Collaborative Science Lessons—Learning and Argumentation in an Interdisciplinary Virtual Laboratory

MARKO TELENIUS, EIJA YLI-PANULA, AIJA AHTINEVA JA MARJA VAURAS

mjtele@utu.fi
Turun yliopisto, opettajakoulutuslaitos

Abstract

This study investigates upper secondary school students' argumentation and interdisciplinary integration of biology and chemistry in a virtual environment. Pre and post-tests (content knowledge, scientific reasoning) were carried out. Students (n=35) collaborated in small groups to generate hypotheses, study plans and presentations which were analysed by SOLO taxonomy and videotaped spoken products were presented. Results of this qualitative study showed that the students were able to generate hypotheses despite the non-supportive curriculum and were able to integrate biological and chemical concepts and spontaneously argue in virtual environment. The results are discussed with respect to collaborative argumentation and interdisciplinary integration.

Keywords

Argumentation, interdisciplinary learning, students, virtual laboratory

Introduction

This study aimed to investigate how upper secondary school students integrate knowledge from two separate disciplines (biology and chemistry) and spontaneously argue for or against content issues through collaborative interdisciplinary learning. The virtual learning programs are usually designed to individually study issues of one discipline. However, in this study the virtual environment was designed for collaborative working to enhance students´ learning, e.g. collaborative thinking and multi-perspective taking in small groups across two disciplines.

According to Taber (2008) to be able to connect new material to existing knowledge and to evaluate new theories, it is essential that students are able to *integrate knowledge* across disciplines. To see how biology and chemistry contents are integrated to one another, students need to understand how key concepts of the disciplines are interrelated and how they are built on in other disciplines among science (Simon, 2008). Also reasoning skills are required in order to understand interrelated scientific concepts and theories (Lawson, 2004). Learning key topics and key concepts from the perspectives of different disciplines is more inspiring, and highlights science's conceptual integration (Taber, 2015).

Collaboration helps students to learn from and about *scientific argumentation* and to make their own thinking visible. Using appropriate evidence and reasoning helps students to complete inquiry practices; and as evaluating scientific arguments has proved difficult for students, they should be encouraged to collaborate with each other on tasks requiring the generation of a scientific explanation (Sampson & Clark, 2008). To strengthen the requirement for collaboration and scientific argumentation, and because the conditions for science learning improve through integration across different disciplines (Duschl, 2008), the virtual laboratory environments are interdisciplinary in nature.

As it is difficult to develop a curriculum that reconciles the requirements of context-based education with the reality of the classroom (Knutsson, Blåsjö, Hållsten & Karlström, 2012; Pilot & Bulte, 2006), collaboration and virtual laboratories may serve to address this issue.

Spontaneous argumentation in the context of a virtual laboratory

In general, argumentation is an attempt to establish truth, containing a claim supported by data exposed to rebuttals or counter-arguments (Osborne, 2010;

Aydeniz, Pabuccu, Cetin & Kaya, 2012; Robertson & Shaffer, 2014). In the educational literature, argument has been defined in either “rhetorical” (individual) or “dialogical” (social) terms. While the first category emphasises argument as a tool to persuade other people to adopt one’s way of thinking, the second focuses on consideration of alternative positions to reach agreement, either as an individual or within a social group (Driver, Newton & Osborne, 2000). These forms of argument are closely related and partially overlapping, since the skills that characterise social argumentation are also entailed in rhetorical argument (Kuhn, 1993). The present study leans on the “dialogical” or “multivoiced” interpretation of argument, as students generated their arguments within small groups. Here, however, the term *spontaneous argumentation* is used to highlight the fact that students were not specifically instructed in argumentation, and the digital tool did not explicitly support this type of learning. Instead, students supplied evidence in support of their claims without any prompts from the teacher or from the digital tool. A spoken product is a reasoned piece of discursive discussion in which a claim has been justified (Berland & McNeill, 2010; Kuhn, 2010).

Argumentation has been included in science education to meet two goals: acquisition of scientific concepts and engagement in scientific discourse (Kuhn, 2010; Bricker & Bell, 2008). This trend is informed by the notion that the learner develops new understanding through a cognitive process of comparing old models (i.e. those they already have) against new models, which they are given. It follows that argumentation—the skill to judge why ideas are wrong or right—plays a crucial part both in learning to think and in developing new understandings. (Osborne, 2010). Argumentation in science education is referred to as scientific argumentation and contributes to learning goals by helping students to understand, use and generate scientific explanations and arguments (Duschl & Osborne, 2002).

Role of argumentation in promoting content knowledge (CK), scientific reasoning (SR) and collaborative learning in a virtual laboratory

In collaborative study settings, learners can engage in constructive peer interactions that help them to move towards high-level understanding (Zhang, Scardamalia, Lamon, Messina & Reeve, 2007). In such situations, collaborative or “dialectical” argumentation offers a means of augmenting both motivation and CK (Chinn & Clark, 2013). There is also evidence that inquiry learning can be applied in the context of computer-supported collaborative learning, while students work together towards a deeper understanding (Sampson & Clark, 2008). In the present study, *collaborative inquiry learning* refers to students' joint intellectual efforts when students working in groups of two or three struggle to understand the topic, reason about the environmental situation and follow scientific procedures to conduct experiments and create outcomes in a virtual laboratory (Hodson, 2014). Students´ collaboration may importantly support engagement and interest in science subjects.

Argumentation also supports engagement in science learning (Driver, Newton & Osborne, 2000). In general, there are two theoretical frameworks that explain the benefits of argumentation: cognitive accounts and sociocultural accounts. The cognitive account focuses on individual learning and explains learning in terms of improved individual representations. The sociocultural account considers learning through argumentation as adoption of social practices (Ryu & Sandoval, 2015). The present study focuses on the social level of a group activity in which learning takes place and is therefore a sociocultural account. In this study, sociocultural learning is understood as the need to think together, to make thinking processes clear, and to use scaffolds provided by the teacher or peer students. This ability to discuss visibly of one's knowledge is the pathway to argumentation.

When engaging in argumentation, students engage in a range of explicit elaborative processes, which are known to promote learning (Chinn & Clark, 2013; Aufschnaiter et al., 2008). Additionally, during the process of argumentation, students learn from their peers; the process of providing evidence for claims gives learners a better understanding of what they are learning and requires them to articulate reasons in support of the claims they develop (Chinn, O´Donell & Jinks, 2000; Osborne, 2010).

Research Questions (RQ)

The following research questions underpinning the study were:

> RQ1. What was the level of group learning outcomes in terms of collaboratively constructed hypotheses, writing up of study plans and written presentations?
>
> RQ2. What was the level of students´ scientific reasoning and their content knowledge of biology and chemistry in pre and post-tests?

RQ3. How did integration of biology and chemistry arise in collaboratively constructed written text and spontaneous argumentation in the videotape excerpts in the context of a virtual laboratory?

Issues studied in RQ1 formed the basis for selecting two groups of students for further studies. A further issue was whether students in the two selected groups (which differed in their outcomes) had different skills and knowledge profiles in relation to SR and disciplinary CK within and between the two groups (RQ2). SR was, however, tested only as a pre-test. The RQ3 focused on interdisciplinary integration and spontaneous argumentation of the two selected groups.

Research Design and Methodology

Study background and participants

This is a qualitative case study describing the phenomenon "interdisciplinary argumentation" in collaboratively supported virtual environment using mixed methods (Cohen, Manion & Morrison, 2011). The present data was collected from two Finnish rural schools (H and P, three different classes). The participants were upper secondary school students aged 16–17 (n = 35; 16 boys, 19 girls) and four teachers. Two of the teachers taught biology (BI), and two taught chemistry (CH). To support consistency of experiences, teachers from both schools were instructed separately and provided with the same information and guidelines. The students' tasks were organized according to the National School Curriculum for upper secondary schools (FNBE, 2003), each involving an optional course in either BI or CH. The students were accustomed to using computers and internet-based resources, as they used laptops during their normal lessons. However, this was the first time they had studied in a virtual laboratory environment or used the scientific approach (i.e. an inquiry-based study procedure including hypotheses and virtual data gathering) in integrated BI and CH courses.

The teachers divided students into pairs or threes. To enhance collaborative practices (Vauras & Volet, 2013), each pair or small group worked on a shared laptop, on which the virtual laboratory was installed. The teachers formed groups to ensure favourable group dynamics, and that at least one student in each group possessed good English language skills. In total, there were 13 small groups: nine groups of three and four groups of two students. From here on, the term *study* *group* refers to all students and *group* will be used to refer to all small groups, irrespective of the number of members. Integration skills of two disciplines and their spontaneous argumentation were chosen to be analysed from two groups H4 and P3.

Measurements, procedures and virtual studies

Virtual environment

The virtual laboratory used in this study was the Virtual Marine Scientist (VMS) (Fauville, 2013) <https://ipkl.gu.se/english/Research/research_projects/vms>. This software was designed to encourage collaborative learning and problem solving and, in particular, to introduce the way scientists conduct research to students. For present purposes, virtual studies focusing on the impact of ocean acidification on mussels' and sea stars' vitality and growth speed provided the interdisciplinary context. The studies in VMS were all in English and all students' actions were videotaped.

Pre and post-tests

The study (Table 1) included three separate tests and began with *pre-tests* (in Finnish), measuring biology content knowledge (BI-CK), chemistry content knowledge (CH-CK) and SR. The students were gathered in an IT classroom and they completed the tests there. At the end right after their PowerPoint presentations, students were asked to complete the *post-test*s, which involved only BI-CK and CH-CK, but no SR test due to the lack of time.

Table 1. Structure of the case study (VMS= Virtual Marine Scientist laboratory)

The issue and time	The tasks
Pre-tests (no time limit)	* online biology and chemistry content tests (BI-CK, CH-CK) * the classroom test of scientific reasoning (SR) * prior to the VMS based lessons without time limitations.
VMS lessons 1–2 (2 x 90 min)	* key theories and logic of the experiments studied using the virtual book with videos and passages * generation of the study plans and grant applications to get access to the virtual laboratory * the virtual experiments started if time available
VMS lesson 3 (1 x 90 min)	* the virtual experiments were finalised * results analysed and interpreted * preparation of PowerPoint presentations started
Presentations	* presentations were finalised using the structure of a scientific article, and presented to the class
Post-tests (no time limit)	* online BI-CK and CH-CK tests * right after the presentations without time limitations

The BI-CK test consisted of seven multiple choice (i.e. choose the completely correct answer of these three given ones) and nine open questions. The CH-CK test consisted of 14 multiple choice and no open questions. The pre and post-tests in CK followed the same design, assessing students' understanding of basic BI and CH concepts. The content of questions related either directly or closely to the themes of the virtual laboratory: pH, ocean acidification, climate change, and population dynamics. Both CK tests measured recall and interpretation of graphically presented data. These CK tests were designed by two BI and two CH university teachers and were scrutinized by the research team. Along with the VMS, the tests were piloted (Yli-Panula, Hiilovaara & Vauras, 2015) in one school by International Baccalaureate (IB) students prior to the study. Based on that pilot study, some items were slightly modified. In addition to scientific content, students' SR was also assessed prior to group work, using an existing questionnaire Lawson's Classroom Test of Scientific Reasoning (previously validated; Lawson, 2000). This classroom test was used for categorization. All tests were converted for each student into electronic format using the ViLLE platform (http://villeteam.fi/en). Students completed the tests with no time limits.

Data analyses

Flow of Data Analysis

The data included students' VMS lessons on ocean acidification, their PowerPoint presentations of study outcomes, videotaped student classroom collaborations, pre and post-tests (Table 2).

Table 2. Flow of the data analysis

The issue and time	**The tasks**
Marking the pre- and post-tests	* test were graded based on criteria * the criteria for the tests set by the research team established by three researchers * the criteria for the ready-made tests from literature * all the scores checked by two researchers * the relation between students' scientific reasoning and content knowledge was analysed
Grading the PowerPoint presentations:	* the grading scale for the presentations was established by two researchers independently (inter-reliability 98 %) * inductive content-based analysis and SOLO-taxonomy were used
The group selection	* best and average group were chosen for precise, more detailed analysis * two group selection based on level of the presentations and the quality of videotapes (how data was documented in the videos)
The video clip selection and video analysis	* analyses of the videotapes of the two groups * the extracts of students discussing their hypotheses and result interpretations edited * the best and worst argumentative sessions were chosen and analysed
Grading the hypotheses – experimental plans (cf. table 1)	* analyses of the hypotheses and experimental plans (using the equal SOLO taxanomy scale cf. presentations) * only levels 3–5 for accepted grant applications; declined applications were not graded * the grading conducted by three researchers independently (inter-reliability 98 %)
Integration – spontaneous argumentation analyses	* the integration of biology and chemistry analyses from content test questions, hypotheses, experimental plans and video clips * scientific argumentation analyses from hypotheses and experimental plan and video clips

To answer the three research questions, data were obtained from assessments of the following three data sets: 1) collaboratively written virtual study plans, including hypotheses; 2) PowerPoint presentations detailing groups' procedure; and 3) videotaped group work. To illustrate collaboration and argumentation while learning in the VMS, two groups (H4 and P3) of the 13 were selected on the basis of their distinct collective outcomes in terms of final presentations and technical quality of video material. The high-performing group (H4) scored 5, and the average-performing group (P3) scored 3. For present purposes, integration of BI-CK and CH-CK and spontaneous argumentation were analysed for these two groups. All students' outcomes in VMS (collaboratively written study plans, hypotheses and PowerPoint presentations) were assessed using content analysis (Tuomi & Sarajärvi, 2011). The three researchers analysed the outcomes independently; agreement rate was 98%, and disagreements were resolved through discussion.

Analysis of the pre and post-test

Scientific CK tests were scored as follows. Multiple-choice items scored 0–2 (0 completely wrong, 1 partially right and 2 completely correct answer). Open questions were assessed by the first two authors independently on a scale from 0–3 (1 point per correct issue; questions asked for three concepts). Any disagreements were resolved in discussion. The maximum possible score for the BI-CK was 32; for the CH-CK it was 28; for SR test (Lawson, 2000) it was 28. For the study, all students were divided into two groups (low = < 15 points and high = ≥ 15 points), based on SR skills results. The paired t-test was carried out to find if there is a significant difference between the means of the high and low groups regarding the BI-CK and CH-CK scores in pre and post-tests.

Analysis of written group outcomes

In generating the categories for student groups' final presentations, a six-level scale was introduced. The scale adopts the five categories of the Structure of Observed Learning Outcomes (SOLO) taxonomy (Biggs & Collis, 1982; Chan, Tsui, Chan & Hong, 2002), with the addition of one further category for completely unstructured presentations (referred to Level 0, "unstructural"; Table 3). The same scale was used to assess the accepted hypotheses and study plans; for these, only levels 3–5 were used (Tables 3 and 6). Integration of BI and CH was considered a crucial element of the grading system for hypotheses and study plans.

Table 3. Grading scale for written hypotheses, study plans and final presentations

Level	SOLO Category	Description
0	*"Unstructural"*	*the purpose of the work (PW) understood incorrectly or not understood *language and colloquial phrases or wordings used *the course of a scientific experiment not perceptible and reasoning illogical.
1	*Prestructural*	The PW faultily understood *simple language mostly used *the course of a scientific experiment and reasoning vaguely perceptible *the presentation follows the instructions without understanding of the material *the narration mostly descriptive.
2	*Unistructural*	The PW partly faultily understood *language of the presentation mostly following the traditions of scientific language *the course of a scientific experiment and reasoning is evident *a personal consideration perceptible.
3	*Multistructural*	*the PW is understood *scientific language is used *the course of a scientific experiment followable, and deduction shows only minor groping *a personal consideration sometimes shown but not justified.
4	*Relational*	*the PW understood well *scientific language used *the course of a scientific experiment well perceptible, and deduction shows no gaps *the presentation firmly based on the instructions *the meaning of the titles understood and their contents deepened to some extent *a personal consideration plentiful and justified *integration between biology and chemistry occurs occasionally and on a very superficial level.
5	*Extended abstract*	*the PW is thoroughly understood *scientific language used appropriately *the course of a scientific experiment clearly perceptible and reasoning logical *the given draft extended and the titles describe the contents well *consideration profound and justified *he presentation shows great personal consideration *rich Integration between biology and chemistry

Analysis of group discussions

The group discussions for H4 and P3 were analysed by the first three authors from the videotaped lessons. For these analyses, two videotaped excerpts were independently selected by other project team members on the basis of…. In these excerpts, the groups discussed their study plans and results regarding the virtual experiments. For the preliminary analysis, the quality of each group's spontaneous argumentation was rated as either high or low. Argumentative discussion was rated as high if the students combined BI and CH in their argumentation and provided clear evidence in support of their claims; as low if no interdisciplinary discussion occurred and no evidence or only very weak evidence to support their claims was provided. Finally, three examples have been quoted to illustrate students´ spontaneous argumentation and integration.

Results

Collaboratively constructed written hypotheses, study plans and presentations

There was relative consistency among the group learning outcomes in terms of collaboratively constructed written hypotheses, writing up of study plans and written presentation (cf. RQ1) with the exception of group H1 (Table 4). The outcomes of the group performance reflected students' ability to integrate their knowledge by groups in BI and CH.

Table 4. Assessed levels (1-5 modified SOLO taxonomy) of hypotheses, study plans and presentations for the 13 groups, the two selected groups in bold

Group	H1	H2	H3	**H4**	H5	H6	H7	H8	H9	P1	P2	**P3**	P4
Hypothesis	5 *1	3 *2	3 *2	**4 *3**	4 *2	5 *1	5 *2	5 *1	5 *1	3 *0	3 *0	**5 *1**	4 *0
Study plans	3	5	3	**5**	4	4	5	5	4	3	3	**4**	4
Presentation	1	3	2	**5**	3	3	3	3	4	4	3	**3**	4

Note: *0-3 Number of times the hypothesis of the student group failed by the teacher

Six of the 13 groups generated level five hypotheses, and four produced level five study plans. For seven of the groups, the level was the same (3, 4 or 5) for hypotheses and study plans. A well-structured hypothesis or study plan did not guarantee a high-grade final presentation, and the number of declined hypotheses was not associated with the level of final outcomes in presentation. Based on these results the high performing H4 and average-performing P3 groups were chosen for further investigation. These two groups did not differ significantly in terms of their hypotheses and study plans, but their final outcomes were distinctly different.

CK and SR skills

According to t-test (Table 5) students with high SR scores (n=21) had significantly higher content knowledge of BI and CH both in pre and post-tests than those with low reasoning skills (n=14). At study group level, however, no uniform score changes between pre and post-tests were observed.

Table 5. Content knowledge of BI and CH measured by pre and post-tests in students with low (n=14) and high (n=21) scores on scientific reasoning skills

Groups	n		CH pre-test	CH post-test	BI pre-test	BI post-test
Low	14	Mean	16.2	16.9	18.8	18.4
		SD	0.8	1.2	1.2	1
High	21	Mean	17.3	17.5	20.6	20.2
		SD	0.7	0.6	0.7	0.6
p-value			0.0027	4.8·10-9	0,00058	6.7·10-9

(SD=Standard deviation; paired t-test)

The results indicate clear differences between the two selected student groups (H4 and P3). Group H4 returned higher mean scores than group P3 on the SR test (M = 17.3, SD = 0.3, and M = 14.7, SD = 0.3, respectively), and their final presentation was the only one awarded the maximum five points. Within both groups, individual members returned very consistent scores in each category, with one remarkable exception (S8). In group H4, student S8 raised her scores by 9 points in the CH-CK (from 12 in the pre-test to 21 points in the post-test, Appendix, Table 6). The results concerning the individual students´ level in CK, SR skills, and the group outcomes provided the baseline for the evaluation of students' spontaneous argumentation.

Spontaneous argumentation and integration of BI and CH

Spontaneous argumentation highlighted the interdisciplinarity of the virtual laboratory context, as the students had to justify their study methods and study plans using integrative reasoning (cf. RQ3). The presented two quotations showed that group H4 students employed integrative thoughts and wide causalities in their discussion, while group P3 focused on questions raised by the background material. In group H4's collaborative and argumentative discussion, S3 raised a number of questions that drove the work, offering clear answers and justifying her arguments and causing others to define their arguments. The teacher's brief interventions supported and maintained this process. The following is a short discussion between S3 and S1.

> *S3: "Okay. What are we studying now? What are we changing from the variables?" S1: "You can change a couple of those, for example pH two times and a*

couple of replicates…" S3: "Aren't these [variables]…?" S1: "What was a replicate?" S3: "A repeated experiment." [Tells the word in Finnish] S1: "Replicates, for example three replicates. Exactly. So, there are three replicates. There was something about those…" S1: "There. Three replicates. When there are [several repeated experiments], yep." S3: "As we do in physics lessons, we took three measurements of the pendulum so that it would make the measurement more accurate." S1: "Yes." S3: "Okay, yeah."

In this quotation, the students even referred to concepts learned in their physics lessons to justify their decisions. Integration of physics was also a means of understanding the need for a control—a term most often used in BI. In other words, students applied knowledge from their physics lessons to justify their claim and to help other students to understand. Importantly, group H4's success in the written presentation owed in part to the teacher's heavy argumentative support. The group was asked to rework their presentation four times before it was accepted, and the teacher gave them face-to-face advice and feedback, enabling them to revise and strengthen their arguments.

In the average group P3 two of the three students usually took part in the argumentative discussion. The teacher did not argumentatively support the students but provided the group with content knowledge and direction. The conversation showed that S1 controlled the discussion while the other students followed along and asked some counter questions. In the video excerpt, this active student wrote the hypothesis down after asking the questions. The student did not wait long for the answers. It also seems that S1 formulated the hypothesis on the basis of her own ideas, disregarding the input from the two other students. S2 and S3 were passive during the discussion and seemed to communicate a lack of knowledge. To summarize, both two groups had three students, but the groups differentiated from each other in the way the students contributed to the goal. The group H4 was led by one student, and despite this, all students' contributions were incorporated, common knowledge was identified with reference to the biology experiment and the teacher encouraged the students to move on with their reasoning. On the contrary to the group H4, P3 was completely dominated by one student leaving no room for the others to participate in discussion, students' common points were not identified, and the teacher didn't contribute to the results. The fundamental difference between groups H4 and P3 was that while the H4 students collaborated towards a common goal, the P3 discussion was controlled by one student.

When spontaneous argumentation arose in the written presentations of both groups, there was integration of biology and chemistry in the hypotheses and study plans. The students justified their arguments using biological and chemical concepts (e.g. using the concepts of population, chemical and physical environmental factors [temperature, pH] and the nauplius stage).

Groups H4 and P3 belonged to the seven out of 13 groups which showed remarkable *interdisciplinary integration* skills; the six (H1, H6, H7, H8, H9, P3) with level five in written hypotheses, four (H2, H4, H7, H8) in written study plans and one group in written presentation (H4). The crucial collaboratively constructed outcome of this study related to the integration of BI-CK and CH-CK, referring to occasions when students used both BI and CH concepts in discussing the various topics. Integration in written text typically involved justifying biological claims with chemistry-based support when generating study plans. Group discussions in the video excerpts regarding H4 and P3 showed further evidence of integration.

In the following quotation the high-performing group H4 integration occurred simultaneously, as students (S) generated hypotheses and study plans and analysed the data sets (dialogue between S1 and S3 below, S2 did not participate). The group explained their biological observations using chemical concepts, and members appeared to approach questions from a broad context. The teacher allowed the students to interact on their own.

> *S3: "So, if the temperature is raised, the blue mussels grow faster and so do the sunstars."*
>
> *S1: "Should we say that acidification…? Or that the rise of temperature…? Or that because raising temperature…?" S3: "Shouldn't that [the effect of temperature on blue mussels] be already known?" S1: "Acidic water affects those adults, the blue mussels." S3: "But will we say that…? That adult blue mussels suffer from it [acidification]?" S1: "Yep."*

Further the students in H4 combined biological and chemical concepts to form hypotheses and referred to scientific ideas to justify their claims. This passage exemplifies the simplest and most obvious interface for integration of BI and CH, where students demonstrated their understanding that pH (a chemical concept) affects the marine animal population (a biological concept).

In the average group P3, integration was demonstrated when the students wrote the study plan. The group's integrative discussion was restricted to information provided by the virtual laboratory. The lack of integration is shown in the discussion between S1 and S3.

S1: "No you are right. Where is the application? Let's start with writing that the number of people increases the amount of carbon dioxide. The amount of carbon dioxide raises the temperature of the seas…" S1: "Which leads to…" S3: "Which leads to…" S1: "The decrease of pH."

S3: "Ocean acidification."

In this dialogue, the students considered only chemical concepts when formulating their hypothesis; they did not bind the idea to the biological issues in the sea, and they did not consider possible biological causes. This group also connected the increased amount of carbon dioxide directly to the temperature of the seawater, revealing the potential for misconception. They mentioned people as a cause to ocean acidification, however, link to the socio scientific issue —reasoning/argumentation/justifying the claims —was missing. The teacher did not participate in this discussion.

Discussion

This study investigated students´ interdisciplinary integration and spontaneous argumentation for or against content issues in a virtual learning environment, which were a thoroughly new learning environment for the students. The validity of this qualitative case study was strengthened by triangulation comprising the methods, researchers and study materials (Tuomi & Sarajärvi, 2011). Validity was also strengthened by the feedback obtained from a preliminary study and by quoting students´ interdisciplinary argumentations. The used videotaped excerpts of the students´ group discussions and in addition the pre and post-tests gave a sufficient overall picture of the studied phenomenon.

The argumentative discussions occurred spontaneously in the context of generating hypotheses and study plans. The group discussions gave rise to spontaneous interdisciplinary argumentation, although this was not encouraged or supported beforehand. Of the two groups, H4 resorted to interdisciplinary argumentation to support their views, to justify their hypotheses and study settings, and to explain their results. P3's input in this field was much weaker focusing more on chemical concepts but showed traces of argumentation that could easily have been strengthened by the teacher. The shared argumentative discussion associated with CK helped students to better understand complex environmental issues (cf., Celik & Kihc, 2013; Osborne, 2010). Argumentation offered a means of social interaction and engaging in collaborative learning and was observed among students in the groups and between students. The themes of the virtual studies (climate change and, more specifically, ocean acidification) stimulated this kind of argumentation, as they cannot be properly addressed within one discipline.

The students made use of *interdisciplinary integration,* that is, they discussed the given topics using concepts from both disciplines. It could be argued that this integration occurred because the marine biology material contained chemical concepts, and students needed chemical reasoning to formulate hypotheses and to analyse study results in the VMS environment. One possible reason for the integrated use of concepts from the two disciplines is that the virtual studies were undertaken during BI and CH courses, and the two subject teachers collaborated well. Thus, in this study virtual laboratory environment was able to foster interdisciplinary reasoning and learning. These outcomes align with de Magistris (2005), who examined the integration of mathematics and physics for university students. The course was also based on a virtual laboratory, and the outcomes showed increased student comprehension of theoretical subjects as…. Taber (2008) has also emphasised that interdisciplinary integration is important in learning science and constructing coherent science knowledge systems. As even successful learners can have difficulties with interdisciplinary learning, the results of this study are very promising because students were able to integrate scientific concepts across biology and chemistry (and even physics).

Students with high reasoning skills performed better in the content tests than those with low reasoning skills, across all groups. This result is in line with Lawson (2000). He found a relationship between hypothesis testing skills and lecture content exams, and he ended up highlighting the importance of argumentation in the collaborative learning process. If students are to understand complex concepts and theories, teachers must ensure that students develop the requisite reasoning abilities (Lawson, 2004; Kuhn, 2010). Interestingly, individual group members' performance in the pre and post-tests were not reflected in the group outcome. In general, students' CK was average or good (> 15 points) as assessed on biology and chemistry CK tests.

Despite having to use a foreign language in their science studies for the first time, upper secondary school students were able to collaboratively study in VMS (e.g. generate hypotheses, study plans), following the regular national curriculum (FNBE, 2003). However, by comparison with the pilot study IB students (Yli-Panula et al., 2015), who normally used inquiry-based *missing word*, the students following the regular curriculum were unfamiliar with this kind of learning method. This difference was reflected in their uncertainty about what was expected of them. It should be noted that this effect was not apparent in the groups' collective presentations, where outcomes were mostly at level three or four. In addition, a precisely structured hypothesis or study plan did not guarantee a high-grade final presentation, and the number of declined hypotheses was not associated with the level of final outcomes in presentation. In that sense, *the VMS environment resembled a real-life laboratory situation*, where it is possible to plan a good experiment but then fail to execute it well.

Examining the impact on college students of argumentation-based education in chemistry, Aydeniz et al. (2012) showed that students in the intervention group performed significantly better than control group students, as argumentation created a social context for learners to elaborate and receive feedback on pre-existing ideas. The present study aligns with these and other findings that students spontaneously turn to argumentation when taught in an environment that facilitates public discussion among members, where students encounter criticism and tolerance for changing views, and teachers act as intellectual authorities (Duschl, 2008; Chinn & Clark, 2013; Osborne, 2010; Ryu & Sandoval, 2015).

Regarding the limitations, both teachers and students communicated uncertainty in the new learning environment. The use of identical pre and post-tests and this very short timeframe of the learning period may also have affected students' motivation to thoroughly answer questions the second time, which possibly explains the weak learning effects.

Conclusions and Implications

The outcomes of this study were encouraging despite the considerable demands on students, who had to cope with a new type of learning environment, learn in a foreign language and co-construct knowledge through arguing in group collaboration. However, dealing with content knowledge through a foreign language can support dialogue and collaboration between students as they check the meanings of different terms and understanding (Nikula, 2017). In summary, a virtual laboratory can serve to promote spontaneous scientific argumentation and the integration of two different disciplines. A virtual laboratory can be understood as an inquiry task environment, in which students need to collaborate to generate explanations, justifying these explanations and their reasoning according to scientific practices. This kind of platform is great for sociocultural purposes also and makes it visible part of the classroom learning. This study supports the idea to develop and implement a new collaborative interdisciplinary study platform embedded in digital software influencing student group collaboration and scientific argumentation, in addition supportive material for teachers is needed. In future research it will be important to study both students and teacher's collaborative argumentation in virtual environment.

References

Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, *45*, 101–131. DOI 10.1002/tea.20213

Aydeniz, M., Pabuccu A., Cetin, P. S., & Kaya, E. (2012). Argumentation and students' conceptual understanding of properties and behaviors of gases. *International Journal of Science and Mathematics Education*, *10*, 1303–1324. DOI: 10.1007/s10763-012-9336-1

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning the SOLO taxonomy (Structure of the Observed Learning Outcome)*. New York: Academic Press.

Berland L. K., & McNeill, K. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*, 765–793. DOI 10.1002/sce.20402

Bricker, L. A., & Bell, P. (2008). Conceptualizations of argumentation from science studies and the learning sciences and their implications for practices of science education. *Science Education*, *92*, 473–498. DOI 10.1002/sce.20278

Celik, A. A., & Kihc, Z. (2013). The impact of argumentation on high school students' conceptual understanding, attitude towards chemistry and argumentativeness. *European Journal of Physics and Chemistry Education*, *6*, 58–75.

Chan, C. C., Tsui, M. S., Chan, M. Y. C., & Hong, J. H. (2002). Applying the structure of the observed learning outcomes (SOLO) taxonomy on student's learning outcomes: An empirical study. *Assessment & Evaluation in Higher Education*, *27*, 511–527. doi.org/10.1080/0260293022000020282

Chinn, C. A., & Clark, D. B. (2013). Learning through collaborative argumentation. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. M. O'Donnell (Eds.), *International Handbook of Collaborative Learning* (pp. 314-332). New York: Taylor & Francis.

Chinn, C. A., O'Donnell, M. A., & Jinks, T. S. (2000). The structure of discourse in collaborative learning. *The Journal of Study Education*, *69*, 77–97.

Cohen, L., Manion, L. & Morrison, K. (2011). *Research methods in education*. New York: Routledge.

de Magistris, M. (2005) A MATLAB-based virtual laboratory for teaching introductory quasi-stationary electromagnetics. *IEEE Transactions on Education*, *48*, 81–88.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*, 287–312.

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, *38*, 39–72. DOI.org/10.1080/03057260208560187

Fauville, G., Lantz-Anderson, A., & Säljö, R. (2013). ICT tools in environmental education: Reviewing two newcomers to schools. *Environmental Education Research*, *20*, 248–283. DOI.org/10.1080/13504622.2013.775220

Finnish National Board of Education (FNBE). (2003). *National core curriculum for upper secondary schools*. Helsinki: Opetushallitus.

Hakkarainen, K. (2003). Progressive inquiry in a computer-supported biology class. *Journal of Research in Science Teaching*, *40*, 1072–1088.

Hodson, D. (2014). Learning science, learning about science, doing science: Different goals demand different learning methods. *International Journal of Science Education*, *36*, 2534–2553. DOI.org/10.1080/09500693.2014.899722

Knutsson, O., Blåsjö, M., Hållsten, S., & Karlström, P. (2012). Identifying different registers of digital literacy in virtual learning environments. *Internet and Higher Education*, *15*, 237–246. DOI.org/10.1016/j.iheduc.2011.11.002

Kuhn, D. (2010) Teaching and learning science as argument. *Science Education*, *94*, 810–824. DOI:10.1002/sce.20395

Kuhn, D. (1993) Science as argument: Implications for teaching and learning scientific thinking. *Science Education, 77*, 319–337. DOI: 10.1002/sce.3730770306

Lawson, A. E. (2000). Development of scientific reasoning in college biology: Do two levels of general hypothesis-testing skills exist? *Journal of Research in Science Teaching*, *37*, 81–101. DOI:10.1002/(SICI)1098-2736(200001)37:1<81: AID-TEA6>3.0.CO;2-I

Lawson, A. E. (2004) The nature and development of scientific reasoning: A synthetic view. *International Journal of Science and Mathematics Education*, *2*, 307–338. DOI:10.1007/s10763-004-3224-2

Nikula, T. (2017) 'What's the Moment Thingy?'–On the Emergence of Subject-Specific Knowledge in CLIL Classroom Interaction. In J. Langman & H. Hansen-Thomas (eds.) *Discourse Analytic Perspectives on STEM Education* (11–29). Basel: Springer International Publishing.

Pilot, A., & Bulte, A. M. W. (2006). The use of "contexts" as a challenge for the chemistry curriculum: Its successes and the need for further development and understanding. *International Journal of Science Education*, *28*, 1087–1112. DOI.org/10.1080/09500690600730737

Robertson, A. D., & Shaffer, P. S. (2014). "Combustion always produces carbon dioxide and water": a discussion of university chemistry students' use of rules in place of principles. *Chemistry Education Research and Practice*, *15*, 763–776. DOI:10.1039/C4RP00089G

Ryu, S., & Sandoval, W. A. (2015). The influence of group dynamics on collaborative scientific argumentation. *Eurasia Journal of Mathematics, Science & Technology Education*, *11*, 335–351. DOI.org/10.12973/eurasia.2015.1338a

Sampson, V., & Clark, D. (2008). The impact of collaboration on the outcomes of scientific argumentation. *Science Education*, *93*, 448–484. DOI:10.1002/sce.20306

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. DOI.org/10.1080/17437270802417176

Taber, K. S. (2008). Exploring conceptual integration in student thinking: evidence from a case study. *International Journal of Science Education*, *30*, 1915–1943. DOI.org/10.1080/09500690701589404

Taber, K. S. (2015). The Role of Conceptual Integration in Understanding and Learning Chemistry. In J. García-Martínez & E. Serrano-Torregrosa (Eds.), *Chemistry Education: Best Practices, Opportunities and Trends* (pp. 375–394). Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA. DOI:10.1002/9783527679300.ch15

Tuomi, J. & Sarajärvi, A. (2011). *Laadullinen tutkimus ja sisällön analyysi* [Qualitative study and content analyses]. Helsinki: Tammi.

Vauras, M., & Volet, S. (2013). The study of interpersonal regulation in learning challenges the research methodology. In S. Volet and M. Vauras (Eds.), *Interpersonal Regulation of Learning and Motivation: Methodological Advances* (1–13). New York: Routledge.

Yli-Panula E, Hiilovaara-Teijo M. & Vauras M. 2015. High school students' inquiry based collaborative learning in virtual marine science laboratory. Kauppinen M, Rautiainen M & Tarnanen M. *Suomen ainedidaktisen tutkimusseuran julkaisuja, Ainedidaktisia tutkimuksia 8*: 135-152.

Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Sociocognitive dynamics of knowledge building in the work of nine- and tenyear olds. *Educational Technology Research and Development*, *55*, 117–145. DOI: 10.1007/s11423-006-9019-0