I think the best solution for our old computers is, because
2. What do you think about recycling our old computers in India?
 I think recycling our old computers, because
3. What do you think about child labor?
 I think child labor is, because

References

- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Erlbaum.
- Brown, A. (1997). Transforming schools into communities of thinking and learning about serious matters. *American Psychologist*, 52, 399–413.
- Burbules, N. C., & Berk, R. (1999). Critical thinking and critical pedagogy: relations, differences, and limits. In T. S. Popkewitz, & L. Fendler (Eds.), *Critical theories in education* (pp. 45–65). New York: Routledge.
- Dekker, R., Elshout-Mohr, M., & Wood, T. (2004). Working together on assignments: multiple analysis of learning events. In J. van der Linden, & P. Renshaw (Eds.), *Dialogic learning* (pp. 145–170). Dordrecht, The Netherlands: Kluwer.
- Dennick, R. G., & Exley, K. (1998). Teaching and learning in groups and teams. *Biochemical Education*, 26, 111–115.
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84, 278–312.
- Garside, C. (1996). Look who's talking: a comparison of lecture and group discussion teaching strategies in developing critical thinking strategies. *Communication Education*, 45, 212–227.
- Giroux, H. A. (1994). Toward a pedagogy of critical thinking. In K. S. Walters (Ed.), *Re-thinking reason: New perspectives in critical thinking*. Albany, NY: SUNY.
- Halpern, D. F. (2003). *Thought and knowledge: An introduction to critical thinking* (4th ed.). Mahwah, NJ: Erlbaum.
- Haworth, A. (1999). Bakhtin in the classroom: what constitutes a dialogic text? Some lessons from small group interaction. *Language and Education*, 13(2), 99–117.
- Johnson, D. W., & Johnson, R. T. (1993). Creative and critical thinking through academic controversy. *American Behavior Scientist*, 37, 40–53.

- Kellogg, R. T. (1999). *The psychology of writing*. New York: Oxford University Press.
- Kerr, D. (1999). Changing the political culture: the Advisory Group on Education for Citizenship and the teaching of democracy in school. *Oxford Review of Education*, 25(1/2), 274–285.
- Klaczynski, P. A. (2001). Analytic and heuristic processing influences on adolescent reasoning and decision making. *Child Development*, 72, 844–861.
- Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development*, 74(5), 1245–1260.
- McNeil, J. D. (1996). *Curriculum. A comprehensive introduction* (5th ed.). New York: Harper Collins.
- Meijer, J., Elshout-Mohr, M., & Van Hout-Wolters, B. H. A. M. (2001). An instrument for the assessment of cross-curricular skills. *Educational Research and Evaluation*, 7(1), 79–107.
- Mercer, N. (2000). *Words and minds: How we use language to think together*. London: Routledge.
- Noddings, N. (1992). *The challenge to care in schools: An alternative to education*. New York: Teachers College Press.
- O'Donnell, A. M., & King, A. (Eds.). (1999). *Cognitive perspectives on peer learning*. Mahwah, NJ: Erlbaum.
- Onderwijsraad. (2003). *Onderwijs en burgerschap. Advies [Education and citizenship. Advice]*. Den Haag: Onderwijsraad.
- Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, 41(10), 994–1020.
- Paul, R. C. (1992). *Critical thinking: What every person needs to survive in a rapidly changing world* (2nd revised ed.). Santa Rosa, CA: Foundation for Critical Thinking.
- Paul, R. C. (1994). Teaching critical thinking in the strong sense. In K. S. Walters (Ed.), *Re-thinking reason: New perspectives in critical thinking* (pp. 181–198). Albany, NY: SUNY.
- Renshaw, P. (2004). Dialogic learning, teaching and instruction: theoretical roots and analytical frameworks. In J. van der Linden, & P. Renshaw (Eds.), *Dialogic learning* (pp. 1–15). Dordrecht, The Netherlands: Kluwer.
- Rychen, D. S., & Salganik, L. H. (2003). *Key competencies for a successful life and well-functioning society*. Göttingen, Germany: Hogrefe & Huber.
- 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(1), 112–138.
- Salomon, G., & Perkins, D. N. (1998). Individual and social aspects of learning. In P. D. Pearson, & A. Iran-Nejad (Eds.), *Review of research in education* (pp. 1–24). Washington, DC: American Educational Research Association.
- Schuitema, J., ten Dam, G. T. M., & Veugelers, W. (in press). Teaching strategies for moral education: a review. *Journal of Curriculum Studies*.
- Siegel, H. (1992). The generalizability of critical thinking skills, dispositions and epistemology. In S. P. Norris (Ed.), *The generalizability of critical thinking* (pp. 97–108). New York: Teachers College.
- Solomon, D., Watson, M., & Battistich, V. (2001). Teaching and schooling effects on moral/prosocial development. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 566–603). Washington, DC: American Educational Research Association.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: implications for the rationality debate? *Behavioral and Brain Sciences*, 23, 645–665.
- Stark, R., Mandl, H., Gruber, H., & Renkl, A. (1999). Instructional means to overcome transfer problems in the domain of economics: empirical studies. *International Journal of Educational Research*, 31, 591–609.
- Ten Dam, G., & Volman, M. (2004). Critical thinking as a citizenship competence: teaching strategies. *Learning & Instruction*, 14(4), 359–379.
- Ten Dam, G., Volman, M., & Wardekker, W. (2004). Making sense through participation: social differences in learning and identity development. In J. van der Linden, & P. Renshaw (Eds.), *Dialogic learning* (pp. 63–85). Dordrecht, The Netherlands: Kluwer.
- Terenzini, P. T., Springer, L., Pascarella, E. T., & Nora, A. (1995). Influences affecting the development of students' critical thinking skills. *Research in Higher Education*, 36, 23–39.
- Toplak, M. E., & Stanovich, K. E. (2002). The domain specificity and generality of disjunctive reasoning: searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, 94, 197–209.
- Torff, B. (2005). Developmental changes in teachers' beliefs about critical-thinking activities. *Journal of Educational Psychology*, 97(1), 13–22.
- Tsui, L. (1999). Courses and instruction affecting critical thinking. *Research in Higher Education*, 40, 185–200.
- Tynjälä, P. (1998). Traditional studying for examination versus constructivist learning tasks: do learning outcomes differ? *Studies in Higher Education*, 21, 173–189.
- Van Emmeren, F. H., Grootendorst, R., Snoeck Henkemans, F., Blair, J. A., Johnson, R. H., & Krabbe, E. C. W., et al. (Eds.). (1996). *Fundamentals of argumentation theory. A handbook of historical backgrounds and contemporary developments*. Mahwah, NJ: Erlbaum.
- Veugelers, W. (2000). Different ways of teaching values. *Educational Review*, 52, 37–46.
- Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: an empirical investigation of a possible sociocultural model of cognitive development. *Learning and Instruction*, 9(6), 493–516.
- Wells, G. (2000). Dialogic inquiry in education. Building on the legacy of Vygotsky. In C. D. Lee, & P. Smagorinsky (Eds.), *Vygotskian perspectives on literacy research. Constructing meaning through collaborative inquiry* (pp. 51–85). Cambridge, UK: Cambridge University Press.
- Yore, L., Bisanz, G. L., & Hand, B. (2003). Examining the literacy component of science literacy: 25 years of language arts and science research. *International Journal of Science Education*, 25(6), 687–725.
- Zohar, A., Degani, A., & Vaakin, E. (2001). Teachers' beliefs about low-achieving students and higher-order thinking. *Teaching and Teacher Education*, 17, 469–485.
- Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39(1), 35–62.