Challenges in Biology Education Research

Edited by Niklas Gericke and Marcus Grace

A selection of papers presented at the XIth conference of European Researchers in Didactics of Biology (ERIDOB)

Challenges in Biology Education Research

Edited by
Niklas Gericke
Marcus Grace

A selection of papers presented at the XIth conference of
European Researchers in Didactics of Biology (ERIDOB)

Challenges in Biology Education Research

Edited by
Niklas Gericke
Marcus Grace

The XIth ERIDOB Conference

ACADEMIC COMMITTEE
Dr Marcus Grace, Southampton University, UK (Chair)
Dr Marida Ergazaki, University of Patras, Greece
Dr Niklas Gericke, Karlstad University, Sweden
Dr Maria José Gil Quílez, University of Zaragoza, Spain
Dr Marie-Christine Knippels, University of Utrecht, Netherlands
Dr Grégoire Molinatti, University Montpellier II, France
Dr Tali Tal, Technion, Israel Institute of Technology, Haifa, Israel
Dr Jörg Zabel, University of Leipzig, Germany

LOCAL ORGANIZING COMMITTEE
Dr Niklas Gericke (Chair)
Dr Nina Christenson
Dr Birgitta Mc Ewen
Lic Karin Thörne
Dr Susanne Walan
Maria Kull (Conference Administration)

ERIDOB Conferences	Published book including selected papers from the conference
2016 Karlstad, Sweden:	*Challenges in Biology Education Research*, 2018. Edited by Niklas Gericke and Marcus Grace.
2014 Haifa, Israel:	*The Future of Biology Education Research*, 2016. Edited by Tali Tal and Anat Yarden.
2012 Berlin, Germany:	*Research in Biological Education*, 2014. Edited by Dirk Krüger and Margareta Ekborg.
2010 Braga, Portugal:	*Authenticity in Biology Education: Benefits and Challenges*, 2011. Edited by Anat Yarden and Graça S. Carvalho.
2008 Utrecht, The Netherlands:	*The Nature of Research in Biological Education: Old and New Perspectives on Theoretical and Methodological Issues*, 2009. Edited by Marcus Hammann, Arend Jan Waarlo and Kerst Boersma.
2006 London, UK:	*Biology in Context: Learning and Teaching for the Twenty-first Century*, 2008. Edited by Marcus Hammann, Michael Reiss, Carolyn Boulter and Sue Dale Tunnicliffe.

2004 Patras, Greece:

Trends in Biology Education Research in the New Biology Era, 2005. Edited by Marida Ergazaki, Jenny Lewis and Vassiliki Zogza.

2002 Toulouse, France:

Biology Education for the Real World: Student – Teacher – Citizen, 2003. Edited by Jenny Lewis, Alexandra Magro and Laurence Simonneaux.

2000 Santiago de Compostela, Spain:

Proceeding of the III Conference of European Researchers in Didactic of Biology, 2001. Edited by Isabel García-Rodeja Gayoso, Joaquín Díaz de Bustamante, Ute Harms and María Pilar Jiménez Aleixandre.

1998 Göteborg, Sweden:

Research in Didaktik of Biology, 2000. Edited by Björn Andersson, Ute Harms, Gustav Helldén and Maj-Lis Sjöbeck.

1996 Kiel, Germany:

What? – Why? – How? Research in Didaktik of Biology, 1998. Edited by Horst Bayrhuber and Fred Brinkman.

Table of contents

Introduction – challenges, trajectories and opportunities for biology education research

Niklas Gericke[1] and Marcus Grace[2]

niklas.gericke@kau.se

This volume consists of original papers presented at the 11th Conference of European Researchers in Didactics of Biology (ERIDOB) organized by the academic committee of ERIDOB and the Centre of Science, Mathematics and Engineering Education Research (SMEER) at Karlstad University in Sweden. The conference took place on 5-9 September 2016 with 165 participants representing 24 countries. There were 77 oral presentations, including four symposia, and 52 poster presentations.

Since 1996, European Researchers in Didactics of Biology have met every two years to present and discuss research findings related to teaching and learning biology. As the research community has expanded so have the conferences, and we now also have participants from Asia, Africa, North and South America. However, the conference still remains sufficiently small, transparent and welcoming to keep the famous ERIDOB spirit, with talks of excellent quality in an atmosphere filled with joy and generosity.

After the conference the presenters were asked to send in extended papers for this book. From the submitted papers, 24 were selected for this volume, after being independently reviewed by at least two members of the ERIDOB Academic Committee. The ERIDOB conference has been collaborating with the *Journal of Biological Education* (Grace & Gericke 2018) and three of these papers were selected to be published in the journal (Harris & Winterbottom 2018; Lombard et al. 2018; and Scheuch et al. 2018). These papers are therefore only available as abstracts in this volume to avoid double publication.

---

[1] Karlstad University, Sweden

[2] University of Southampton, UK

Contributions to the ERIDOB conference should fit into one of the following strands within biology education:

- Students' conceptions and conceptual change
- Students' interest and motivation
- Students' values, attitudes and decision-making
- Scientific thinking, nature of science and argumentation
- Teaching strategies and teaching environments
- Teaching and learning with educational technology
- Environmental education
- Health education
- Social, cultural, and gender issues
- Practical work and field work
- Research methods and theoretical issues

At the conference, we had contributions within all these strands, though *Students' conceptions and conceptual change, Students' interest and motivation, Scientific thinking, nature of science and argumentation* and *Teaching strategies and teaching environments* were the strands generating the greatest interest and receiving more contributions than the other strands. This is also the case for this volume of selected papers, which we have organized around these four strands under section one. Each strand is presented as a chapter and all the papers within each chapter are in alphabetical order.

In the first chapter of section one the contributions revolve around *students' conceptions and conceptual change*. Georgios Ampatzidis and Marida Ergazaki explore students' conceptual reasoning about disturbed and protected ecosystems after a teaching intervention. In a paper by Pierre Clément and colleagues, teachers' conceptions of environment is investigated and compared between five Sub-Saharan African countries. They found great differences which were dependent upon the sociocultural contexts of the countries. The contribution by François Lombard et al. outlines the Reference Map Change Coding (RMCC) method for revealing students' conceptual progression at a fine-grained level. Concepts and causal links expressed in students' learning progression is charted onto reference model maps. Learners'

diverse and often surprising conceptual paths challenge the view of learning as a linear process. In the fourth contribution in this chapter Edward Mifsud and Sue Dale Tunnicliffe conceptualise natural history habitat dioramas as potential models for biological learning of local flora and fauna for primary school children. In the final paper of the chapter Nadine Tramowsky and Jorge Groß analyse moral arguments of students (10-16 years old) in respect to meat consumption and livestock farming. The results illustrate the Educational Moral Metaphors System as an analytical tool and how different moral arguments and students' conceptions can be related.

The second chapter of section one focuses on *students' interest and motivation*. Petra Duske and Michael Ewig first investigate whether Content and Language Integrated Learning (CLIL) teaching in biology influences students' motivation and their acquisition of knowledge in biology. In a CLIL approach, a foreign language is used for the learning and teaching of both content and language. The results show that no correlation was found between the students' growth of knowledge during the CLIL unit and their motivation or competence in English. In the second paper Alexander Eckes and colleagues report on motivational effects of structure and autonomy support on 11-12 year old students in non-formal learning at a university exhibition. Following up the two previous studies Annika Rodenhauser and Angelika Preisfeld explore intrinsic motivation in bilingual courses on molecular biology and bionics in an out-of-school laboratory. In the final paper of this chapter Mariella Roesler and colleagues study the influence of interest and motivation on grade 10 students' performance in written tests with biological content and contexts. The results indicate that there are no differences in average task difficulty across contexts. Nevertheless, students' interest and motivation vary depending on the context and content of a task, and the health context turned out to be particularly interesting and engaging for the students.

The third chapter of section one addresses *scientific thinking, nature of science and argumentation*. In the first paper Andreani Baytelman and colleagues investigate the contribution of pre-service primary teachers' epistemological beliefs to their informal reasoning, while trying to manage health socio-scientific issues. The results indicated that relatively sophisticated epistemological beliefs, especially about the structure of knowledge, positively predicted the quality of arguments posed. Alexander Bergmann and Jörg Zabel then investigate

students' (13-16 years old) perspectives on the ethical, social, and legal dimension of neuroscience by involving the students in an everyday-like discourse. The results show that 'misuse of science' is a central frame the students use when discussing science and its relation to society. Birgitta Berne reports on the impact of peer discussions on 14-15 year old students' arguments when addressing socio-scientific issues in biotechnology relating to GMO. The results show that recurrent peer discussions support the students in advancing their argumentation. Marida Ergazaki and her colleagues report on a case study that explores young children's (age 5-5.5) reasoning ability to categorize biological and non-biological entities presented to them as pictures or 3-dimensional objects. The contribution by Anne Laius and her colleagues shows that the implementation of a new competence-based biology curriculum has positive effects on students' (16-18 years old) scientific creativity skills and to a lesser extent to their socio-scientific reasoning skills. Ros Roberts then discusses the role of biology context as a way to teach and learn evidence, an essential component of 'scientific practice'. In the final paper of this chapter, Laurence Simonneaux and Jean Simonneaux analyse the views of French students, from different kinds of agricultural training, on the socially acute question of the wolf as a 'pest animal'. In particular, they discuss the registers the students mobilize (cognitive or emotive or both), their engagement in the issue and their ethical position.

The fourth, and largest, chapter of section one presents papers related to *teaching strategies and teaching environments*. Lea Brauer and Corinna Hößle investigate how diagnostic skills are acquired in a learning laboratory. Anne Cohonner and Jürgen Mayer investigate retrieval-based learning in the context of inquiry-based learning and find that inquiry tasks including retrieval-practice and rereading outperform the control condition. Emily Harris and Mark Winterbottom write about an exploratory case study that set out to discover and analyse learning happening within family groups during a visit to a museum natural history gallery, and suggest how family learning can be developed as a learning strategy within biology education. Katharina Luther and her colleagues show that slower learners and visual learners in particular benefit from the clearer structure and the use of realistic video clips for preparation in web-based training during dissections in biology classes. Sonja Schaal and colleagues study the use of mobile technology in biology education, by com-

bining game-based and location-based learning, to provide sensuous experiences of local biodiversity. The results demonstrate a significant increase of biodiversity-related knowledge in general and a stronger connection to nature from pre to post test. Following a citizen science project Martin Scheuch and his colleagues longitudinally investigate biology teachers' PCK development about teaching of biological methods over the two years through the project. Successes as well as failures are reported. Iris Schiffl and Carina Wurdinger then examine whether hands-on experiments are suitable for diagnosing inquiry competencies, and the results show that competency levels were graded lower if competencies were measured using hands-on experiments instead of written test designs. In the final paper of this chapter Ana Valdmann and Miia Rannikmäe find that context-based teaching including socio-scientific health problems and inquiry based learning are effective tools for students to acquire more knowledge in anatomy and physiology.

The theme for the 11th ERIDOB conference was *Challenges of Biology Education Research* – the same as the title of this book of collected papers, and a panel debate around this issue was arranged with William McComas, as organizer. Included in the debate were contributions from seven scholars in the field of biology education research representing seven countries from four continents. In section two of this volume, we include the peer-reviewed versions of these short papers as a joint article commenting on the current challenges, trajectories and opportunities for biology education. As McComas points out in the introduction to the contribution "biology education is a subset of science education with its own context and, to a degree, its own questions", and that is what is explored in this article.

In the first piece of the article, Michal Reiss proposes a framework for biology education research in which the intersection between *biology, education* and *research* is the point of departure. Edith Dempster continues the discussion in the second piece about the qualities of effective biology teacher education programmes and proposes future avenues for research in biology teacher education to follow. In the third contribution Yeung Chung Lee takes a learning perspective based on the well-known learning difficulties in biology education, and based on that fact he proposes that we should reconsider biology education in a constructivist way including three perspectives in biology education: *epistemological, metacognitive*, and *motivational*. In the fourth contribution

Clas Olander addresses the importance of language for learning biology, and the challenge of developing teachers' pedagogical knowledge with respect to language in biology education. Pierre Clément emphasizes the importance of sociocultural contexts for biology education research in piece five, and he introduces two theoretical frameworks. The first framework is the *KVP model* (Knowledge, Values, and Attitudes) that can be used to explore how conceptions taught in biology education are dependent not only on knowledge (K), but also values (V) and social practices (P). The second framework, *didactic delay theory*, revolves around the time delay between the publication of a new scientific concept and its introduction within school. In the sixth piece, Dirk Jan Boerwinkel and Arend Jan Waarlo address the societal applications of biology knowledge and the importance of biology education to support citizenship and empowering students for decision-making through teaching socio-scientific issues. In the final piece of the article William McComas makes some concluding suggestions, based on the above contributions, for future challenges, trajectories and opportunities for biology education research.

We hope that this volume will find its way to biology educators as well as biology education researchers and be a useful contribution to the development of biology education in Europe and the world.

References

Grace, M. & Gericke, N. (2018). JBE and ERIDOB: working together to support biology education research. *Journal of Biological Education*, 52(1), 1-2.

Harris, E. & Winterbottom, M. (2018) "Why do parrots talk?" co-investigation as a model for promoting family learning through conversation in a natural history gallery. *Journal of Biological Education*, 52(1), 89-100.

Lombard, F., Merminod, M., Widmer, V. & Schneider, D.K. (2018) A method to reveal fine-grained and diverse conceptual progressions during learning, *Journal of Biological Education.* 52(1), 101-112.

Scheuch, M., Panhuber, T., Winter, S., Kelemen-Finan, J., Bardy-Durchhalter, M. & Kapelari, S. (2018) Butterflies & wild bees: biology teachers' PCK development through citizen science, *Journal of Biological Education*, 52(1), 79-88.

SECTION ONE –

Selected papers from the ERIDOB 2016 conference

Chapter 1:
Students´ conceptions and conceptual change

Can the idea of the 'balance of nature' be challenged? Students' reasoning about disturbed and protected ecosystems after a teaching intervention, and one year later

Georgios Ampatzidis[1] and Marida Ergazaki[1]

ampatzidis@upatras.gr

Abstract

This paper reports on insights from the 3rd cycle of a developmental research study, which aims to design a learning environment that could support non-biology major students in (a) challenging the idea of 'the balance of nature' and constructing an up-to-date understanding of the function of ecosystems, and (b) using this understanding to advance systems thinking skills. Our focus is on whether and how students' reasoning (predictions/justifications) about the response of ecosystems to human-driven disturbance/protection altered within the 3rd cycle of our study, performed with the 3rd version of our learning environment, and one year later. Informed by constructivism and a problem-posing approach, we developed a computer-supported collaborative learning (CSCL) environment for highlighting the contingent behaviour of ecosystems through the currently valid idea of 'resilient nature'. First year undergraduate educational sciences students (N=44) were introduced to the assumptions of this idea in five 2-hour sessions, by exploring our NetLogo models of protected/disturbed ecosystems aided by worksheets. The analysis of 23 students' responses to certain items of the pre/post/delayed post-test showed that the ideas of protected ecosystems' stability and disturbed ecosystems' full recovery retreated in the post-test and maintained low frequencies one year later. Moreover, the idea of ecosystems' contingent behaviour appeared in high frequencies in the post-test, and in lower but still notable frequencies one year later.

---

[1] University of Patras, Patras, Greece

1. Introduction

Research on students' reasoning about ecosystems and particularly how they respond to disturbance caused, or protection ensured, by humans, reveals an extensive belief in the idea of the 'balance of nature' (or BON-idea) (Zimmerman & Cuddington, 2007). This is a persistent, well known view about the natural world, which implies a predetermined order and stability, assured by the will of a divine power or nature itself (Cooper, 2001; Cuddington, 2001). Although it has been criticized as not representative of natural systems (Cooper, 2001; Cuddington, 2001), the BON-idea seems to prevail in public opinion (Ladle & Gillson, 2009), school science (Jelinski, 2005; Westra, 2008), and students' reasoning about ecosystems' responses to human-driven disturbance or protection (Ergazaki & Ampatzidis, 2012).

Nevertheless, the BON-idea may interfere with our environmental awareness, since it seems to imply that ecosystems have an almost 'magic' ability to restore their initial state and thus it may undermine the significance of not disturbing them (Gunderson et al., 2010; Westra, 2008). Moreover, the BON-idea may hinder conceptual understanding as well, since it opposes the idea of 'resilient nature' which is suggested by up-to-date research on ecosystems (Scheffer, 2009). According to the latter, nature is not considered as 'balanced', while contingency is favoured over purpose and order. Ecosystems are believed to function in multiple alternative states which are self-organized through feedback, and to shift between these states in abrupt, reversible or irreversible ways (Gunderson et al., 2010; Holling, 1973; Scheffer, 2009). The 'resilient nature' idea seems also to offer an appropriate context for fostering systems thinking skills, important for all aspects of life (Richmond, 2004).

Rather surprisingly, there are no reports on 'anti-BON' teaching interventions. Hovardas and Korfiatis (2010) have provided the theoretical underpinnings and the overview of one, but they have not tested it with students. Nevertheless, their intention was to use computer modelling in order to simulate the dynamic processes within ecosystems in comprehensible ways. Computer simulations have been widely used for science teaching and ecology teaching in particular (Rutten et al., 2012), to help visualize important dynamic pro-

cesses taking place at organization levels or time-scales that cannot be easily accessed by students, and promote inquiry-based learning through meaningful interactions among peers and with educational tools (Ergazaki & Zogza, 2008). In fact, computer simulations may function as highly effective educational tools, by more concretely representing aspects of the natural world that are too abstract or complex to deal with, as well as by allowing students to experience cognitive dissonance and reflect on their own initial ideas in its light (Smetana & Bell, 2012).

Considering how important and challenging the BON-idea seems to be, especially in the absence of relevant attempts to address it, and how helpful new technologies might be for this purpose, we decided to perform a developmental research study with 3 cycles of designing-implementing-redesigning a computer-supported 'anti-BON' learning environment that could give students the opportunity for a better understanding of nature's function. Our study addresses the question of whether it is feasible to design a learning environment that could support non-biology major students, (a) in challenging the idea of 'the balance of nature' and constructing an up-to-date understanding about the function of ecosystems, and (b) in using this understanding to advance context-free ideas such as interdependence and reciprocality, which have to do with systems thinking skills. In this paper, we are particularly concerned with (a), while (b) is addressed elsewhere (Ampatzidis & Ergazaki, in press).

More specifically, what concerns us here is to identify whether and how students' reasoning about how ecosystems respond to human-driven disturbance or protection has been altered within the 3rd version of our learning environment, and how stable this change might be one year later. This was developed and tested in the 3rd cycle of our research study. The research questions are:

(1) What kind of predictions do students make about the future of disturbed/ protected ecosystems before, after, and one year after their participation to the 3rd version of our learning environment?

(2) What kind of justifications do they provide for their predictions?

2. Methods

2.1 Overview of the study

This is a mixed-model case study performed in the 3rd cycle of our developmental research programme (Akker et al., 2006), in order to test the 3rd version of our computer-supported, collaborative, 'anti-BON' learning environment. The design of the learning environment was performed in the broader context of social constructivism (Driver et al., 1994) and took into account a problem-posing approach (Klaassen, 1995). Its aim was to highlight the contingent behaviour of ecosystems through the assumptions of the idea of 'resilient nature', in order to support students in (a) challenging the BON-idea and constructing a meaningful, up-to-date understanding of how nature may work, and (b) using this understanding to advance systems thinking skills transferable to everyday life (the latter is not covered here). Apart from designing the learning environment, we developed a pre/post questionnaire with open-ended items in order to collect data about its effectiveness, with the aid of short interviews when required. The questionnaire was delivered to students as: (a) a pre-test one week before the classes, (b) a post-test immediately after the classes, and (c) a delayed post-test after a whole year with no more ecology. Finally, we analysed students' responses using the qualitative analysis software 'NVivo' (Gibbs, 2005) and tested for the statistical significance of their progress using the quantitative analysis software SPSS.

2.2 The participants

The participants of the 3rd cycle of the research were 44 first-year students of the Department of Educational Sciences and Early Childhood Education of the University of Patras (aged 18-19 years), who were attending the optional course 'Essential Concepts of Ecology' offered by the second author. They had all completed the same ecology curriculum for entering university, they were familiar with computers and group work, and they seemed to be rather interested in ecology in terms of raising/answering questions in the course's regular classes. These 44 students from a total of 180 who were enrolled in the course, volunteered to participate in the study, after they had been informed of its goals and time schedule, and they were reassured that they were free to withdraw at any time.

2.3 The learning environment

The learning environment aims to highlight the contingent behaviour of ecosystems through the basic assumptions of the idea of 'resilient nature'. The learning objectives (LOs) are around the understanding these assumptions (LO1-LO4), using them to (a) challenge the notion of balance as an inherent feature of nature, and (b) move to the notion of contingency ($LO_{-contingency}$). In other words, $LO_{-contingency}$ is a 'higher level' learning objective; LO1-LO4 should be first met one by one and then contribute all together in recognizing nature's contingency and thus abandon the idea of 'balance'. An overview of the LOs is presented in Table 1.

Table 1 Learning Objectives (LO).

Learning Objectives	
LO1	Ecosystems may have multiple alternative states
LO2	Each state is self-organized through feedback changing at tipping points
LO3	Shifts between alternative states may be irreversible or reversible based on initial state or handlings
LO4	Reversing the factor that caused the shift, does not necessarily bring the ecosystem to its prior state
LO$_{-contingency}$	Natural systems show a contingent and not pre-determined behaviour ('resilient nature' vs 'the balance of nature')

Students were actively introduced to the target assumptions of the idea of 'resilient nature' in five 2-hour sessions of the optional ecology course. In sessions 1-4, they explored 'NetLogo' models (NM), with the aid of worksheets that required predictions about the ecosystem's behaviour before using the model and explanations afterwards. The four models we developed for the study, simulated terrestrial or aquatic ecosystems faced with internally or externally triggered changes, and were based on findings of current ecological research (Table 2). Each model had two different versions, showing two different trajectories of the ecosystem (NM1-2: state maintenance/change, NM3-4: state recovery/non-recovery), depending on initial conditions (NM1, NM2, NM4) or on certain human actions in the recovery plan (NM3). Half of the students' triads explored the first version while the other half explored the second version. The two different trajectories that were simulated by each model were discussed with the whole class at the end of each session.

Table 2 'NetLogo' models and learning objectives per session.

Sessions (S)	'NetLogo' Models (NM)	Reference Model	Learning Objectives (LO)	
S1	NM1 - Forest: forest maturation two trajectories (initial conditions)	Gunderson et al., 2010	LO1-2	LO-contingency
S2	NM2 - Lake: inflow of nutrients & subsequent termination of it two trajectories (initial conditions)	Scheffer, 2009	LO1-4	LO-contingency
S3	NM3 - Lake: inflow of nutrients & subsequent removal of nutrients & additional corrective actions two trajectories (recovery plan)	Scheffer, 2009	LO1-4	LO-contingency
S4	NM4 - Meadow: removal of spiders & subsequent re-introduction of spiders two trajectories (initial conditions)	Schmitz, 2010	LO1-4	LO-contingency

The interface of the models included the following: (a) a series of boxes showing population size (i.e. the number of individuals) and sometimes also the level of key abiotic factors (e.g. nutrients), (b) a 'simulation window' showing the individuals of the different populations in different shapes and colours, and thus providing students with a relatively concrete visual representation of what happens in the ecosystem as time passes, and (c) a 'graph window' showing changes regarding population size and levels of key abiotic factors with time and thus providing students with a graphical representation of the trajectory of the ecosystem that they are actually required to explore (see Figure 1 from left to right). Finally, in the fifth session, students were engaged in reasoning about the behaviour of ecosystems through 'landscape models' made from plasticine, cardboard and hands-on activities concerning systems thinking.
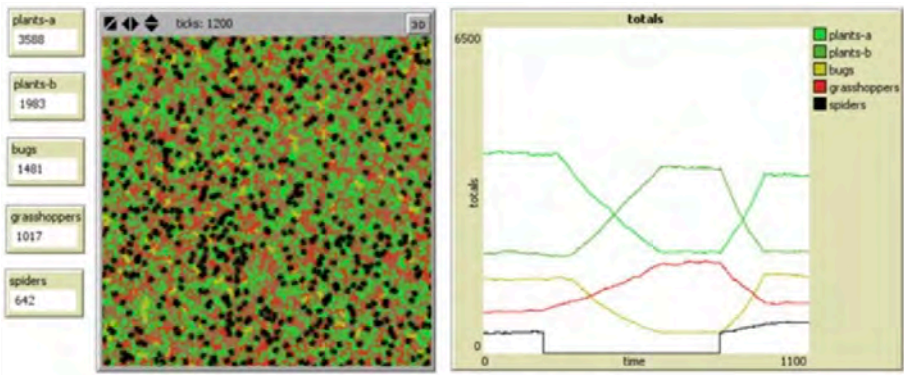
Figure 1 NM4-Meadow model, version 1.

2.4 The pre/post-questionnaire

A pre-, post- and delayed post-test questionnaire was filled in by 23/44 participants (a) before the classes, (b) after the classes, and (c) after a whole year with no more ecology, respectively. The high drop-out rate in the delayed post-test (21/44) was probably because the implementation of the learning environment was completed a long time before, and students were busy with other courses when asked to contribute again to the study. In all three cases, it was explicitly explained to students that the questionnaire was not meant to test *them* like an exam, but to give *us* the opportunity for a valuable insight into their understanding of nature. The first part of each questionnaire included four, equivalent open-ended items about the behaviour of protected or disturbed ecosystems (Table 3), derived from the elaboration of earlier versions we used in the previous research cycles. Items 1 and 4 that concern us here (see Appendix) aimed at probing specific target assumptions (LO1,3,4) *as justifications* for the contingency ($_{J\text{-contigency}}$) of ecosystems' behaviour ('LO1,3,4$_{J\text{-contigency}}$').

Table 3 The pre/post/delayed post-test questionnaire.

Items	Require	Probes …
1. Protected ecosystem	Reasoning about the future of a terrestrial / aquatic national park protected by humans	LO1$_{J\text{-contigency}}$
2. Feedback	Explaining (a) population size control in a swamp / lake through feedback-mediated self-organization, and (b) the loss of control through feedback-change at a tipping point	LO2
3. Disturbed ecosystem (biotic change)	Reasoning about the future of a lake / forest where a new population is added / removed by humans	LO3-4$_{J\text{-contigency}}$
4. Disturbed ecosystem (abiotic change)	Reasoning about the future of a lake where salinity / nutrients are increased by humans (this makes an animal population go extinct) / restored by humans who also re-introduce the extinct population	LO3-4$_{J\text{-contigency}}$

The first author read all the responses as soon as the students had completed the questionnaire, and carried out short interviews with those whose responses needed clarification.

2.5 The analytic procedure

Students' responses to the pre/post/delayed post-questionnaires (and relevant notes from the interviews, where applicable) were transcribed and coded with-

in NVivo (Gibbs, 2005). The several categories that emerged, were organized into a coding scheme, which was divided into: (a) students' 'predictions' about the future of the ecosystem in question (e.g. *'same picture', 'full recovery', 'contingent behaviour'*), and (b) students' 'justifications' for what they predicted (e.g. *'differences in population size', 'recovery mechanisms', 'incomplete recovery process'*). Our prediction and justification categories were mutually exclusive, so each response could be coded as a unique prediction category and a unique justification category. The coding was performed by both authors with a satisfactory agreement: Cohen's Kappa with regard to items 1 and 4 that concern us here was estimated at 0.85.

To test students' progress and its statistical significance, we developed a scoring grid for their responses to each item of the questionnaire (Table 4). The score of each response was the sum of two sub-scores: one for the prediction about the ecosystem's future and another for its justification. The prediction of an 'unpredictable picture' was assigned the highest score, while the predictions of 'same picture'/'different picture' and 'full recovery'/'no recovery' were assigned the lowest score. Similarly, each justification was assigned a score depending on the level of understanding that it showed (Table 4).

Table 4 The scoring grid.

Items	Predictions	P score	Justifications	J score	Total score
Item 1	Unpredictable/contingent picture	3	Possible feedback change	0.5	3.5
			Unpredictable factors	0.25	3.25
	Possible different picture	2	Possible changes in population sizes & environmental factors	0.75	2.75
			Possible changes in population sizes	0.5	2.5
	Different picture	1	Changes in population sizes	0.5	1.5
	Same picture	1	Self-regulated populations if not human-disturbed	0.5	1.5
Item 4	Unpredictable/contingent picture	3	Possible tipping point reached / feedback change	0.75	3.75
			Feedback	0.5	3.5
			Possible differences in recovery handlings	0.5	3.5
			Possible side effects	0.5	3.5
			Unpredictable factors	0.25	3.25
	Possible full recovery	2	Possible recovery process	0.25	2.25
	Full recovery	1	Counteracting feedback loops	0.5	1.5
			Recovery process	0.25	1.25
	No recovery	1	Tipping point reached	0.5	1.5
			Changes in population relationships	0.5	1.5
			Evolution	0.5	1.5
			Human-triggered disturbance	0.5	1.5

The scoring grid was developed so that predictions contributed more than justifications to the total score: satisfactory predictions with non-satisfactory justifications got a higher score than non-satisfactory predictions with satisfactory justifications. Finally, responses with no predictions got zero (0), while responses with unjustified predictions were scored according to the prediction only.

3. Findings

3.1 Students' predictions

Regarding students' reasoning about the future of a protected forest or aquatic park (item 1), we note that the frequency of the 'same picture' prediction decreased in the post-test and *did* remain low in the delayed post-test (pre-

test: 14/23, post-test: 7/23, delayed post-test 8/23) (Figure 2). In contrast, the frequency of the 'different picture' prediction which also decreased in the post-test, *did not* remain low in the delayed post-test (pre-test: 4/23, post-test: 1/23, delayed post-test 6/23) (Figure 2).
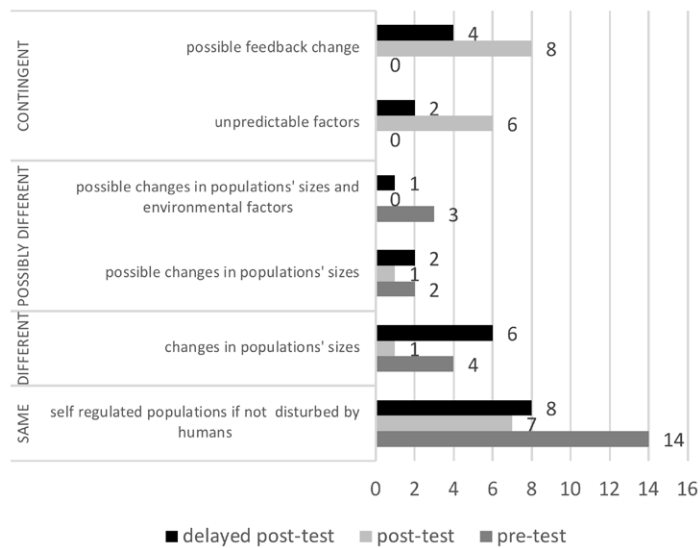
Figure 2 Categories of predictions/justifications about a protected ecosystem (item 1).

Concerning the 'possibly different picture' prediction, we note that its frequency decreased in the post-test and remained rather low in the delayed post-test (pre-test: 5/23, post-test: 1/23, delayed post-test 3/23) (Figure 2). Finally, the target prediction of the ecosystem's 'unpredictable/contingent picture' which was absent from the pre-test, was formulated in *both* the post- and the delayed post-test, reaching its highest frequency in the former (pre-test: 0/23, post-test: 14/23, delayed post-test 6/23) (Figure 2).

Regarding students' reasoning about the future of a lake, where: (a) the salinity or the number of nutrients increased due to human activity leading to the extinction of one population, (b) it was subsequently decreased by human intervention, and (c) the extinct population was reintroduced (item 4), we note that the 'no recovery' prediction, which almost disappeared in the post-test, re-appeared in the delayed post-test with its pre-test frequency (pre-test: 6/23, post-test: 2/23, delayed post-test 6/23) (Figure 3). In contrast, the frequency of the 'full recovery' prediction decreased in the post-test and *did*

remain low in the delayed post-test (pre-test: 16/23, post-test: 3/23, delayed post-test 4/23) (Figure 3).
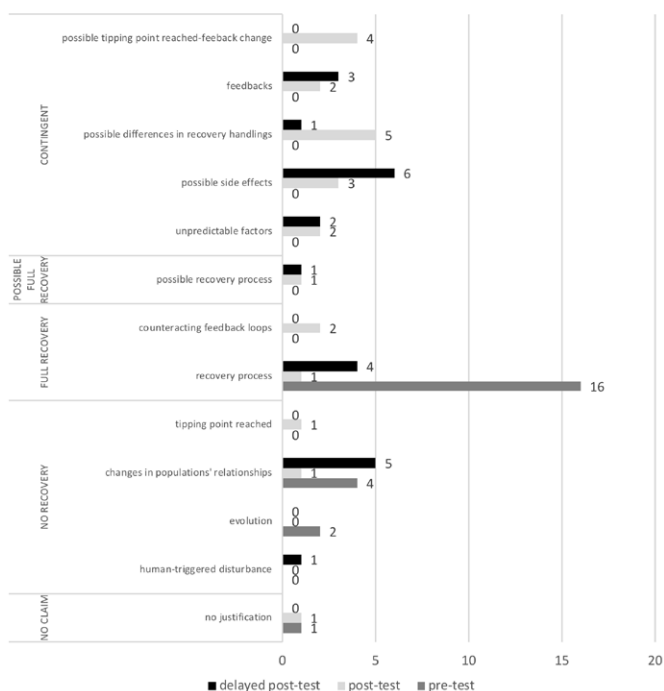
Figure 3 Categories of predictions/justifications about a disturbed ecosystem (item 4).

Concerning the 'possible full recovery' prediction, we note its absence in the pre-test and its appearance with the same low frequency in the post- and the delayed post-test (pre-test: 0/23, post-test: 1/23, delayed post-test 1/23) (Figure 3). Finally, the target prediction of the ecosystem's 'unpredictable/ contingent picture', which was absent from the pre-test, was formulated in *both* the post- and the delayed post-test, reaching its highest frequency in the former (pre-test: 0/23, post-test: 16/23, delayed post-test 12/23) (Figure 3).

3.2 Students' justifications

Students who predicted the 'same picture' for a protected forest or aquatic park justified it through the idea of 'self-regulation of populations if not disturbed by humans' (Figure 2). The 'different picture' prediction was grounded on 'changes in populations' sizes' (Figure 2), as shown in the following extract:

"The populations that live in this aquatic park are in balance due to specific trophic relations among them. I think that some years later the populations will be in a different balance with different sizes than before." (delayed post-test).

The 'possibly different picture' prediction was justified by 'possible changes in populations' sizes', sometimes also combined with 'environmental factors' (Figure 2). The 'unpredictable/contingent picture' prediction was based on 'unpredictable factors' or 'possible feedback change' (Figure 2). The latter is shown in one student's own words: *"Some years later, the aquatic park may be the same or not. There are counteracting feedback loops functioning among the populations of the park; there is a chance that one population may reach a tipping point and as a result a reinforcing feedback loop may start functioning. In that case, the aquatic park will shift state and it will show a different picture."* (delayed post-test).

Concerning the ecosystem which undergoes a human intervention, the students who predicted 'no recovery' for the ecosystem, justified their prediction by appealing to 'human-triggered disturbance', 'evolution', 'changes in populations' relationships', and the 'reaching of a tipping point' (Figure 3). The 'full recovery' prediction was grounded on the 'counteracting feedback loops' or the 'recovery process' (Figure 3). The latter is shown in one student's own words: *"Some years later the forest will be the same and all the populations will be stable, since all the organisms that die are replaced by the ones that are getting born."* (pre-test).

The 'possible full recovery' prediction was justified by a 'possible recovery process' (Figure 3). The target prediction of the ecosystem's 'unpredictable/contingent picture' was based on 'unpredictable factors', 'possible side effects', 'possible differences in recovery handlings', 'feedback', or the 'possible reaching of a tipping point/feedback change' (Figure 3). The latter is shown in one student's own words: *"There are two possibilities. There is a chance that the increase of the nutrients' level has not reached a 'tipping point' and thus the human actions to restore the lake to its initial state may be successful. But, there is also a chance that the changes that happened in the lake due to the increase of the nutrients' level have reached some important 'tipping point' and thus shifting back to its previous state is not possible for the lake even after the human effort."* (post-test).

3.3 Students' reasoning strands: predictions and justifications

The findings of the Wilcoxon signed-rank test we performed to determine whether the scores assigned to students' reasoning strands (i.e. their justified

predictions) for items 1 and 4, differed in a statistically significant way between the pre-, the post- and the delayed post-test, are shown in Table 4.

Table 4 Wilcoxon signed-rank test results.

Items	Pre-test / post-test	Pre-test / Delayed post-test	Post-test / Delayed post-test
Item 1	Pre-test mean= 1.75 Post-test mean = 2.69 Z= -3.336, p<0.05	Pre-test mean =1.75 Delayed post-test mean = 2.14 Z= -2.138, p<0.05	Post-test mean =2.69 Delayed post-test mean = 2.14 Z= -2.085, p<0.05
Item 4	Pre-test mean=1.26 Post-test mean = 2.87 Z= -3.905, p<0.05	Pre-test mean =1.26 Delayed post-test mean =2.51 Z= -3.771, p<0.05	Post-test mean = 2.87 Delayed post-test mean = 2.51 Z= -2.037, p<0.05

For both items, the score difference between the pre- and post-test was found to be statistically significant. The same was valid for the score difference between the pre- and the delayed post-test score. In both of these cases, we recorded a statistically significant improvement in students' reasoning. However, the comparison between the post- and the delayed post-test scores showed a statistically significant retreat: the impact of the learning environment appeared to start fading with time.

4. Discussion

Although students' progress appeared to retreat in the 'post/delayed post' comparisons for both items, the contribution of our anti-BON learning environment was still evident after a whole year with no ecology at all. So, in the delayed post-test the 'same picture' prediction for a protected ecosystem (item 1) almost maintained the much decreased frequency of the post-test. This was also valid for the 'full recovery' prediction in the case of a disturbed ecosystem (item 4): in the delayed post-test, only very few students went back in claiming that a disturbed ecosystem in a process of restoration by humans will 'fully recover' its initial state. More interestingly, the 'contingent/unpredictable picture' prediction for a terrestrial or aquatic ecosystem under human protection (item 1), which reached its highest frequency in the post-test, appeared in a lower but still notable frequency one year later. Again, the same was valid in the case of a disturbed ecosystem (item 4); in the delayed post-test, several students *did* claim that a disturbed ecosystem will have an unpredictable response to human effort to restore its initial state.

Thus, our learning environment appeared to challenge effectively the idea

of stable, 'never changing' nature in the case of protection, as well as the idea of 'always recovering' nature in case of disturbance. However, although it had students consider nature's flux through the assumptions of a 'resilient nature' view, it seems that it did not lead them to *firmly* abandon their *certainty* about an ecosystem's response to protection or disturbance. So, the 'different picture' prediction for a protected ecosystem (item 1), which temporarily retreated in the post-test, found its way back one year later. The same was valid with regard to the 'no recovery' prediction for a disturbed ecosystem: although its post-test frequency was substantially decreased, in the delayed post-test the prediction appeared as frequently as in the pre-test. Considering also the fact that the 'contingent/unpredictable behaviour' prediction was still present in the delayed post-test but less popular than in the post-test itself, one could claim that students *did* build the notion of the ecosystems' contingent behaviour, but not as solidly as we would like.

As expected, destabilizing rather permanently the idea of the 'balance of nature' proves to be hard. This idea is very widespread, while the 'randomness-non-purpose' alternative seems to be strongly counter-intuitive, not just on the conceptual level but possibly on the emotional one as well. It may be that the 'resilient nature' idea challenges students' spirituality or their relationship to nature as a whole, and dealing with this issue too, could possibly enhance further the effectiveness of our learning environment. Nevertheless, since a remarkable part of students' learning gains was still present one year after their engagement with it, we argue that it was a successful one. Even if we take into account that simulation models are just artificial representations of nature and not nature itself, giving students the opportunity to familiarize with the 'resilient nature' idea through them seems to be very important. Thinking about ecosystems in the light of contingency may contribute to a better conceptual understanding as well as to better decisions with regard to environment.

Acknowledgments

References

Akker, J. V. D., Gravemeijer, K., McKenney, S. & Nieveen, N. (Eds.) (2006) *Educational Design Research*. Oxon, England: Routledge.

Ampatzidis, G. & Ergazaki, M. (in press) Using ecology to enhance everyday reasoning: the case of interdependent and reciprocal causality. *Review of Science, Mathematics and ICT Education.*

Cooper, G. (2001) Must There Be a Balance of Nature? *Biology and Philosophy*, *16*(4), 481–506.

Cuddington, K. (2001) The "Balance of Nature" Metaphor and Equilibrium in Population Ecology. *Biology and Philosophy*, *16*(4), 463–479.

Driver, R., Asoko, H., Leach, J., Scott, P. & Mortimer, E. (1994) Constructing Scientific Knowledge in the Classroom. *Educational Researcher*, *23*(7), 5–12.

Ergazaki, M. & Ampatzidis, G. (2012) Students' Reasoning about the Future of Disturbed or Protected Ecosystems & the Idea of the "Balance of Nature." *Research in Science Education*, *42*(3), 511–530.

Ergazaki, M. & Zogza, V. (2008) Exploring lake ecology in a computer-supported learning environment. Journal of Biological Education, 42(2), 90-94.

Gibbs, G.R. (2005) *Qualitative Data Analysis: Explorations with NVivo*. Maidenhead, England: Open University Press.

Gunderson, L.H., Allen, C.R. & Holling, C.S. (Eds.). (2010) *Foundations of Ecological Resilience*. Washington, DC: Island Press.

Holling, C.S. (1973) Resilience and Stability of Ecological Systems. *Annual Review of Ecology and Systematics*, *4*(1), 1–23.

Hovardas, T. & Korfiatis, K. (2010) Towards a Critical Re-Appraisal of Ecology Education: Scheduling an Educational Intervention to Revisit the "Balance of Nature" Metaphor. *Science & Education*, *20*(10), 1039–1053.

Jelinski, D.E. (2005) There is no mother nature-there is no balance of nature: culture, ecology and conservation. *Human Ecology*, *33*(2), 276–285.

Klaassen, C.W.J.M. (1995) *A problem-posing approach to teaching the topic of radioactivity*. Ultrecht, Netherlands: CD-ß Press.

Ladle, R.J. & Gillson, L. (2009) The (im)balance of nature: a public perception time-lag? *Public Understanding of Science*, *18*(2), 229–242.

Richmond, B. (2004) *An Introduction to Systems Thinking*. Lebanon, NH: isee systems.

Rutten, N., van Joolingen, W.R. & van der Veen, J.T. (2012) The learning effects of computer simulations in science education. *Computers & Education*, *58*(1), 136–153.

Scheffer, M. (2009) *Critical Transitions in Nature and Society*. Princeton, NJ: Princeton University Press.

Schmitz, O. (2010) *Resolving Ecosystem Complexity*. Princeton, NJ: Princeton University Press.

Smetana, L.K. & Bell, R.L. (2012) Computer Simulations to Support Science Instruction and Learning: A critical review of the literature. *International Journal of Science Education*, *34*(9), 1337–1370.

Westra, R. (2008) *Learning and teaching ecosystem behaviour in secondary education*. Castricum, Netherlands: Faculteit Betawetenschappen.

Wilensky, U. (1999) NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL.

Zimmerman, C. & Cuddington, K. (2007) Ambiguous, circular and polysemous: students' definitions of the BON-metaphor. *Public Understanding of Science*, *16*(4), 393–406.

Appendix: The pre-post questionnaire

Item 1: Protected ecosystem
Pre-test: Consider a forest that hosts specific populations of plants, animals and decomposers. This forest is a 'national park' – namely a totally protected area where human activities are not allowed. The size of the populations is regularly monitored by the park's scientists who study their course in time. In the absence of natural disturbances, how do you think this forest will look some years later, compared to how it initially looked according to the given description?

Post-test: Consider an aquatic park that hosts specific populations of plants, animals and decomposers. This aquatic park is a totally protected area where any human activities are not allowed. The size of the populations is regularly monitored by the park's scientists who study their course in time. In the absence of natural disturbances, how do you think this park will look some years later, compared to how it initially looked according to the given description?

Item 4: Disturbed ecosystem/Abiotic change
Pre-test: Consider a lake that hosts specific plant populations (phytoplankton, water plants), animal populations (zooplankton, fish, sea birds) and populations of decomposers. The city council of a nearby city decides to throw treated, non-toxic wastewater in the lake for some time. This increases the water-salinity and, as a result, some populations decrease in size, others increase, and a fish population dies out. Sometime later, the water salinity is restored to its initial level due to the efforts of a group of scientists, and the fish population which died out is reintroduced. How do you think this lake will look some years later, compared to how it looked before the salinity-increase according to the given description?

Post-test: Consider a lake that hosts specific plant populations (phytoplankton, water plants), animal populations (zooplankton, fish, sea birds) and populations of decomposers. The city council of a nearby city decides to throw nutrients in the lake for some time and, as a result, some populations decrease in size, others increase, and one fish population dies out. Sometime later, the nutrients of the lake are restored to their initial level due to the efforts of a

group of scientists, and the fish population which died-out is reintroduced. How do you think this lake will look some years later, compared to how it initially did before the nutrient-increase according to the given description?

Teachers' conceptions of environment in five Sub-Saharan African countries

Pierre Clément[1,2], Jérémy Castéra[1], Lawrence Ntam Nchia[3], Laurence Ndong[4], Ivette Béré-Yoda[5], Mame Seyni Thiaw[6] and Mensan Azadzi Dzamayovo[1,7]

clement.grave@free.fr

Abstract

This work attempts to identify teachers' conceptions of Environment in five sub-Saharan African countries: it analyses their similarities and differences; compares them to those of French teachers; and explores if they are influenced by other parameters such as gender, religion, or level of instruction. We used a questionnaire built and validated by the research project *Biohead-Citizen*. Similar samples of teachers filled out the questionnaire anonymously in each country (2144 teachers). Principal Component Analysis shows three poles of values: anthropomorphic; anthropocentric and ecocentric. A strong positive correlation was established between Pro-Genetic Modified Organism (GMO) opinions and anthropocentric values. Between-class analyses and randomization, tests show significant differences among these countries. While most of the teachers' conceptions are ecocentric, there are significant differences among the five African countries. In Togo, where animism is more frequent, teachers are more anthropomorphic. In Cameroon, English-speaking teachers are more anthropocentric and more pro-GMO than their French-speaking colleagues. In Burkina Faso, where French humanitarian help is particularly important, teachers' conceptions are nearer to the French ones. Analysis shows other differences related to the teachers' level of instruction,

[1] Aix-Marseille Université, EA4671 ADEF, ENS de Lyon (Aix-Marseille), France

[2] Honorary in University of Lyon 1, France

[3] ENS Yaoundé, University of Yaoundé I, Cameroon

[4] LARED, ENS de Libreville, Gabon

[5] MESS/DGIFP, Ouaga 01, Burkina-Faso

[6] Faculté des sciences et technologies de l'éducation et de la formation (FASTEF), Dakar, Senegal (died on the 29-01-2017)

[7] Institut national des sciences de l'éducation (INSE), University of Lomé, Togo

or their gender. The specificities of these results could inform stakeholders in designing and implementing environmental education in these African countries.

1. Introduction and research questions

Environmental Education (EE) for Sustainable Development is promoted by UNESCO and by most of the national politics of education and curricula in Africa (UNESCO, 2009a, 2009b; Clément & Caravita, 2011). Teachers are key actors to implement it, and research has analysed teachers' conceptions of environment in several countries (16 countries in Munoz et al, 2009; 24 countries in Clément & Caravita, 2012).

These conceptions are generally characterized by two poles: anthropocentrism and ecocentrism. Anthropocentrism is '*human-centred*', conferring intrinsic value mainly to humans (Callicott, 1984; Larrère, 1987). Ecocentrism is '*ecosphere-centred*', with empathy for nature and emphasising the intrinsic value of the interrelated ecological systems (of which humans are a part). These two poles of attitudes are respectively called "*Utilization*" and "*Preservation*" in the *2-MEV* model (Model of Ecological Values: Wiseman & Bogner, 2003; Bogner & Wiseman, 2004; Munoz et al, 2009; Milfont & Duckitt, 2010). In this model, the two values are independent as it is possible to be at the same time anthropocentric (value Utilization) and ecocentric (value Preservation). The goal of environmental education is to increase the Preservation pole (Bogner & Wiseman, 2004).

In the international comparisons of teachers' conceptions of environment done so far, West or Central Francophone African countries have not be exploited. This present work involves five of these countries (Table 1): three in West Africa (Burkina-Faso, Senegal and Togo) and two in the West part of Central Africa (Cameroon and Gabon). In all these countries the climate and natural landscape are characteristic of an inter-tropical zone.

In this work we specifically added a third pole, "*anthropomorphism*", which holds that any animal can feel and think as human beings do (Quinn et al.,

2016), initially referred to as "*sentimento-centred*" attitude (Khalil et al., 2007; Clément et al., 2011). While researchers in ethology and in cognitive sciences agree that homoeothermic vertebrates (mammals and birds) can feel pain or happiness, it is still under discussion regarding amphibians, fish, or even cephalopods (Chapouthier, 2016); but there is no work to prove any perception of pain or happiness in other Invertebrates (Broom, 2013; Chapouthier, 2016). Nevertheless, some people are convinced they can - being anthropomorphic. Thus we included in our questionnaire questions related to the happiness of flies and of snails, where behaviours are mainly reflexes, with only primitive competences of short memory (Clément et al., 1999).

Given that animist beliefs are still strong in sub-Saharan countries, even when they are juxtaposed with Christian or Muslim beliefs (Sanogo & Coulibaly, 2003), it was interesting to know if animism is linked to more anthropomorphic attitudes in these countries. The environmental beliefs of these people affect the way they exploit their resources and this can be reflected in the Gross Domestic Products (GDP) and economic growth. Table 1 compares the population, GDP and growth of these five countries and France. These African countries are economically poor. In four of them, the GDP / Inhabitant is very low: from $578 in Togo to $1,234 in Cameroon; it is higher in Gabon ($8,580) due to high exports of oil, wood, uranium, manganese for less than one million of inhabitants, with strong inequality among these inhabitants.

Table 1 Summary figures for the five African countries, and France for comparison.

Country	Population (Millions of Inhabitants)	GDP / Inhabitant ($) (Gross Domestic Product)	Growth (2015)
BF BURKINA – FASO	18.5	631	5.0%
SN SENEGAL	14.7	934	5,1%
TG TOGO	7.2	578	5.4%
CM CAMEROON	23.7	1 234	5.3%
GA GABON	0.8	8 580	3.5%
FR FRANCE	64.4	37 728	1.2%

We also added France in this comparison (Table 1), because there are strong economic and cultural links between France and these five African countries. Their educational systems, including most of their textbooks, are influenced by the French educational system, for example, nearly all the official textbooks used in these five countries dealing with environment (Biology,

Geology and Geography) in Secondary Schools (students from 11 to 18 years old), are published by French publishers (Hatier, Bordas, Nathan, Hachette, Belin). Those published by African authors are used in several of these Francophone sub-Saharan Countries, for instance, in Senegal, the textbooks analyzed by M. S. Thiaw on Environment during the Biohead-Citizen project (Carvalho et al., 2008) were written by R.T. Djakou, National Inspector in Cameroon and by S.Y. Thanon, Chief Inspector in Ivory Coast (Djakou & Thanon, 1986, 1996, respectively for students 17-18 and 14-15 years old). These textbooks are currently in use in Togo, as well as in several countries of inter-tropical Africa (for instance Gérenton, 1993, 1995, respectively for students 11-12 and 12-13 years old). Very few textbooks are specific to only one of these African countries. For instance, in Togo, only one textbook, for students 14-15 years old, includes a large part specific to Togo (Géographie, classe de 3ème, L'Afrique Occidentale, le Togo. Hatier, Limusco). In Cameroon, a textbook (Nditafon, 2006) delivering Environmental Education in the formal sector, is in use. In Gabon, the syllabi and textbooks are totally similar to the French ones (cf. for instance Ndong, 2012). Thus, the teaching of environmental education in these countries is strongly influenced by the conception of French authors that the African teachers try to contextualise.

An important trend of research in science education is to analyse the teachers' conceptions as a basis to improve their practices as well as their training (Bruguière et al., 2014, p.5). These conceptions can be analysed as possible interactions between the three poles KVP (Clément, 2006; 2010): K for scientific knowledge, V for values (as Preservation and Utilization in the 2-MEV model) and P for practices (for environment, against pollutions for instance, but also pedagogical practices, as the choice of textbooks and other resources to teach Environmental Education.). French teachers and specialists are often invited to train the teachers in these African countries, given that the insertion of EE in the curriculum as a standalone subject is recent as it was initially taught through Biology, Geography or Geology.

From the above background and literature review, we have developed the following research questions:

- What are the teachers' conceptions of environment in these five African countries, and do they differ among these countries?

- Do the conceptions on the environment of these teachers of sub-Saharan Africa differ from those of their French counterparts?

- Are there differences related to other controlled factors such as: teachers' religion, their level of instruction, or their gender? For instance, ecofeminism claims that women are more concerned about nature and environment than men (d'Eaubonne, 1974 ; Zell, 1998). In consequence, women would be more ecocentric and less anthropocentric than men. Is this verified in these five African countries?

There is a link between these research questions: practising environmental education can be improved when the teachers are more aware of their own conceptions of environment, and of the diversity of these conceptions depending on several parameters.

2. Methodology

We used the questionnaire built and validated by the Biohead-Citizen project (2004-2008: Carvalho et al., 2008). Environmental Education is one of the six topics of this questionnaire, the others being Health Education, Sex Education, Evolution, Human Genetics, Human Brain. We analyze here the teachers' response to 23 statements related to Environment: 8 for ecocentric values; 7 for anthropocentric values; 5 related to opinions about GMO (genetically modified organisms); and 3 related to anthropomorphism. We also use 17 questions related to each teacher's characteristics: gender, age, level of training, religion, political or religious opinions.

A previous international comparison was published for 16 countries (Munoz et al, 2009), including only one sub-Saharan country (Senegal). For the present work, the countries involved are Senegal (SN: N=324, data collected in 2007), Burkina Faso (BF: N=296, data collected in 2008), Cameroon (CM: N=523, data collected in 2010), Gabon (GA: N=269, data collected in 2013) and Togo (TG: N=270, data collected in 2016). The total sample is N = 2,144 teachers. Given that differences among these countries can result from the different dates of collection of these data, we did comparisons of pre-service data collected between 2006 and 2016 in two of the countries involved in the Biohead-Citizen project, and found no differences related to the topic of Environment in Lebanon (Abou-Tayeh et al., 2015, communication in IOSTE

2015) or in Morocco (Agorram et al., 2016, not yet published). Thus, the eventual differences will be more related to differences between countries than to differences between the times of collection of data.

Similar samples were used in each country, that is, six samples, each of about 50 teachers (more in Cameroon to compare two different subsystems of education): In-service Primary school teachers; Pre-service Primary school teachers; In-service Biology teachers in Secondary schools; Pre-service Biology teachers; In-service Language teachers in Secondary schools; and Pre-service Language teachers. In each country, groups of teachers filled out the questionnaire in the presence of a researcher, who immediately collected the completed questionnaire, with special attention to ensuring total anonymity.

The answered questionnaires were coded, then analysed using the software "R" to do multivariate analyses, which are well adapted to interpret this kind of data (Munoz et al, 2009). To identify the teachers' conceptions, we used PCA (Principal Components Analysis) as a qualitative approach to first characterize the main features of the data. To answer the other research questions, we used between-class analyses (Dolédec & Chessel, 1987; Munoz et al., 2009) to investigate whether these conceptions differed among defined groups (such as countries, level of instruction, gender). To know if the differences are significant or not, we used a randomization test (Monte Carlo permutation procedure: Romesburg, 1985; Manly, 1997). The differences between the teachers' answers to each question were also evaluated using a Pearson's Chi-squared test.

3. Results

3.1 African teachers' conceptions of environment

(1a)
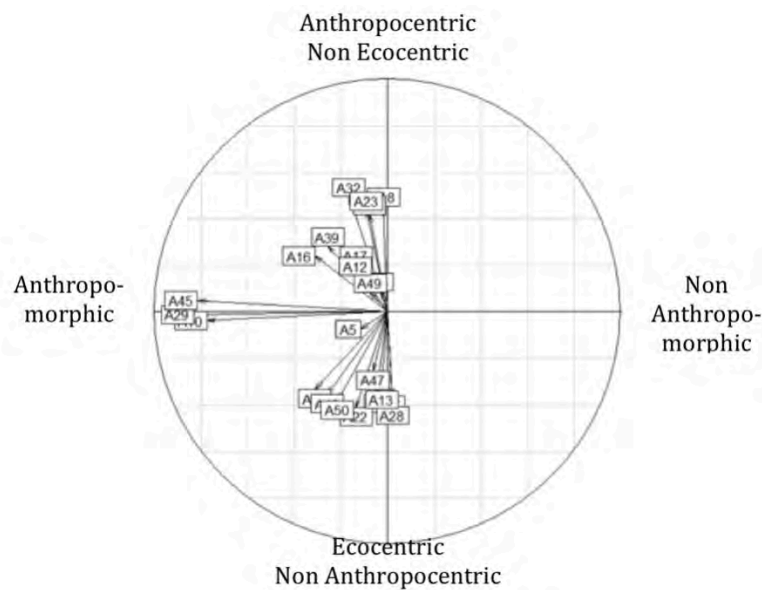
(1b)
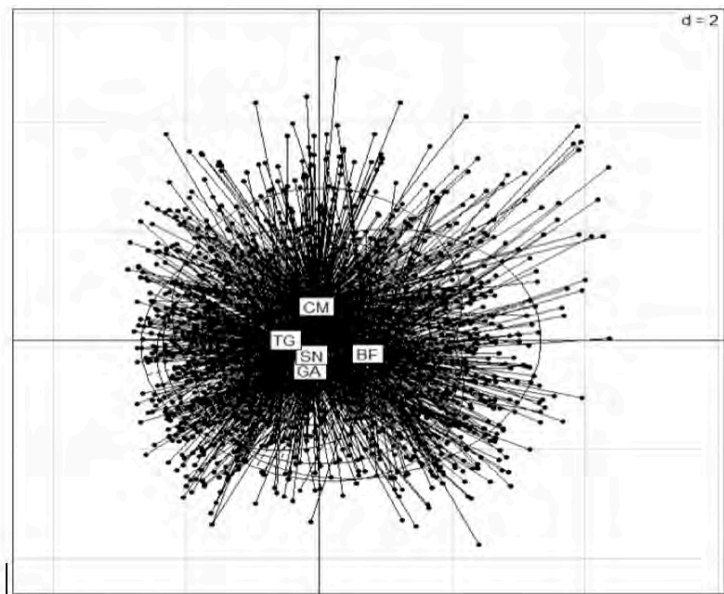
Figure 1 PCA from the teachers' answers in the 5 African countries (N=2144 teachers).

The PCA shows that the teachers' answers mainly propose anthropomorphic as opposed to non-anthropomorphic attitudes (First Principal Component - horizontal axis of Figures 1a and 1b). On the left of the horizontal axis, teachers think a snail, a fly or a frog can feel happiness (anthropomorphic) while those on the right end belief they cannot. Figure 1b shows that there is a great diversity of anthropomorphic values inside each country. Nevertheless, there are some differences among countries, the teachers being more anthropomorphic in Togo, and less anthropomorphic in Burkina Faso.

The second principal component (vertical axis in Figure 1) is related to the other environmental values. It is orthogonal to the horizontal axis, showing that anthropomorphic attitudes are mainly not correlated to anthropocentric or ecocentric values. This vertical axis shows an opposition between anthropocentric and ecocentric values, which is partly a consequence of the great weight of anthropomorphism: after suppressing the 3 items related to anthropomorphism, the first component of this new PCA is Utilization (anthropocentric values) and the second one is Preservation (ecocentric values), that is in conformity with the 2-MEV model. That is also visible in Figure 1a, where some vectors are at the left of the vertical axis: several teachers who agreed with some anthropocentric propositions (value Utilization) such as A16 (*Our planet has unlimited natural resources*) and A17 (*Society will continue to solve even the biggest environmental problems*), also agreed with some ecocentric propositions (value Preservation) such as A11 (*Industrial smoke from chimneys makes me angry)*, A40 (*It is interesting to know what kinds of animals live in ponds or rivers*), and A50 (*All contemporary plant species should be preserved because they may help in the discovery of new medicines*).

The Pro-GMO attitudes are correlated with the Utilization pole (anthropocentric values), and the Anti-GMO attitudes with the Preservation pole (ecocentric values), except the responses to statement A49 which are more related to scientific knowledge (*"If a person eats genetically modified plants, his/her genes can be modified"*).

Figure 1b shows that, inside each country, there is a great diversity of answers with some trends differentiating the five countries, teachers in Cameroon being the most anthropocentric and Pro-GMO, while their colleagues in Gabon are the least.

3.2 Differences among the countries

A between-class analysis, completed by a randomisation test (Monte Carlo), shows very significant differences among the five African countries (p < 0.001). The questions differentiating these countries the most are mainly related to anthropomorphism, anthropocentrism, and attitudes related to GMO. When we consider France alongside these five African countries, there is a spectacular difference between the French teachers' answers and those of their African colleagues: most of the French teachers are less anthropomorphic and less anthropocentric than their African colleagues. Figures 2 to 5 illustrate some of the differences among countries.

Anthropomorphism

In the five African countries, teachers are more anthropomorphic than in France where only 1/4 of French teachers think that flies, snails or frogs are able to feel happiness (Figure 2 for flies). Nevertheless, there are significant differences among the five African countries: from 38% of teachers in Burkina Faso to 70% in Togo (55% to 60% in Cameroon, Gabon & Senegal) in thinking that these animals can feel happiness (Figure 2 for flies: Pearson's Chi-squared test p-value < 0.001). The answers are very similar for snails (A10) and for frogs (A29).
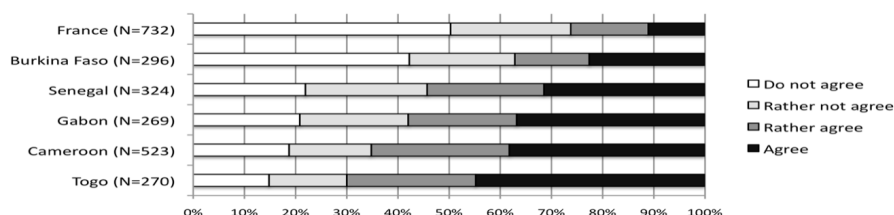
Figure 2 Teachers' answers, grouped by country, to statement A45: "Flies are able to feel happiness".

Anthropocentrism

The difference between France and these African countries is very clear, as illustrated by the Figure 3 for statement A16 (*"Our planet has unlimited natural resources" (with France for comparison"*): only 5% of French teachers agree or rather agree, compared to 39% in Burkina Faso and 58% in Togo and Cameroon. The differences among the five African countries are significant (Pearson's Chi-squared test: p-value < 0.001): Burkinabe teachers are less anthropocentric than their colleagues in the four other African countries, while teachers

are more anthropocentric in Togo and Cameroon.

Nevertheless, for some other questions related to anthropocentrism, a large majority of African teachers are not anthropocentric, even if a little more in Cameroon than in the four other countries.
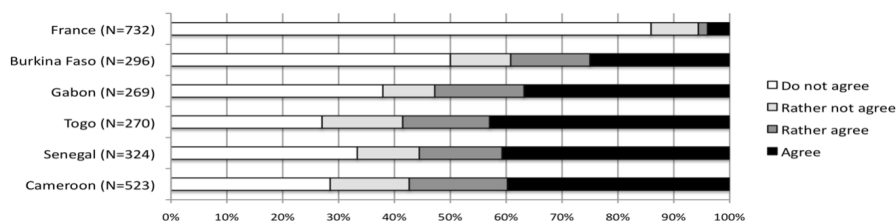
Figure 3 Teachers' answers, grouped by country, to statement A16 "Our planet has unlimited natural resources" (with France for comparison).

Pro or anti-GMO attitudes

Here also there are significant differences among the five African countries (Pearson's Chi-squared test: p-value < 0.001). Figure 4 illustrates these for statement A12. As already shown in Figure 1b, teachers in Cameroon are more pro-GMO than their colleagues of other countries: 68% of them agree or rather agree with proposition A12 *"Genetically modified plants will help to reduce famine in the world"*. In Gabon the teachers are less pro-GMO: 50% of them agree or rather agree with proposition A12. Nevertheless, this opinion is drastically different from those in France where only 35% of the teachers agree or rather agree (Figure 4).
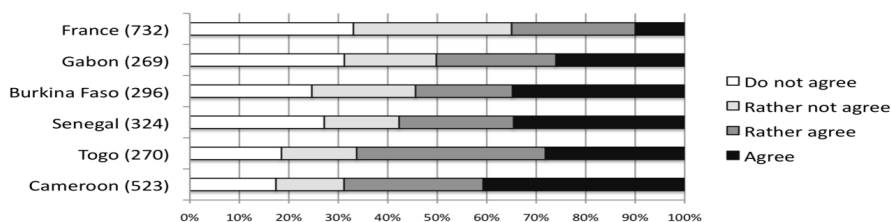
Figure 4 Teachers' answers, grouped by country, to statement A12 "Genetically modified plants will help to reduce famine in the world".

Ecocentrism

The items related to ecocentric values are less differentiated among the five African countries. For most of the items, a majority of teachers are ecocentric.

For instance, responding to statement A22 *"I enjoy trips in the countryside"*, most of teachers, from 88% in Burkina Faso to 97% in Senegal or in Togo, ticked I agree or I rather agree, as in France (96%). The difference among the five African countries is significant, but not so important. However, with respect to statement A28 *"It makes me sad to see the countryside taken over by building sites"*, only about half of African teachers agreed or rather agreed (from 48% in Cameroon to 71% in Gabon), compared to 85% in France. This item probably has a different meaning in less economically developed countries.

Differences between two regions in Cameroon
A peculiar situation was reported in Cameroon where half of the sample was English speaking. According to the results of complementary analyses restricted to Cameroon that were recently presented as a communication in IOSTE 2016 (Nchia et al., 2016), significant differences between French and English speaking regions of Cameroon were established, with the Francophone teachers' conceptions being less anthropocentric and less pro-GMO than their Anglophone colleagues.

3.3 Differences linked to other controlled parameters such as the level of training and the gender
Although **the religious practices** are very different in the five African countries in our sample, being Muslim or Christian (Catholic or Protestant), sometimes agnostic or atheist or even animist, we did not find any significant influence of the various religious groups on teachers' conception of the environment.

 Comparing the six samples (pre- or in-service teachers, Primary or Secondary schools teaching biology or language), the differences are clearly significant: Primary teachers (pre-service as well as in-service) are more anthropocentric and more pro-GMO than secondary teachers. This difference is linked to **the level of instruction of teachers**. When they are trained during 1 or 2 years at University, they are more anthropocentric and more pro-GMO than their colleagues trained during 3 years, 4 years or more at University. Figure 5 illustrates this difference for statement A16 *"Our planet has unlimited natural resources"*.
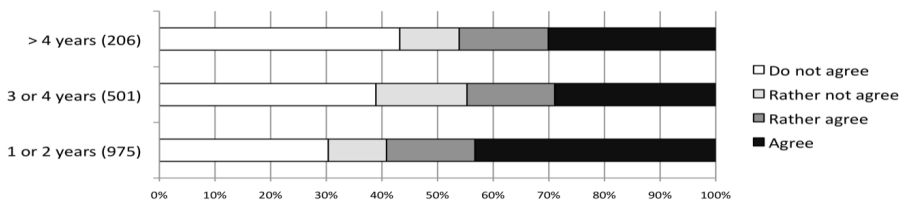
Figure 5 African teachers' answers, grouped by level of instruction, to statement A16 "Our planet has unlimited natural resources".

A significant and unexpected **gender effect** was found. For the total sampling of the five African countries, female teachers believe more than their male colleagues that our planet has unlimited resources (Figure 6). Fewer female teachers think that "*Humans will die out if we don't live in harmony with nature*" (A7) (80% agreed or rather agreed compared to 85% of male teachers: p-value = 0.012). Fewer female teachers agreed or rather agreed with "*Genetically modified organisms are contrary to nature*" (A13) (58% of females compared to 64% of males: p-value = 0.004)
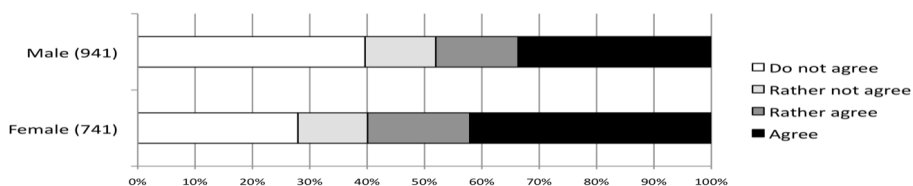
Figure 6 African teachers' answers, grouped by gender, to statement A16 "Our planet has unlimited natural resources" (p-value < 0.0001).

4. Discussion

Our results show a diversity of conceptions inside each country (Figure 1b), with nevertheless some trends and differences that we are going to discuss. These conceptions partly confirm the 2-MEV model (Bogner & Wiseman, 2004, Munoz et al., 2009) with two main environmental values, Utilization and Preservation, which are not opposed for several teachers. Nevertheless, for some other teachers, these values are opposed, ecocentric values being correlated with non-anthropocentric values. Moreover, we showed a significant link between the pro-GMO attitudes and anthropocentric values. Until now, the relationship between the 2-MEV values and the attitudes related to GMO was observed only at some national levels as in Lebanon (Khalil et al., 2007),

Poland (Clément et al., 2011), and more recently Australia (Quinn et al., 2016). It is interesting to find it in African countries, which until now had not been investigated.

4.1 Differences between French and African teachers' conceptions of environment

There are important differences between France and the five African countries. The main ones are related to anthropomorphism, with a clear opposition between France and the African countries in our sample. Anthropomorphism is higher in Africa. It is rooted in the African animism, which is influential even if combined with Christian or Muslim religious beliefs and practices (Sanogo & Coulibaly, 2003). For animism, any living being has a soul, and in consequence can suffer or be happy. Thus, Africans take care of nature and environment, understand ecological interactions, and live in harmony with their environment which is a characteristic of ecocentric values (Sanogo & Coulibaly, 2003). That is coherent with our results related to large ecocentric conceptions of African teachers. In contrast, most French teachers are not anthropomorphic, expressing a Cartesian tradition of seeing animals as machines, already shown when comparing some European countries (Forissier, 2003). The roots of French ecocentrism are not the same as in Africa. According to Mazrui & Wagaw (1986), traditional African values are sociocentric, theocentric and ecocentric, while the French ones are mainly egocentric and ecocentric.

In the five African countries, as well as in France, the large ecocentrism of teachers is also probably a result of the positive impact of the ecological approach to environmental education based on the principle of Sustainable Development promoted by UNESCO (2009a). This large ecocentrism of teachers contradicts the apparent problems of preservation in these countries, particularly in Africa where it is frequent to find a lot of plastic bags or other waste littering the natural landscape (e.g. Bassole, 2016). Three hypotheses can be proposed to interpret this apparent contradiction: (1) being more educated, teachers are more sensitive to preservation than the general population; (2) environmental education for sustainable development is relatively new in West and Central Africa, with limited effects on young people and not yet on the large older population; (3) possibly, the declarative responses of teachers do not correspond to their real practices. Complementary research is necessary to test these hypotheses.

Our results show differences linked to the teachers' level of instruction: the higher the level, the less the teachers' conceptions are anthropocentric and pro-GMO (Figures 4 & 5).

These two sets of values clearly differentiate French and African conceptions of environment, with some possible complementary explanations:

(1) As indicated in Table 1, the Gross Domestic Product/Inhabitant is very much higher in France than in African Countries. In Africa, food insufficiency, as well as access to potable water (WHO/UNICEF, 2015), are still major problems, with a large part of agricultural products being exported to developed countries. In sub-Saharan Africa, a significant part of the population is still undernourished, according to the United Nations Food and Agriculture Organization (FAO, 2015): 7% in Gabon, 8% in Cameroon, 11.5% in Togo, 11% in Senegal, 20% in Burkina Faso.

(2) In many African countries, agriculture still occupies more than 2/3 of the working population[8] while only 2.8% in France[9].

(3) The political contestation of GMO is higher in Europe, particularly in France (e.g. Foucart & Thiberge, 2016; Horel, 2016; Garric, 2013); and the multinational firms selling GMO are still influential in these African countries. However, some anti-GMO demonstrations have already taken place in Burkina Faso (Van Eeckhout, 2015).

All these parameters seem to have a greater impact on the African teachers' conceptions than the similarities in the syllabi and textbooks used in France and sub-Saharan African countries presented in the introduction of this paper. These similarities could account mainly for the largely ecocentric teachers' conceptions in all these countries, even if, as noticed above, the roots of these ecocentric conceptions can be partly different in France and in these African countries.

4.2 Differences among the five African countries

The differences between the five African countries are not easy to explain. There is no clear difference between the three West African countries (Burkina-Faso, Senegal and Togo) and the two Central African ones (Cameroon and Gabon), other than the Gross Domestic Product/Inhabitant.

However, there are two main differences among these five African coun-

---

[8] http://www.fao.org/docrep/V4870F/V4870F04.htm

[9] INSEE 2016 : https://www.insee.fr/fr/statistiques/1906677?sommaire=1906743

tries. Firstly, the Burkinabe teachers' conceptions are the most similar to those of French teachers, whereas Burkina Faso is poorer than Gabon, Cameroon or Senegal. A possible explanation could be the large French humanitarian aid to Burkina-Faso for several years through non-governmental organisations (NGOs), as well as direct help to the Burkinabe system of education. "*Burkina Faso is a place of predilection for non-governmental organizations (NGOs) and other humanitarian organizations*"[10]. Today in Burkina Faso 35 humanitarian Associations are listed, 28 in Senegal, 16 in Togo and only 9 in Cameroon[11].

Secondly, in Cameroon and in Togo, the teachers' conceptions are more anthropomorphic, more anthropocentric and more pro-GMO than in the three other African countries.

In Togo, the high proportion of anthropomorphic conceptions is related to the highest amount of animist beliefs: 34% of the population are of traditional African religions, while 1% in Senegal, 4.9% in Gabon, 11.4% in Cameroon, 15% in Burkina Faso[12]. Moreover, in Togo, Christian and Muslim believers often participate in traditional African rituals and customs[13].

In Cameroon, half teachers in our sample are English speaking. Compared to the half French speakers, the difference is significant: more anthropocentric and more pro-GMO (Nchia et al., 2016, communication in IOSTE 2016, Braga). This difference is probably not a consequence of the respective languages, but of the French or English systems of education, which respectively influence these two regions of Cameroon; but we need more investigation to compare these two systems of education and their eventual influence on teachers' conceptions of environment. Until now, we did not find significant difference between English and French teachers related to their conceptions of nature and environment (Clément & Caravita, 2012).

4.3 A specific gender effect

As mentioned above, eco-feminist women are expected to be more ecocentric and less anthropocentric than men. Nevertheless, McEwen et al. (2015), using the Biohead-Citizen questionnaire, found no gender difference when com-

---

[10] http://www.urgenceafrique.org/sites/default/files/presse-rapports-files/rapport2012_urgence_afrique_r.pdf

[11] http://associations-humanitaires.blogspot.fr/p/blog-page_5.html

[12] http://www.liberte-religieuse.org/fiches-pays/afrique/ (15-7-2017)

[13] http://www.liberte-religieuse.org/togo/ (15-7-2017)

paring male and female teachers' conceptions of environment in Sweden and France.

Our results illustrate the contrary of the eco-feminist predictions: in the five sub-Saharan African countries, for some items the female conceptions significantly differ from those of their male colleagues, being less ecocentric, more anthropocentric and more pro-GMO. That is not a surprise considering the fundamental role of women in agriculture in these countries. According to a report of the World Bank and the Food and Agriculture Organization of the United Nations (FAO), in sub-Saharan Africa, women produce up to 80 per cent of food for household consumption and for sale in local markets (Ben-Ari, 2014). "*In the case of crops such as rice, wheat and maize, which account for about 90% of the food consumed by rural dwellers, it is mainly women who plant seeds, take care of weeding, grow and harvest and sell the surplus*" (Ben-Ari, 2014). For secondary crops (legumes and vegetables, for example), FAO reports that "*Women's contribution ... is even more importan*t". In consequence, it is normal that African women have more conceptions of environment related to production, utilisation, and anthropocentric perspectives. The notion of eco-feminism comes mainly from Western developed countries and largely ignores the hard life of African women as mothers, housekeeping and main workers on the farms.

5. Conclusions

It is with caution that environmental education in sub-Saharan Africa should use the resources forged in France, such as school textbooks or teacher training courses. Our results show that some specificity must be taken into account.

Eco-feminism is one of the gaps differentiating the Western and African ways of thinking and living. Our results show other differences between France and sub-Saharan Africa. An important one is anthropomorphic attitudes rooted in African animism. For Nyéréré (1972), Mazrui & Wagaw (1986) and Ki-Zerbo (1990), the Western-style teaching in sub-Saharan Africa removes "school children from African norms and values, cultivating contempt for traditions and promoting elitism and individualism in a society where Community life is a tradition. Hence the question: how can animism as an African cultural, philosophical and religious background contribute to forging the strategies and tools of African development? By teaching values of respect for nature and all that it contains, animism is preaching for ecology which

has become an obvious fact for all developed and non-developed countries" (Sanogo & Coulibaly, 2003). We don't think that environmental education, in these African countries, must promote animism, but it has to take into account the specificities we started to identify in this paper: differences between France and these sub-Saharan countries, and also the specificities of each of these African countries, which need to be further investigated.

References

Abou-Tayeh, P., Clément, P. & Khalil, I. (2015) Is there an evolution, from 2006 to 2015, of the Lebanese pre-service teachers' conceptions related to questions about biology, health and environment? Communication in *IOSTE Eurasian Regional Symposium* (24-26 April 2015), Bahçe ehir University, Istanbul, Turkey.

Bassole, H. F. (2016) Elimination des sachets plastiques au Burkina Faso, où en est-on avec le projet ? http://www.mediaterre.org/afrique-ouest/actu,20160718183908.html

Ben-Ari, N. (2014) L'agriculture, une affaire de femmes. *Afrique renouveau*, Édition Spéciale Agriculture 2014, http://www.un.org/africarenewal/fr/magazine/%C3%A9dition-sp%C3%A9ciale-agriculture-2014/l%E2%80%99agriculture-une-affaire-de-femmes

Bogner, F. & Wiseman, M. (2004) Outdoor ecology education and pupils' environmental perception in preservation and utilization. *Science Education International*, 15(1), 27-48.

Broom, D.M. (2013) Le bien-être des invertébrés : insectes, araignées, escargots et vers. In T. Auffret Van des Kemp & M. Lachance (dir.), *Souffrance animale. De la science au droit*. (pp. 155-175).Cowansville, Québec: Yvon Blais,.

Bruguière, C., Tiberghien, A. & Clément, P. (2014) Part I, Overview of the book. In Bruguière C., Tiberghien A. & Clément P. (eds.), *Topics and Trends in Current Science Education* (pp. 3-18). Netherlands: Springer.

Callicott, J. B. (1984) Non-anthropocentric value theory and environmental ethics. *American Philosophical Quarterly*, 21(4), 299-309.

Carvalho, G., Clément, P., Bogner, F. & Caravita, S. (2008) *BIOHEAD-Citizen : Biology, Health and Environmental Education for better Citizenship*, Final Report. Brussels : FP6, Priority 7, Project N° CITC-CT-2004-506015

Chapouthier, G. (2016) Souffrances animales. In K. L. Matignon (dir.), *Révolutions animales* (pp. 197-209). Paris : Arte Éditions / Éditions les liens qui libèrent .

Clément, P. (2006) Didactic transposition and the KVP Model: Conceptions as interactions between scientific knowledge, values and social practices. (pp.9-18). *Proceedings summer school ESERA*. IEC, Univ. Minho Portugal.

Clément, P. (2010) Conceptions, représentations sociales et modèle KVP. *Skholê*, 16, 55-70.

Clément, P. & Caravita S., (2011) Education pour le Développement Durable (EDD) et compétences des élèves dans l'enseignement secondaire – Rapport d'une étude commandée par l'UNESCO (88 pp.), http://hal.archives-ouvertes.fr/hal-01026073

Clément, P. & Caravita, S. (2012) Diversity of teachers' conceptions related to environment and human rights. A survey in 24 countries. In C.Bruguière, A.Tiberghien & P.Clément (eds), Ebook *Proceedings of the ESERA 2011 Conference*, *Science Learning and Citizenship, Part 8* (p. 42-48). http://

hal.archives-ouvertes.fr/hal-01054209

Clément P., Hergueta S. & Buisseret P. (1999) Pas si bêtes ! 1000 cerveaux, 1000 mondes. À chacun son monde. *Biologie-Géologie, bulletin APBG (Association des Professeurs de Biologie – Géologie)*, 3, 527-546.

Clément, P., Laurent, C. & Samonek, E. (2011) Polish teachers' conceptions related to environment. *Annales Universitatis Paedagogicae Cracoviensis, Studia ad Didacticam Biologiae Pertinentia I*, 86, 104-115.

d'Eaubonne, F. (1974) *Feminism or death*. Paris: Pierre Horay.

Djakou, R. & Thanon, S. Y. (1996) *Géologie. Biologie*. Paris : Nathan (first ed. 1988, Paris : Bordas)

Djakou, R. & Thanon, S. Y. (1986) Écologie, Géologie. Afrique intertropicale. Paris : Bordas.

Dolédec, S. & Chessel, D. (1987) Rythmes saisonniers et composantes stationnelles en milieu aquatique I- Description d'un plan d'observations complet par projection de variables. *Acta Oecologica, Oecologia Generalis*, 8(3), 403-426.

FAO (Food and Agriculture Organization of the United Nations) (2015) *The State of Food Insecurity in the World 2015*. FAO Hunger Map. http://www.fao.org/hunger/en/

Forissier, T. (2003) Les valeurs implicites dans l'Education à l'Environnement. Analyse de la formation d'enseignants de SVT et des conceptions de futurs enseignants français, allemands et portugais. PhD thesis, University Lyon 1 (France).

Foucart, S. & Thiberge, C. (2016) L'évaluation de la toxicité des plantes transgéniques remise en cause. *Le Monde* 20-12-2016, p.6.

Garric, A. (2013) Monsanto recule sur les OGM en Europe. *Le Monde* 18.07.2013

Gérenton, O. (1993) *Biologie 6ème - Afrique et pays intertropicaux*. Paris: Hatier.

Gérenton, O. (1995) *Biologie 5ème - Du Sahel à la forêt. Paris:* Hatier.

Horel, S. (2016) OGM : la recherche entachée de conflits d'intérêts. *Le Monde* 20-12-2016, p.6.

Khalil I., Clément P. & Laurent C., (2007) - Anthropocentrées, écolocentrées ou sentimentocentrées : Les conceptions d'enseignants et futurs enseignants libanais sur la nature et l'environnement. *Feuilles Libanaises (Ligue des Professeurs de l'Université Libanaise)*, 29, 67-92.

Ki-Zerbo, J. (dir.) (1990) Éduquer ou périr, Paris, Unicef-Unesco.

Larrère, C. (1987) *Les philosophies de l'environnement*. Paris: PUF, Philosophies

Manly, B. F. J. (1997) *Randomization, Bootstrap and Monte Carlo Methods in Biology*. 2nd ed. London: Chapman & Hall.

Mazrui, A.A. & Wagaw, T. (1986) Vers une décolonisation de la modernité : éducation et conflits culturels en Afrique de l'Est, in *Le processus d'éducation et l'historiographie en Afrique, (Proceedings colloque Unesco Dakar, 1982)*, pp. 35-64, Paris: Unesco.

McEwen, B., Clément, P., Gericke, N., Nyberg, E., Hagman, M. & Landström, J. (2015) Female and male teachers' pro-environmental behaviour, conceptions and attitudes towards nature and the environment do not differ: Ecofeminism put to the test. *Asia-Pacific Forum on Science Learning and Teaching*, 16, (1), 1-25.

Milfont, T. L. & Duckitt, J. (2010) The environmental attitudes inventory: A valid and reliable measure to assess the structure of environmental attitudes. *Journal of Environmental Psychology*, 30(1), 80-94.

Munoz, F., Bogner F., Clément P. & Carvalho G .S. (2009) Teachers' conceptions of nature and environment in 16 countries. *Journal of Environmental Psychology*, 29: 407-413. http://hal.archives-ouvertes.fr/hal-01024981

Nchia, L, N., Clément, P., Nditafon G. N., Fonkeng, E. G. & Tamesse, J. L. (2016) Primary and Secondary School Teachers' Conception of Environmental Education in the English and French Speaking

Subsystems of Education in Cameroon. *XVII International Organisation for Science and Technology Education (IOSTE) Symposium*, Braga (Portugal), 11-16 July 2016. http://www.eventos.ciec-uminho.org/ioste2016/

Nditafon, G, N. (2006) Delivering Environmental Education: A guide for the formal Sector: A teacher's Guide. Living Earth Foundation Cameroon Programme. *Cameroon Environmental education Support Project phase II.*

Ndong, L. (2012) Malnutrition protéino-énergétique et enseignement de la nutrition et de l'alimentation au Gabon. In Regnier J.-C., Bailleul M., Gras, R. (Eds). *L'Analyse Statistique Implicative : de l'exploratoire au confirmatoire* (pp. 263-289). France : Universités de Caen et Lyon 2

Nyéréré, J. (1972) *Indépendance et éducation.* Yaoundé (Cameroun): Éditions CLE

Quinn, F., Castéra, J., & Clément, P. (2016) Teachers' conceptions of the environment: anthropocentrism, non-anthropocentrism, anthropomorphism and the place of nature. *Environmental Education Research*, 22(6), 893-917.

Romesburg, H.C. (1985) Exploring, Confirming, and Randomization Tests, *Computers & Geosciences,* 11(1), 19-37.

Sanogo, Z. Y. & Coulibaly, N. B. (2003) Croyances animistes et développement en Afrique subsaharienne. *Horizons philosophiques*, 132, 139-152.

UNESCO (2009a) *Contextes et structures de l'Education pour le développement durable*, DEDD 2005-2014, page 9.

UNESCO (2009b) *The ESD Lens. Results of a review process by seven Southern Africa countries.*

Van Eeckhout, L. (2015) Manifestation au Burkina Faso contre les OGM de Monsanto. *Le Monde*.fr, 23.05.2015: http://www.lemonde.fr/afrique/article/2015/05/23/manifestation-au-burkina-faso-contre-les-ogm-de-monsanto_4639357_3212.html#rWbhpvbRVHtbehi4.99

WHO/UNICEF (2015) *Water sanitation hygiene. Key facts from Joint Monitoring Programme 2015 report.* http://www.who.int/water_sanitation_health/monitoring/jmp-2015-key-facts/en/

Wiseman, M. & Bogner, F. X. (2003) A higher-order model of ecological values and its relationship to personality. *Personality and Individual Differences*, 34, 783–794.

Zell, S. K. (1998) Eco-feminism and the science classroom: A practical approach. *Science & Education*, 7(2), 143-158.

A method to reveal fine-grained and diverse conceptual progressions during learning

François Lombard[1,2], Marie Merminod[1], Vincent Widmer[2] and Daniel K. Schneider[2]

francois.lombard@unige.ch

Abstract

Empirical data on learners' conceptual progression is required to design curricula and guide students. In this paper, we present the Reference Map Change Coding (RMCC) method for revealing students' progression at a fine-grained level. The method has been developed and tested through the analysis of successive versions of the productions of eight cohorts (N=100 total) of high school biology students groups involved in a year-long, inquiry-based learning design. Concepts and causal links expressed in students' gradually refined explanations of biological phenomena are charted onto reference model maps. Trends within variability in all cohorts are measured by a consolidated Prevalence Index (cPI) counting the occurrence of each item across all versions of the students' explanations. Results of a case study presented reveal great variability in patchwork progressions. Learners' diverse and often surprising conceptual paths challenge the view of learning as a linear process. For example, some items consistently appear later, thereby offering empirical evidence of slow spots that require attention. We discuss possible causes, educational implications, and show that our method offers crucial insight into the process of learning as it happens. We finally argue that RMCC also could become a follow-up tool for interested teachers.

---

[1] IUFE - Institut universitaire de formation des enseignants, Université de Genève, Geneva, Switzerland

[2] TECFA - Technologies de formation et d'apprentissage, Faculté de psychologie et des sciences de l'éducation, Université de Genève, Geneva, Switzerland

Keywords: assessing students progression, conceptual change, conceptual obstacles, inquiry, immunology, models

Learning at natural history dioramas: a model for interpreting museum biological settings

Edward Mifsud[1] and Sue Dale Tunnicliffe[1]

edward.mifsud@um.edu.mt

Abstract

This research attempts to conceptualise natural history habitat dioramas as potential models for biological learning of local flora and fauna. A cohort of 9-year-old students from Malta (N=63) was asked to draw in their class a place with local animals and plants, they then made another drawing at the Natural History Museum of Malta, before observing the dioramas, and then making a third drawing of their favourite habitat diorama. One-to-one interviews were conducted, during which the children were asked to elaborate on their drawings and comment on their choices. Drawings were analysed semi-quantitatively and qualitatively using the data analysis package ATLAS.ti. The inclusion of particular types of fauna (especially birds) in drawings was influenced by prior knowledge and culture. A progression from drawing from imagination in class and at the museum before viewing the diorama, to increasingly drawing from observation was noted. The educational potential and role in biological learning of dioramas has previously been reported in literature, but our results indicate the sound potential of natural history habitat dioramas to act as scientific models for biological learning. A major contribution of this research was the creation of a novel interpretation model for museum settings.

Keywords: dioramas, informal learning, interpretation model, museum learning, primary science

---

[1] University College London, Institute of Education, London, UK

1. Introduction

Habitat dioramas are three-dimensional museum displays presenting imitations of biological landscapes. They are often found in in natural history museums (NHM), and these displays typically show preserved animals in their natural foreground with freeze-dried or modelled flora of some form set against a painted background. The diorama's integrated montage of animals with their surroundings is a means of bringing natural history to 'life', but at the same time exposing human attitudes towards nature and so also performing a function in the cultural construction of our world (Wonders, 2003).

Various researchers have recently documented the educational potential and role of dioramas in biological learning (Ash, 2004; Insley 2007, 2008; Peart & Kool, 1988; Piqueras et al., 2008; Reiss & Tunnicliffe, 2007; Scheersoi, 2009; Tunnicliffe, 2002, 2005, 2007). Skilfully constructed natural history dioramas can still provide a significant opportunity for fundamental acquisition of science knowledge (Stern, 2009).

While limited research has dealt with the educational value and role of habitat dioramas, it has not considered how dioramas can enable the visualization of animals and plants. This study is an attempt to conceptualise habitat dioramas as a potential model for biological learning of local flora and fauna (Gilbert, 2005, 2008).

The main research questions are:

- Do children's drawings about animals and plants in local habitats change after their interaction with NHM dioramas? If so, in what way do drawings change?
- Which aspects of their experiences do visitors rely on to interpret a museum setting?

2. Theoretical framework

2.1 Dioramas and mental models

Natural history dioramas are still an underutilised educational resource and have been dismissed as old fashioned and irrelevant by non-educator management officials enticed by effective technological innovations. However, other authors (e.g. Tunnicliffe, 2009) have shown that dioramas can be a powerful

potential tool in science education with great potential for learning in biology, particularly in aspects of biodiversity, ecological relationships and ecosystem ecology.

When properly designed, dioramas allow lone visitors and small groups to carry their own interests to the exhibit and to connect with them in a way that provides a measure of control (Tunnicliffe & Reiss, 2006). Through their 'stillness', dioramas offer opportunities to 'stand and stare' and serve as a focus for biological understanding in an out-of-school environment. Dioramas potentially motivate visitors to stay longer at an exhibit and to facilitate their understanding of the object's functions, meanings or associations. Visitors may also relate their previous experiences to the scenes and artefacts presented in the diorama, which thus become "appealing, invite exploration and therefore facilitate learning" (Scheersoi & Tunnicliffe, 2009).

Sociocultural and constructivist frameworks inform this research. Learning in a constructivist manner may be understood as the building and refining of mental models, while also acknowledging the importance of social interaction in developing these models. The link between constructivism and sociocultural theory is interesting due to its potential to explain children's development of knowledge in terms of its individual and social construction, under the influence of social and cultural practices (Jaworski, 1996). Perhaps Falk & Dierking (2000) offer the most explicit combination of the two theoretical perspectives in their Contextual Model of Learning (CML), which states that learning is personally, socially, and physically situated. Students and teachers need to understand how science and science education are always a part of larger communities and their cultures, including the sense in which they take sides in social and cultural conflicts that extend far beyond the classroom (Lemke, 2001; O'Loughlin, 1992).

When children are allowed to interact with the dioramas, they stop to observe, notice the different forms of animals and plants, the anatomical features of each organism and possible relationships between animals and plants, or animals and animals. The child forms his or her concept of animals and plants in general, and more specifically a concept of the particular organisms featured in the exhibit. The child's personal knowledge of a phenomenon or main features of an object are held in the mental model formed. When asked, the child can produce a representation from the mental model (Cox, 1992; Reiss & Tunnicliffe, 1999).

During learning we habitually use external representations surrounding us to construct internal representations in our minds. We use external representations, such as habitat dioramas, to build an internal representation or mental model. There is no direct evidence of the existence of mental representations, other than when these are reproduced in some external form (Rapp & Kurby, 2008). Often we are called to convert our mental representations into external presentations during communication, for example when writing a scientific paper or composing an email (Gilbert, 2008). Representations are very frequently used in science education and we can make sense of these when we are able to visualise what the representation is meant to show. Any person studying science needs to be able to visualise a phenomenon through a representation such as a model. The key aspect of 'metavisualisation' is the ability to make meaning of (visualise) a representation in the different special dimensions it may occur (Gilbert, 2008). Observation, visualisation and learning are closely linked.

2.2 Interpretation models

Some major and well accepted interpretative models commonly encountered in science education and informal learning literature are: a) Contextual Learning Model (Falk & Dierking, 2000); b) Acuity Model (Patrick, 2006); c) Model Based Learning (Buckley & Boulter, 2000) and d) Activity Theory (Leont'ev, 1974; Engeström et al., 1999). We consider the Activity System that originates from Activity Theory (Leont'ev, 1974; Engeström et al., 1999) as the most appropriate and adaptable in the case of interpreting a museum exhibit such as a habitat diorama. Based on this, we propose our own model, which we present in the Results section. Figure 1 is an illustration of the components of the system and the relationship between them. Here are the components described:

- 'Object' is the objective of the activity system. Object refers to the objectiveness of the reality; items are considered objective according to natural sciences, but also have social and cultural properties. In interpreting a museum exhibit, this becomes the *focus* or the main theme of the exhibit.
- 'Subject' is the actor or actors engaged in the activities. Likewise, the *subject* is the person interacting with the exhibit.
- 'Community' is the social context; all actors involved in the activity system. For a museum exhibit this would be the *group* with whom the visitor

is viewing the exhibit such as a family or a class.

- 'Mediating artefacts' (or concepts) used by actors in the system. These are tools influencing actor-structure interactions and they change with accumulating experience. Tools are influenced by culture, and their use is a way for the accumulation and transmission of social knowledge. The museum exhibit or diorama is the *artefact* that is conveying the message or the theme.
- 'Division of labour' refers to social strata, hierarchical structure of activity, or the division of activities among actors in the system. This is not applicable to museum exhibit interpretation.
- 'Rules' are the conventions, guidelines and rules regulating activities in the system. There are social norms and practices that apply to museums.

The Diorama Interpretation Model is derived from the Activity System, which has been adapted to become applicable to museum exhibits with the inclusion of other features that are presented and explained later in this paper.
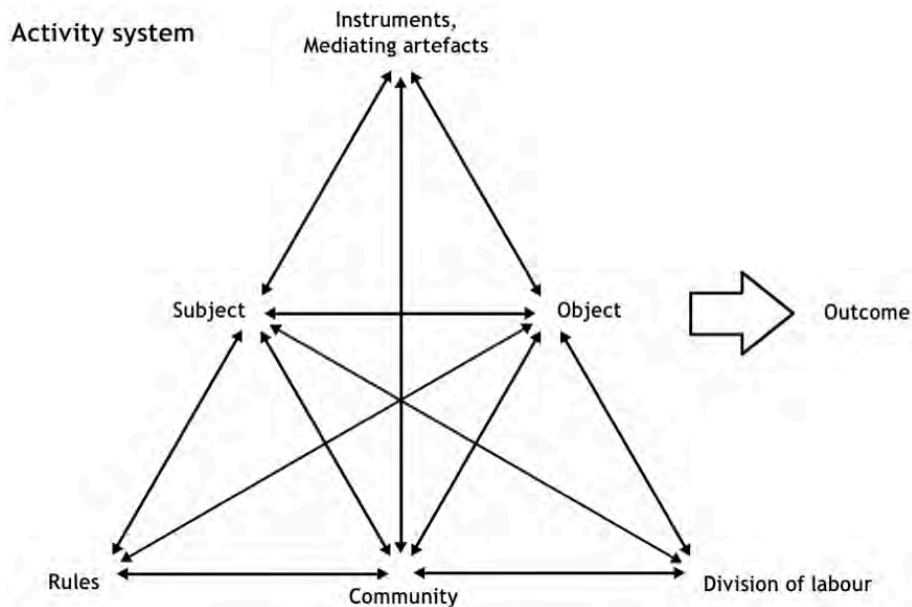
Figure 1 Activity System.

3. Research design and methodology

This research is a mixed method quantitative/qualitative case study. The participants were mixed ability 9-year-old students (N=63) from a state co-educational primary school in central Malta. The pupils came from the middle to working class social strata. These pupils did very little biological science in school and their performance was not determined in a formal way. A few had relatives who owned a piece of land and so they had experience of the countryside, but most pupils had little direct contact with nature. None of the pupils had ever been to the NHM in Malta before this study. Malta was the context of the study since one of the authors is Maltese and no such research had even been carried out in Malta. The school was selected after the authorities granted formal permission, and the pupils selected formed the entire fifth grade cohort.

Children were first asked to draw 'a place with local animals and plants' in class. A fortnight after they visited the NHM of Malta, where they first produced another drawing before observing the five dioramas in groups of four pupils selected at random. Right after leaving the diorama area, the pupils were asked to draw their favourite diorama. A one-to-one interview was conducted a week after the visit, with each interview lasting 8-10 minutes. Each student was asked to explain the three drawings produced and comment on the choice of features and other points of significance to the student. Interviews were transcribed and content analysis conducted to elicit common categories that inform the interpretation model.

Drawings were analysed semi-quantitatively and qualitatively using the data analysis package (CAQDAS) ATLAS.ti (Friese, 2012). A coding method was developed for analysing the drawing, in principle similar to emergent analytic coding developed by Haney et al. (2004). A list of features that the drawings contain was made, and each feature was assigned a specific code as shown in Figure 2. Main code categories, such as 'animal' and 'plant', were assigned. In the case of animals, taxonomic sub-categories were added to better classify the organisms included. Each animal included in the drawing was linked to the appropriate taxonomic sub-category for example 'mammal', 'insect' or 'bird'.
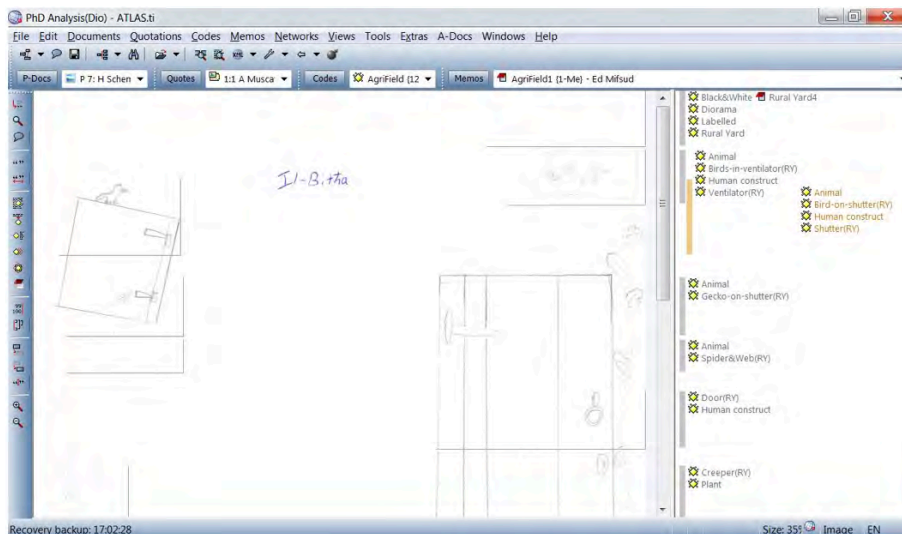
Figure 2 Coding using ATLAS.ti.

We analysed the three drawing set produced by each pupil to determine the sort of developmental progression (if any) from Class, through Pre-diorama to Diorama. We based our analysis on a similar one performed for a large-scale study at the London Zoo aimed to assess the formal learning programme offered by the London Zoo (Jensen, 2011). We then studied the drawings for changes in quantity and categories of biota presented as well as abiotic features and diorama details. Noticeable development was found in the types of fauna and flora pupils drew, how they elaborated the biota and whether or not these were presented in a habitat.

4. Results

In the class and pre-diorama tasks, children were asked to draw 'a place with local animals and plants' and they drew more animals than plants (mainly birds, mammals and arthropods), but also wrote down far more animals than plants. The ability to recall animals in preference to plants could be due to various factors such as a greater knowledge about animals, and a general disregard and lack of sensitivity for plants. However, Maltese children mentioned more animals than they actually included in their drawings, with much greater variety likewise dominated by birds and mammals. This emulates the famous

Ausubelian maxim (Bell 1993; Freeman, 1997) that "the child knows more than he draws". The proportions of animals, plants, and the other objects written in the webs (81%, 13% and <1% respectively) were almost identical to those given by Yorek et al. (2009).

The animals were generally represented in a similar form to that reported by Golomb (2004). Trees were seen in a 'lollypop' shape, with disproportionate trunks, and flowers as 'sunflowers' with a long stalk and prominent petals. Animals were almost only drawn in their standard sideways orientation, highlighting the distinctive features of the subject. Four-legged animal drawings showed some degree of figural differentiation and display the right-angular directions seen in humans. For example, mammals were drawn with their bodies displayed horizontally in side view, head in frontal view, four straight legs and an occasional tail. Fish were drawn as an oval with the usual sideways fishmouth, one or two eyes and a tail. Birds were typically shown in aerial view with head, body and tail aligned horizontally, with wings extending vertically.

The results from this research point towards a more comprehensive idea of what qualifies as an 'animal' for Maltese 9 year olds than reported in literature (Bell, 1981). Compared to their foreign peers, Maltese students refer to a greater variety of species that they consider as an animal. The general archetypal animal of most pupils is the large terrestrial, four-legged vertebrate, mostly mammal species (e.g. cow, cat, lion, elephant), and animals found at home as pets, on a farm or in the jungle. The class drawings presented the subordinate groups (taxon) *amphibian, bird, cnidarian, crustacean, fish, insect, mammal, mollusc* and *reptile*, while the museum drawings presented the subordinate groups *bird, echinoderm, fish, insect, mammal, mollusc* and *reptile*. Birds (37%), mammals (24%), arthropods (16%) and fish (13%) were the animals mainly drawn in class, while birds (60%), arthropods (18%) and mammals (13%) were mainly drawn at the museum. This research also shows that arthropods such as ant, bee, butterfly, ladybird and spider ranked more highly in frequency among Maltese children compared to foreign peers.

In this study, differences in number of species was not so striking. Class drawings yielded 15 different species of mammals, 11 species of bird, 5 species of arthropods, while museum drawings yielded 16 different species of bird, 10 species of mammals, and 4 species of arthropods. The mammals drawn were mostly endemic or domesticated species (cat, cow, dog, donkey, hamster,

horse, bat, rabbit, rat), with far fewer exotic species (elephant, leopard, lion, tiger, monkey, kangaroo, squirrel). In accordance with another study (Tunnicliffe et al., 2008), Maltese children very rarely drew aquatic/semi-aquatic mammals such as dolphins, whales and seals.

In both class and museum, children drew far fewer plants than animals with far less variety and mainly seeded types. In class drawings, the following plants were noted: moss, palm, tulip, reed, sunflower, daffodil, apple, olive, orange, peach and pine tree, compared with apple, cherry, pine, rose and sunflower in museum drawings. When they could not give a particular exemplar, children referred to the vegetative specimen as simply 'plant' (Bell, 1981). The findings seem to strengthen the view that plants are of no immediate importance to children (Bowker, 2007; Johnson, 2004), and they seem to have what has been referred to as 'plant blindness' (Wandersee & Schussler, 2001). The general trend showing that animals are noticed much more than plants was observed in each diorama drawing. A good number of drawings (75%) featured a least one plant, but the total number of plants (14%) was half that of animals (32%). Human artefacts (human-made structures) seem to be more important to children than plants. Few drawings showed anthropomorphic features. In class, drawings of human features were present in 63% of drawings (33% in museum), most of which were human-made objects such as a boat, rubble wall, house, glasshouse, road, aquarium, airplane, barn and tools. Looking at each child's set of three drawings we could notice appreciable changes that occurred throughout the tasks set. Almost half the pupils (47%) showed a development towards greater variety and increased habitat representation, while almost a third (28%) showed greater detail in organism representation without enhanced habitat representation. The rest of the pupils' drawings (25%) showed an opposite effect, that is, loss of habitat, reduced variety or more basic organism representation.

We consider the *Activity System* to be most appropriate and so have based our model on it, with the inclusion of additional features as suggested by empirical evidence from our research. Figure 3 and Table 1 below shows the Interpretation Model as we propose it, which presents six interrelated factors. 'Focus', 'Artefact', 'Group' and 'Subject' emerge from Activity Theory, while 'Culture', 'Previous Knowledge', 'Mental Model' and 'Expressed Model' emerge from data. This model may be used to interpret museum objects or artefacts, and particularly in this case the habitat dioramas.

Figure 3 Interpretation Model for cultural tools.
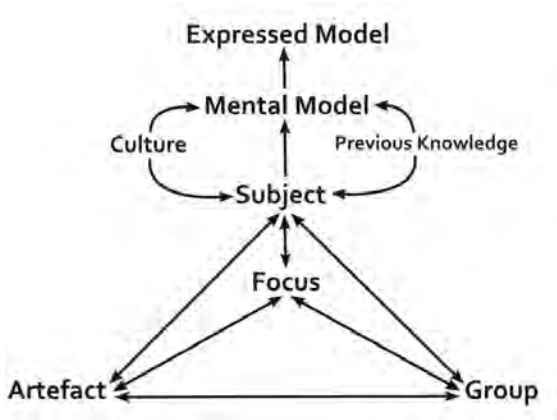
Table 1 Interpretation Model terms defined.

Subject	The person observing the diorama, i.e. student or visitor.
Artefact	The mediating tool; a diorama, picture, 3D model or other media forms.
Focus	The idea, topic or location represented by the artefact and of interest to the subject and/or group, such as 'habitat'.
Group	The group of people i.e. friend or family, with whom the subject experiences the artefact.
Mental model	The personal representation of the artefact or focus held in the subject's mind.
Expressed model	The external representation of the mental model.
Culture	The sociocultural imprint of the family, country and society.
Previous knowledge	What the subject already knows about the focus.

4.1 The interpretive model applied

Figure 4 shows the Sand Dune Diorama and the representation of it in a drawing. Figure 5 shows how the interpretation model can be used to show how the pupil interpreted this particular setting.

Figure 4 Sand Dune diorama and drawing.
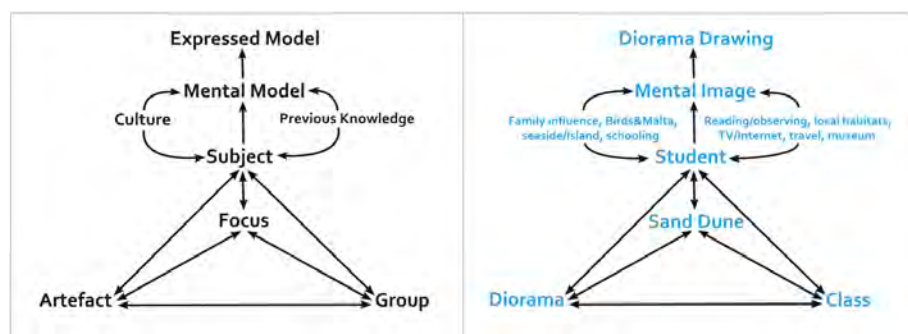
Figure 5 Applying the Interpretation Model.

In this case the model was applied (right hand side of Figure 5) to the drawing produced by a boy who chose to represent the sand dune habitat diorama as show in Figure 4. Here the model was employed on this diorama as the museum artefact, with the biological focus of a sand dune that was viewed and studied as a group of students (the class). The subject (person) viewing it is the pupil whose mental image is influenced by the cultural elements of family, school, birds in Malta and sea around the island. It is also influenced by knowledge from reading, observation, local habitats, TV and internet, travel to foreign countries and museums.

5. Discussion

This research investigates the influence of natural history museum habitat dioramas on knowledge of biota and how this is expressed in drawings. Did the viewing of the dioramas in any way affect the way children drew? Drawing

is certainly a valid representational and research tool, but it may be limited in showing a child's comprehensive knowledge. We should not assume that children's drawings are print-outs of mental images (Jolley, 2010) and that children never just copy (Kress, 1997). Although we cannot exclude the idea that some children do tend to copy from their peers' work, each drawing expresses a unique context for the visual forms and structures that are copied (Hopperstad, 2010; Kress, 1997). There is little evidence in our results that children actually copied and very few drawings by different pupils show clear similarities.

Some important conclusions that emerge from this research are that Maltese children have a better understanding of what an animal is by mentioning a wider range of species of organisms they think are animals. They also differ from their foreign peers in that they included more birds than mammals, and also arthropods feature quite well in drawings. Malta has its own different culture, with a long and deep-rooted tradition of bird trapping and hunting. The importance of birds is quite clearly evidenced in the data; birds are consistently the most frequently drawn animals in the three drawing tasks, which vary from what was reported in other countries, in that mammals are the preferred class of animals. The apparent disregard of children for plants has been previously reported in literature and was confirmed in this research. Wandersee & Schussler (2001) coined the term *plant blindness* and argued that two possible indications of this might be: a) the idea of plants as just the backdrop for animals and b) failing to notice plants in the environment. Plants in the local habitat dioramas at the NHM in Malta are located in prominent positions and not just serving as a background for animals.

Studying carefully the drawings produced in class and the museum, we noted important changes. There is a progression from drawing from imagination in class and at the museum before viewing the diorama, to increasingly drawing from observation, but still showing signs of imagination in the diorama drawings. There is a greater sense of intellectual realism rather than visual realism as evidenced by the 9 year old participants in this research. For almost half the pupils, the viewing of the dioramas had a positive effect on the drawings produced. The class drawings are an expression of their present mental model, of the flora and fauna of Malta. When observing and interpreting the dioramas, or any other museum exhibit, the visitor draws on his or

her existent mental model. In other words, the visitor observes and interprets the dioramas through this conceptual lens. The novel museum environment was expected to have an effect on the mental model. Drawing at the museum confirmed this and for most children the mental model expressed did to a certain extent change, even if not for all in the same way. The mental model expressed in the drawings seems to be influenced by the place where children settle to create their drawings.

A development noted was the change from iconic to more realistically represented animals and in some cases also plants. This shows a desire to capture the object and represent it was evident in those pupils that drew one or two animals in greater detail (Golomb, 2004). Different children perceive and represent natural objects uniquely. Their attention to detail varies, with some children taking a more generalized view of the 'scene' depicting broad shapes and borders with little detail of plants and animals, but still using their own schematic graphics to represent them in some way. Others, albeit fewer, show greater detail as their attention is captured by the features of the organisms they observe. In this way, children's drawings became unique and personal, making generalization difficult while analysing the drawings.

From the example provided, we show which elements seem to be invoked while interpreting a museum artefact such as a habitat diorama. These are presented in the Interpretation Model developed from the Activity System, which is not directly applicable in our case but in a modified form it may be applied to interpret museum settings. The strengths of the model lie in the manner it links together the elements involved in the interpretation of an artefact (mediating tool) to understand the message it conveys, for example Natural History Dioramas present flora and fauna in their habitat showing possible ecological relationships. It elucidates how a learner may understand a topic as mediated by an artefact, to construct an intangible mental model to create a tangible expressed model (a drawing). The interaction with peers, the cultural baggage possessed and knowledge held, may influence the mental model constructed. Potentially, this may apply to various topics as presented or modelled by 2D, 3D or virtual mediating tools. This may be done in different learning situations in science and other areas in formal, non-formal and informal settings.

References

Ash, D. (2004) How families use questions at dioramas: ideas for exhibit design. *Curator*, 47(1), 84-100.

Bell, B.F. (1981) When is an animal, not an animal? *Journal of Biological Education*, 15(3), 214-218.

Bell, B.F. (1993) *A constructivist view of learning in Children's Science, Constructivism and Learning in Science*. Waikato: Deakin University.

Bowker, R. (2007) Children's perceptions and learning about tropical rainforests: an analysis of their drawings. *Environmental Education Research*, 13(1), 75-96.

Buckley, B. C. & Boulter, C. J. (2000) Investigating the Role of Representations and Expressed Models in Building Mental Models. In J. K. Gilbert & C. J. Boulter (Eds.), *Developing Models in Science Education*. London: Kluwer Academic Publishers.

Cox, M.V. (1992) *Children's Drawings*. London: Penguin Books.

Engeström, Y., Miettinen, R. & Punamäki, R.L. (1999) *Perspectives of Activity Theory*, Cambridge University Press.

Falk, J.H. & Dierking, L.D. (2000) *Learning from Museums: Visitor Experiences and the Making of Meaning*. New York: AltaMira Press.

Freeman, N.H. (1997) Identifying Resources from which Children Advance into Pictorial Innovation. *Journal of Aesthetic Education*, 31(4), 23-33.

Friese, S. (2012) *Qualitative Data Analysis with Atlas.ti*. London: Sage.

Gilbert, J.K. (2005) Visualization: A metacognitive skill in science and science education. In J. K. Gilbert (Ed.), *Visualization in Science Education (Vol. 1)*. Dordrecht: Springer.

Gilbert, J.K. (2008) Visualization: An Emergent Field of Practice and Enquiry in Science Education. In J. K. Gilbert, M. Reiner, & M. Nakhleh (Eds.), *Visualization: Theory and Practice in Science Education (Vol. 3)*. Dordrecht: Springer.

Golomb, C. (2004) *The Child's Creation of a Pictorial World (2nd ed.)*. New Jersey: Lawrence Erlbaum Associates, LEA.

Haney, W., Russel, M. & Bebell, D. (2004) Drawing on Education: Using Drawings to Document Schooling and Support Change. *Harvard Educational Review*, 74(3), 241-271.

Hopperstad, M.H. (2010) Studying meaning in children's drawings. *Journal of Early Childhood Literacy*, 10(4), 430-452.

Insley, J. (2007) Setting the Scene. *Museums Journal*, 2(107), 33-35.

Insley, J. (2008) Little Landscapes: dioramas in museum displays. *Endeavour*, 32(1), 27-31.

Jaworski, B. (1996) *Constructivism and Teaching: The socio-cultural context*. www.grout.demon.co.uk

Jensen, E. (2011) *Learning about Animals, Science and Conservation at the Zoo: Large-scale survey-based evaluation of the educational impact of the ZSL London Zoo Formal Learning programme*. London: University of Warwick and Zoological Society of London.

Johnson, S. (2004) Learning science in a botanic garden. In M. Braund & M. Reiss (Eds.), *Learning Science Outside the Classroom*. London: Routledge Falmer.

Jolley, R.P. (2010) *Children and Pictures: drawing and understanding*. Oxford: Wiley-Blackwell.

Krampen, M. (1991) *Children's Drawings: Iconic Coding of the Environment*. London: Plenum.

Kress, G. (1997) *Before Writing. Rethinking the Paths to Literacy*. London: Routledge.

Lemke, J.L. (2001) Articulating Communities: Sociocultural Perspectives on Science Education. *Journal of Research in Science Teaching*, 38(3), 296-316.

Leont'ev, A.M. (1974) *General Notion of Activity*. Moscow: Nauka.

O'Loughlin, M. (1992) Rethinking Science Education: Beyond Piagetian Constructivism, Toward a

Sociocultural Model of Teaching and Learning. *Journal of Research in Science Teaching*, 29(8), 791-820.

Patrick, P.G. (2006) *Mental Models Students Hold of Zoos*. (PhD), Unpublished PhD Thesis. University of North Carolina, Greensboro.

Peart, B. & Kool, R. (1988) Analysis of a natural history exhibit: are dioramas the answer? *The International Journal of Museum Management and Curatorship*, 7, 117-128.

Piqueras, J., Hamsa, M.K. & Edvall, S. (2008) The practical epistemologies in the museum: A study of students' learning in encounters with dioramas. *Journal of Museum Education*, 33(2), 153-164.

Rapp, D.N. & Kurby, C.A. (2008) The 'Ins' and 'Outs' of Learning: Internal Representations and External Visualizations. In J. K. Gilbert (Ed.), *Visualization: Theory and Practice in Science Education (Vol. 3)*. London: Springer.

Reiss, M.J. & Tunnicliffe, S.D. (1999) Conceptual Development. *Journal of Biological Education*, 34(1).

Reiss, M.J. & Tunnicliffe, S.D. (2007) *Opportunities for learning in biology*. Paper presented at the NARST: April 15-17, New Orleans.

Scheersoi, A. (2009) The important role of Natural History dioramas in biological learning: Biological interest development at Natural History dioramas. In S. D. Tunnicliffe & A. Scheersoi (Eds.), *The important role of Natural History dioramas in biological learning (Vol. 29*, pp. 1-40): ICOM Natural History Committee.

Scheersoi, A. & Tunnicliffe, S.D. (2009) Editorial. In S. D. Tunnicliffe & A. Scheersoi (Eds.), *The important role of Natural History dioramas in biological learning* (Vol. 29, pp. 2-4): ICOM Natural History Committee.

Stern, T. (2009) An afternoon among dioramas at Yale Peabody Museum. In S. D. Tunnicliffe & A. Scheersoi (Eds.), *The important role of Natural History dioramas in biological learning* (Vol. 29, pp. 14-15): ICOM Natural History Committee Newsletter.

Tunnicliffe, S.D. (2002) The educational value of natural history collections in learning about biodiversity. *The Biology Curator* (22), 27-40.

Tunnicliffe, S.D. (2005) What do Dioramas Tell Visitors? A Study of the history of Wildlife Diorama at the Museum Of Scotland. *Current Trends in Audience Research and Evaluation*, 18, 23-31.

Tunnicliffe, S.D. (2007) The potential of natural history dioramas in developing science inquiry. *Gemnews,* (Winter), 7.

Tunnicliffe, S.D. (2009) Inquiry at Natural History Dioramas: useful resources in science education. In S. D. Tunnicliffe & A. Scheersoi (Eds.), *The important role of Natural History dioramas in biological learning* (Vol. 29, pp. 16-20): ICOM Natural History Committee Newsletter.

Tunnicliffe, S.D., Gatt, S., Agius, C. & Pizzuto, S.A. (2008) Animals in the lives of young Maltese Children. *Eurasia Journal of Mathematics, Science & Technology Education*, 4(3), 215-221.

Tunnicliffe, S.D. & Reiss, M. (2006) *Drawing breath: the use of drawings and interviews in a six-year longitudinal study of 5-11 year-olds' understandings of what is inside themselves*. Paper presented at the ERIDOB: London.

Wandersee, J.H. & Schussler, E.E. (2001) Toward a theory of plant blindness. *Plant Science Bulletin*, 17(1), 2-9.

Wonders, K. (2003) Habitat Dioramas and the Issue of Nativeness. *Landscape Research*, 28(1), 89-100.

Yorek, N., Sahin, M. & Aydin, H. (2009) Are animals 'more alive' than plants? Animistic-anthropocentric construction of life concept. *Eurasia Journal of Mathematics, Science and Technology Education*, 5 (4), 369-378.

Conceptions in the flesh: the educational moral metaphors system, using the example of meat consumption

Nadine Tramowsky[1] and Jorge Groß[1]

nadine.tramowsky@uni-bamberg.de

Abstract

The aim of this research is to analyse moral arguments of students in respect to meat consumption and livestock farming. In order to encourage discussion and reflection processes about morality and socio-scientific issues in biology classes, we explored the structure of moral conceptions based on a theory of mind. It is already claimed that human thinking is based on experience, and moral conceptions are also embodied and structured by metaphors (Lakoff & Johnson, 1999). Based on this theory, this paper will develop and give a perception of the Educational Moral Metaphors System as an instrument to identify and analyse moral conceptions. Upon this theory, we have developed and evaluated an evidence-based instrument to explore moral conceptions. The findings result from 5 teaching experiments with 15 students (10-16 years old). Therefore, in this article specific students' conceptions and metaphors are examined on the basis of moral arguments and judgements by means of Qualitative Content Analyses. The results illustrate the Educational Moral Metaphors System as an analytical tool and how different moral arguments and students' conceptions can be, as well as which recurring metaphorical structures can be found. In this paper we will illustrate that moral thinking is expressed in language by metaphors.

Keywords: cognitive metaphor theory, livestock farming, meat consumption, moral arguments, students' conceptions

---

[1] Department of Science Education, University of Bamberg, Germany

1. Introduction

Research on moral conceptions in the area of biological and ecological ethics gained increasing importance with the implementation of the new educational standards, such as decision-making competence in Germany (KMK, 2004). Therefore, education is more than just knowledge. Decision-making competence in biology classes has already been reviewed in depth, for example about abstract topics such as genomics or abortion and in connection with moral thoughts (e.g. Hößle, 2003; Boerwinkel & Waarlo, 2010). On the one hand, socio-scientific issues (hereinafter referred to as SSIs) are more authentic problems and provide suitable topics for argumentation (Zeidler & Sadler, 2008). On the other hand, morality is linked to SSIs, which can be an instrument for contributing to decision-making (Sadler & Zeidler, 2003). For example, people argue about SSIs such as organic food on the strength of their personal experiences and make decisions based on their values (Rundgren & Rundgren, 2010). These researchers developed the SEE-SEP model as an analytical framework for SSIs, which combines different areas (ethics and moral, environment, science, policy, sociology and culture, economy) with the aspects of personal experience, knowledge and values (Rundgren & Rundgren, 2010). In biology, school students are already confronted with SSIs. Here it is effective to focus on topics, which are already part of biology classes and to enrich these issues by allowing more room for moral thoughts (Boerwinkel & Waarlo, 2010).

During teaching such topics, teachers could support moral arguments and reflection in order to promote decision-making competences. Since diet is a key issue in biology class, and more than just a question of nutrition, it is a suitable object for analysing moral conceptions. In industrialised countries, for many people eating habits are part of lifestyle and well-being. The patterns of consumption also influence the environment, health, science, economy, politics and culture – and are linked to our morality. Therefore, diet is a personal, but not a private matter. The field of science education in respect to meat consumption and livestock farming is linked to SSIs (e.g. Morin et. al, 2014; Brocos & Jiménez-Aleixandre, 2016). But in total, few specific researches in this field exist. So far, animal ethics issues play a subordinate role in biology classes and teacher training courses (Binngießer, 2013). In order to understand moral conceptions, a theory of mind, such as the Cognitive Metaphor Theory (Lakoff & Johnson, 1999) is necessary. Based on this theory, our research

mainly focuses on developing (Substudy 1) and evaluating an instrument to identify and analyse moral conceptions about livestock farming and meat consumption under consideration of values, experience and knowledge (Substudy 2). In Substudy 1, we developed an instrument for analyzing students' moral conceptions by adapting Lakoff & Johnson´s (1999) 20-metaphor scheme to our own animal-related topic. In Substudy 2 we utilize this instrument in a new qualitative case study with 15 students.

2. Theoretical background and research questions

Morality includes individual or collective conceptions and beliefs, according to which people value actions as morally good or evil. But the connection between morality and good and evil is remarkable because moral principles can create both good and bad things. According to the trend towards moral scepticism, there is no valid morally good or bad and the answers to this question are controversial. Humans are able to make moral statements, act morally and express values. But what has qualified us to do this? The *rationalistic* approaches according to Kohlberg (1976) and Piaget (1997) assume that moral thinking develops progressively and irreversibly through rational thinking and empathy in stages. Unlike Piaget and Kohlberg, Haidt (2012) turns against the purely rationalistic traditional model of moral judgement. From his *intuitionist* perspective, moral judgements are the result of a quick and automatic assessment, an intuition. From a *cognitive-linguistic* perspective, morality includes abstract ideas and can therefore be understood in a metaphorical form. Based on this theory, the paper will show that moral conceptions are structured by metaphors (Lakoff & Johnson, 1999; Gropengießer, 2007).

In the sense of a moderate constructivism, students are understood as self-controlled, social and actively constructing subjects with prior experience (Phillips, 2000; Gerstenmaier & Mandl, 1995). Ready in 1999, the philosophers and linguists Lakoff and Johnson claimed that human thinking is based on experience and our neuronal network is structured metaphorically (e.g. Gallese & Lakoff, 2005). Such implicit metaphors are not flashy rhetorical tools such as the explicit metaphors. It is more a discreet structure of thinking. Humans can understand things such as up, down, behind and in front because they have made bodily or social experiences in the source domain. This perennial experience has become a kinaesthetic schema, which can be

transferred imaginatively by metaphors on an abstract target domain in order to understand it. The result that metaphors provide an approach to understand moral thoughts and influence our actions has already been published in various research papers (e.g. Meier et al., 2007; Zhong & Liljenquist, 2006). These conceptions are expressed by means of language, and affect the actions of people. According to Lakoff & Johnson (1999), the direct social environment – such as the family – is the place where children experience what is right, wrong, good or bad. However, experiences in families can vary: This depends on different family models with (1) rather incontestable persons of authority in hierarchical communication structures, moral role models, praise/punishment and obedience (strict parent family morality) or (2) more partnership-democratic and non-violent structures with responsibility, empathy and dialogue-based communication structures (nurturing parent family morality). Early childhood experiences create (to 1), conservative or (to 2) progressive value orientations. Moral conceptions are structured by many metaphors and they also result in different value orientations. Since experiences in the social environment can be both conservative and progressive, many people have both orientations.

In order to develop an Educational Moral Metaphors System (hereinafter referred to as EMMS; Figure 2), several of the more than over 20 moral metaphors described by Lakoff & Johnson (1999), were sorted into five superior moral metaphors (Tramowsky et al., 2016). The five moral metaphors were deductively developed by systematically categorizing the existing metaphors (moral categories) and their examination based on individual cases. For this purpose, we analysed the students' conceptions from our preliminary study (n=6) by means of the numerous existing basis metaphors (Table 3). Students use certain moral metaphors in certain target domains. In order to apply this to educational processes, matching metaphors were merged to produce superior moral metaphors. The five metaphors are formed on an empirical basis through the creation of categories. According to the Cognitive Metaphor Theory, morality grows (as with other conceptions) on the basis of physical and social experiences acquired in our environment. In order to identify moral metaphors, a theory-based frame was developed, which is called EMMS. Metaphors are summarized according to the EMMS in moral conceptions, which are characterized by specific structures. The *Ruler Metaphor* is based on expe-

riences with regimes and authorities: God is "naturally" more powerful than humans. Humans are more powerful than animals, adults are more powerful than children, and men are more powerful than women. This "natural" order can become a corresponding moral order (Lakoff & Johnson, 1999, pp.298-304). In order to the *Be Good Metaphor,* moral people are described as healthy and pure ("you have a *clean* vest") and immoral people as sick, infectious and unclean. Each object has a moral essence that establishes moral behaviour. The *Be Good Metaphor* is structured by the metaphor of moral essence, the metaphor of moral purity and the metaphor of moral health. Another form of moral thoughts is summarized as an *Accounting Metaphor*. Here, economic concepts are applied to moral problems. Moral actions can also be understood as an accounting process to balance moral accounts. It is moral to pay debts and immoral not to do so. Fairness is an imaginative process to balance moral accounts; fair treatment and distribution are moral actions, whereas unfair treatment and distribution are immoral actions (Lakoff & Johnson, 1999, pp.292-298). Here, the experience with justice is transferred to other moral areas. This is about the equitable distribution of objects or goods (e.g. smartphones, money, a pretty pebble found on the beach) or immaterial objects (e.g. career prospects, participation, responsibility, power). Moral in this case is the equitable distribution. But what is meant by simple distribution is a subject of dispute (ibid.). The *Freedom Metaphor* can structure decisions by establishing freedom as moral, and restrictions as immoral (ibid., p.304). The *Empathy Metaphor* forms a basis for moral empathy. The community is understood metaphorically as family, the moral agents as the nurturing parents, the humans needing help as children, moral actions as nurturing acts and the well-being of others as one's own well-being (ibid., pp.309-310).

On this basis, we have developed EMMS, and this paper will provide a perception of our findings according to this system. The research questions we chose were as follows:

- What moral conceptions can be found in students' arguments about meat consumption and what is their genesis?
- To what extent is the Educational Moral Metaphors System suitable to identify and structure moral conceptions about meat consumption?

3. Research design and methods

The Model of Educational Reconstruction (hereinafter referred to as MDR) has been developed as a theoretical framework for research and development in science education (Duit et al., 2012). Our qualitative case study is based on the MDR. The following research tasks have been processed:

- clarification of scientific content
- investigating students' perspectives to understand the process of comprehension of students' conceptions, and
- construction of learning environments in a process-oriented design framework in which findings of the first und the second task are included.

In the teaching experiment (according to Komorek & Duit, 2004), we questioned 15 students (10-16 years old, 8 females and 7 males) of the 5th, 9th, and 10th grade (Secondary School, 3 vegetarians and 12 non-vegetarians) by using a standardized guideline throughout the interview (60 to 90 minutes long). The participation was voluntary. By consultation with a biological teacher, we went into a German Secondary School for volunteer participants. The teacher, students and parents were informed about the aim of the study (improvement of biology education with regard to ethical issues), data recording, storage and protection. We paid attention to a balanced gender distribution. Our focus was on students at the beginning of the Secondary School (5th grade) and at the end of the Secondary School (9th/10th grade). Students at the end of secondary school can already make self-determined decisions on meat consumption. These more independent statements can be compared with the statements of the students at the beginning of Secondary School. The topic of the study was unknown for the students. All personalised data were made anonymous.

The teaching experiment consisted of two parts:

- Firstly, the students were asked in individual interviews or group interviews specific questions about different types of livestock farming and meat consumption (investigation phase). Also the students were asked about different effects of meat consumption on the ecological, economic and social environment. On the other hand, we asked about personal experience,

values, and attitudes as well as ratings and decisions (for example, "How do you rate that?"). The students of the 9th and 10th grade were questioned in individual interviews about their own family models.

- Secondly, the students could make reconstructed interventions in the group interviews (conciliation phase). For instance, to make moral conceptions about the human-animal relationship (Ruler Metaphor) discussible, the students were given six pictures. Depicted were a father, a mother, a pig, a farmer, teens, a child and a card with the task: Who decides over whom? Arrange the pictures! With the aid of these interventions, the students could reflect and enlarge on their own perspectives. Table 1 contains more information about the profile of the participants and the teaching experiments.

Table 1 The profile of the participants and the teaching experiments.

Participant	Teaching Experiment 1			Teaching Experiment 2			Teaching Experiment 3			Teaching Experiment 4			Teaching Experiment 5		
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15
Name (anonymised)	Zara	Marie	Greta	Jakob	Malte	Peter	Luisa	Helena	Nora	Eva	Paul	Milan	Noah	Mira	Maximilian
Age	11	11	11	11	11	11	14	14	15	15	14	15	15	15	16
Gender	Female	Female	Female	Male	Male	Male	Female	Female	Female	Female	Male	Male	Male	Female	Male
Vegetarian	No	No	No	No	No	Yes	Yes	No	Yes	No	No	No	No	No	No
Grade	5th			5th			9th			9th			9th/10th		
Special subject	No			No			Economy			Economy			Science		
Investigation phases															
Individual interview (30 mins.)	No			No			Yes			Yes			Yes		
Group interview (60 to 90 mins.)	Yes			Yes			Yes			Yes			Yes		
Drawings	Yes			Yes			No			No			Yes		
Conciliation phases															
Ruler Metaphor	Yes			Yes			No			Yes			No		
Accounting Metaphor	Yes			Yes			No			No			No		
Freedom Metaphor	No			No			No			No			Yes		
Empathy Metaphor	Yes			Yes			Yes			No			No		
Consumption decision	No			No			Yes			No			No		
Dilemma (resources)	No			No			No			Yes			No		
Consequences of meat consumption	No			No			No			No			Yes		

The revised interview guide of the preliminary study was used to identify and interpret the structuring of the moral conceptions. The collected data was analysed according to the Qualitative Content Analysis (Gropengießer, 2008): Firstly, the recording of sound and images takes place. The statements and actions of the students are documented consistently. Secondly, the statements are transcribed literally into a readable form with the help of transcription rules. Thirdly, we reduced the statements in consideration of the context (redacted utterances). The fourth step (context-oriented utterances)

is the beginning of the data evaluation. In consideration of the research questions, we deductively arranged the statements to summarize meaningful sections. Fifthly, we explicate the statements by analysing linguistic aspects according to the Cognitive Metaphor Theory (Gropengießer, 2007; Lakoff & Johnson, 1999) and the EMMS. This paper focuses on the development and perception of the EMMS as an analytical tool for identifying and analysing moral conceptions. We sequenced and illustrated the statements by examples. Sixthly, we summarized the statements by individual concepts.

Findings

After analysing the conducted interviews, we were able to identify fundamental conceptions. We found statements about knowledge, personal experience, values, and attitudes. According to our research question (1), we explicitly focused on the identification of the moral conceptions in students' arguments about meat consumption, which are connected to their fundamental conceptions (Figure 1). Those students who used more progressive value orientations also described their family models as more partnership-democratic, without violent structures (e.g. "At my home there is no boss") and vice versa (e.g. "My father is the boss of the family").
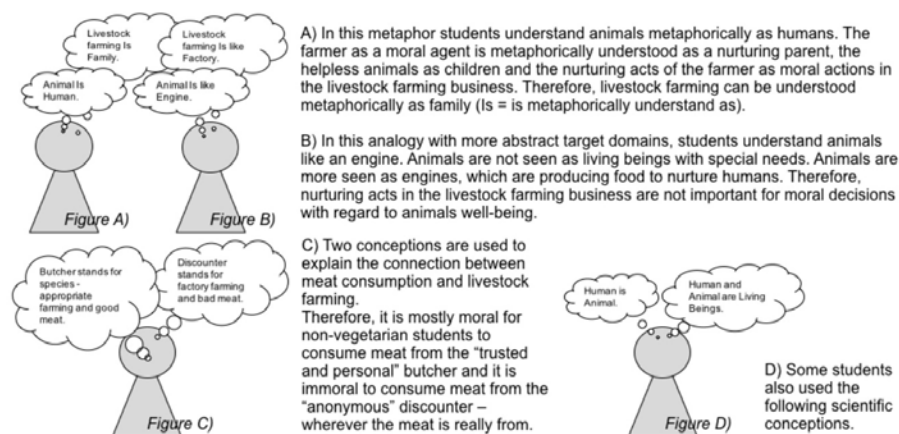
Figure 1 Students' conceptions about meat consumption.

The following section illustrates the investigated students' moral conceptions, several structures in moral conceptions and the process of comprehension (Table 2).

Ruler Metaphor: Asked to justify her meat consumption, Eva (15 years old) expresses herself as follows: "I think humans are on *top* here on earth, because there's also God, but he's not directly on earth. Humans are on *top* of this world and they can *decide* about everything and man has continuously *evolved* in that way. I think that puts humans *above* animals". Eva uses terms such as "top" and "above", which indicate the use of the Ruler Metaphor as well as an underling kinaesthetic schema, such as the Up-Down Schema (Table 3). Luisa (14 years old) argues: "Some people think that animals are *subordinate* to people and we have the *right* to treat animals like this (...) that they are not *worth* as *much* as people. (...) I think that animals are also living things (...) and we can also feed ourselves with vegetarian products. In my opinion, animals should have *equal rights*". With this statement, Luisa represents the concept of equality between humans and animals. In contrast to Eva, the vegetarian Luisa uses the concept of morality in a progressive form and uses the scientific conception that *Humans and Animals are Living Beings* (Figure 1). Luisa understands the human-animal relationship not as "superior or inferior" but as "equivalent". But in all cases, order is articulated in terms of morality.

Be Good Metaphor: If you ask Peter (10 years old) to explain the reason for his vegetarian lifestyle, he argues as follows: "I have been thinking every day how animals are killed. For example, the chickens in America are killed horribly and that is not *hygienic* for human beings". Here, Peter uses the term "hygienic", which indicates the use of the metaphor of moral purity and health, described by Lakoff & Johnson (1999). Students like Peter understand immoral people as sick, infectious and unclean.

Accounting Metaphor: If one asks students about livestock farming, conventional factory farming is mostly evaluated as "unfair". Students account for the actions of human beings in order to reach a moral equilibrium. Milan (15 years old) likes to eat meat: "I think it's not *fair* that animals live in factory farming, because animals do not *deserve* to live so badly, because they have done nothing wrong to get *punished*". Milan uses terms such as *deserve*, which indicates the use of the Accounting Metaphor. By accounting values, he tries to balance moral accounts with the concept of restitution. Peter (10 years old) accounted according to another concept: "If a human being is charged for murder he/she will be sent to prison right away, but if humanity kills billions of animals each year then not a single person is sent to jail. A living being is a

living being, right? (...) But someday nature *will get its own back with* hurricanes and tornados. Many people will die. That's a kind of nature's *revenge*, I`d say. I guess nature wants to say, it is not *right* to kill animals and so you *get punishment*". Peter uses the concept of revenge by balancing something bad "kills billions of animals" with something bad "nature gets its own back". Just as the vegetarian Luisa, he uses the concept of morality in a progressive form and the scientific conception of *Humans and Animals are Living Beings*.

Freedom Metaphor: Marie (11 years old) articulates her conceptions as follows: "If a lot of animals are in a stable and they are not allowed to *go out*, then in my opinion [livestock farming] it's stupid. But if the animals have enough *space* to move about and they are always allowed to *go out* [into nature] and *into* [the stable], then in my opinion it's good". Marie recognized moral behaviour as positive external freedom, which consists of the absence of mental and physical barriers. The limitation of freedom is evaluated as negative external freedom and therefore immoral for her. From a cognitive linguistic perspective, it can be interpreted that morality is metaphorically understood as freedom of action. Freedom of action and movement can be perceived as a moral act and the restriction of freedom as an immoral act.

Empathy Metaphor: If you ask students about their arguments on livestock farming, the vegetarian Nora (15 years old) expresses her idea as follows: "The animal feels bad because *I would also feel* bad if I must be cooped up in such a small space with lots of others. That is not *fair* to animals because we would not treat ourselves like that". Nora argued by showing compassion and created an analogy between the animals' and her own discomfort. The vegetarian Nora uses the concept of morality in a progressive form and connects it with the conception that *Animal Is Human*.

Table 2 Moral conceptions about meat consumption.

Moral Categories	Moral Conceptions		
(1) Ruler Metaphor	human-animal relationship as equal order	human-animal relationship as hierarchical order	human-animal relationship as circulation
(2) Be Good Metaphor	moral behaviour is purity and health		immoral behaviour is dirt and disease
(3) Accounting Metaphor	justice by restitution	justice by exchange	justice by revenge
(4) Freedom Metaphor	moral is freedom of movement and action		moral is control
(5) Empathy Metaphor	egocentric empathy (anthropomorphic)		absolute empathy (anthropocentric)

4. Discussion

With respect to our research question (1), the following section will discuss the relation between students' conceptions and moral conceptions, the genesis of moral arguments and the perception of the EMMS: Lakoff & Johnson (1999) explain moral conceptions made by human beings towards human beings, see also Table 3. Except in the metaphor of natural order, they are not explicitly including animals in their theory. However, students use moral metaphors towards animals and include animals in their morality. One reason for this could be that some students understand animals metaphorically as human beings (e.g. *Animal Is Human*) and use an egocentric empathy (Empathy Metaphor). Some students do not use the conception of a special position of human beings in their arguments and in these arguments, human beings do not have more rights than other animals. We already know that social discriminations such as racism, homophobia, or chauvinism can be overcome if groups of people give up their conceptual "elevated" status compared to other groups such as slaves, foreign nations, homosexuals or women. Moral philosophers such as Reagan (2001) have already attempted to overcome this conceptual "heightened" position of human beings over animals for the last 40 years. For instance, they campaign for special animal rights, therefore including a moral status of animals. Scientists such as Reagan or also Darwin (Whye, 2002), as well as students make moral decisions with the use of metaphors, for instance the Ruler Metaphor. Scientists and students basically place human beings in the animal kingdom. Some students understand the human-animal relationship as hierarchical and do not assign humans to the animal kingdom.

This means that scientists and students use metaphors. The main difference with using metaphors is the more or less reflected ways of taking scientific conceptions into account.

Some students used more progressive value orientations and described their family models as more partnership-democratic. Especially vegetarian students used altruistic and progressive motives in their arguments and more often the Empathy Metaphor. This finding is consistent with Lakoff & Johnson (1999), in which social and physical childhood experiences are closely connected to moral thinking and behaviour. The study by Samuel & Oliner (1988) presents the motives and background of people who displayed civil courage in Nazi Europe: Rescuers were able to share feelings with victims, they were more empathetic and sensitive than non-rescuers, they bore more social responsibility and gained their first social experiences in non-violent families, and this is also a confirmation of our findings. Therefore, experience with "nurturing parent family morality" can rather be an important requirement of creating altruistic arguments, and experience with "strict parent family morality" can rather promote hierarchical thinking and actions. This is because children not only gain these experiences in their families, but also in other environments, such as in school. As the school is also a place for gaining experience, one should take this into consideration, for instance in the teacher-student relationship and the structure of learning environments. According to our second research question, this section will discuss the categories themselves. The results are used to review the EMMS.

The results of the *Ruler Metaphor* are consistent with the results of Meier et al. (2007). The researchers found with the aid of four experiments that people associate abstract things such as God, power and the morally good with "up", and this spatial elevated position has a measurable positive impact on their moral actions (including cooperative, helpful, empathetic, prosocial behaviour). Several students understand human beings to be the legitimate rulers of an unmanageable but adaptable and exploitable environment. Existing publications in Germany about teaching evolution (e.g. Groß, 2007; Zabel & Gropengießer, 2015), show that students have the conception that the evolutionary development of living beings has a goal, and humans are classified as being at the "top". A disagreement with scientific ideas exists in both cases. For instance, Darwin wrote: "It is absurd to talk of one animal being *higher*

than another." (Whye, 2002). But on the other hand, several (mostly vegetarian) students do not place human beings in a higher position but rather give animals basic rights such as freedom. The Ruler Metaphor was used to explain the relationship between human beings and animals.

The *Be Good Metaphor* is characterized by the evaluation of moral or immoral behaviour (the evaluation of human traits) and the quest for moral goodness. For instance, cleanliness (e.g. "ethnic cleansing") is used as a metaphor for moral statements, and immoral behaviour is understood as experienced impurity (e.g. Kattmann, 2013; Zhong & Liljenquist, 2006).

Accounting Metaphor: Interpreted by Piaget, Peter is in the phase of transition. He emphasizes justice and equality and recognizes that a group determines rules. Students also use utilitarian arguments, in which nature appears as the avenger in other contexts such as stem cell therapy. Metaphors such as the "right to life" are found in several biology teaching studies (e.g. Hößle, 2003, p.65).

Freedom Metaphor: In the western world, freedom is one of the highest values. For Kohlberg, freedom plays a central role in his undetectable sixth stage of moral development.

Empathy Metaphor: Students transfer their own childhood experiences, which reduce or create a feeling of well-being in respect to livestock farming in order to justify their judgement. Students can either use egocentric empathy (orientation on their own value system) or the opposite, such as absolute empathy (orientation on a foreign system of values) (Lakoff & Johnson, 1999, pp. 309-310). By using egocentric empathy, animals are metaphorical understood as humans with analogous needs, feelings and behaviour (*Animal Is Human*). However, absolute empathy is used in connection with the scientific concepts such as *Humans and Animals are Living Beings* with special needs, feelings and behaviour.

Table 3 Essence of moral metaphors.

Moral Metaphors	Basis Metaphors (Lakoff & Johnson, 1999)	Source Domains (experiences with well-being)	Underlying Kinaesthetic Schemas*	Target Domains	Examples (from students)
(1) Ruler Metaphor	Metaphors of moral authority - order - strength	- family models - power relations - physical strength - role models - verticality	- Up-Down Schema - Part-Whole Schema - Linear Order Schema	human - animal relationship (quantity)	humans are above animals
(2) Be Good Metaphor	Metaphors of moral essence - purity - health	- purity - cleanliness - dirt - health - disease - transfer	- Donor-Donation-Recipient Schema	good, bad behaviour for humans against animals	she has a heart of gold; this behaviour is not hygienic
(3) Accounting Metaphor	Metaphors of moral accounting - reciprocation - retribution/ revenge - altruism, fairness - restitution - „Turning the Other Cheek" - Karma - rights	- (im-)materiel values - accounting, transfer - credits, depts - balance	- Donor-Donation-Recipient Schema - Balance Schema	justice, balance in the human - animal relationship	we have to give something back to the animals
(4) Freedom Metaphor	Metaphors of moral freedom - bounds - rights	- freedom of movement and action - limits of freedom - bounds	- Source-Path-Goal Schema - Container Schema	freedom, rules for animals and humans	animals have the right to be free
(5) Empathy Metaphor	Metaphors of moral empathy - nurturance	- parental care, nurturance and empathy - dependence of kids on their parents	- Link Schema - Person Schema	empathy, behaviour against animals and humans	we have to treat animals like we would treat humans

* based on: Up-Down Schema (Lakoff, 1987, pp. 275-278, Johnson, 1987, p. 119; Groß, 2007, p. 172); Container Schema (Lakoff, 1987, pp. 272-273; Lakoff & Johnson, 1999, pp. 31-32; Lakoff & Johnson, 2003, pp. 23-27; Source-Path-Goal Schema (Lakoff, 1987, p. 275; Lakoff & Johnson, 1999, pp. 32-34; Gropengießer, 2007, p. 109); Part-Whole Schema (Lakoff, 1987, pp. 273-274); Linear Order Schema (Lakoff, 1987, p. 274); Linear Order Schema (Lakoff, 1987, p. 275); Balance Schema (Johnson, 1987, p. 74); Person Schema (Lakoff & Johnson, 1999, pp. 34-36; Lakoff & Johnson, 2003, pp. 28-29; Donor-Donation-Recipient Schema (Gropengießer, 2007, p. 109).

The described EMMS (Figure 2) has been developed on the basis of Lakoff & Johnson´s moral metaphors as well as empirically by means of creating five categories. The EMMS is useful to identify and analyse moral statements and includes all moral metaphors described by Lakoff & Johnson (1999). Sometimes, especially in the Empathy and Freedom Metaphor, overlaps exist. These overlaps show that the moral categories are neither completely separable nor completely linked. However, there is still a possibility to identify further possible, not assignable moral metaphors.
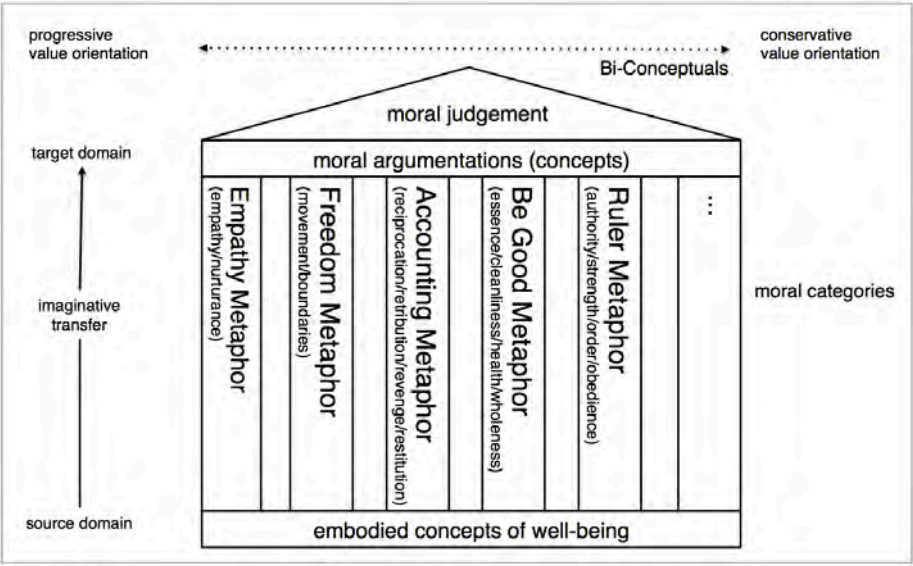
Figure 2 Educational Moral Metaphors System.

5. Conclusions

This study has illustrated that a cognitive-linguistic approach can be used to develop an educational tool for the interpretation of moral conceptions about SSIs, and that their genesis and moral conceptions for livestock conservation and meat consumption are experience-based. Both Rundgren & Rundgren (2010 in respect to the SEE-SEP model as an analytical framework for SSIs), and Lakoff & Johnson (1999 in respect to metaphorical structures in our moral thoughts) recognise and take into account the context of morality and personal experiences. Ways of thinking are structured by five moral metaphors. Moral arguments are not static and cannot be considered as a linear model – they

are highly dynamic. A specified number of moral concepts exist like the keys of a piano. But during decision-making, moral conceptions can be used in different ways to play different moral tunes and melodies. With the help of this perspective, it is possible to analyse our students' moral conceptions. These educational offers can help with regard to the description of ethical problems, the reflection of modes of action, respect characteristic moral conceptions in order to assist students to make responsible consumer decisions, including different viewpoints and perhaps one way to make morality discussible. With regard to metaphor pluralism, new educational opportunities for teaching biology are realisable. Through the knowledge of the importance of experiences and their imaginatively transfer to new target areas, the influences of moral development, which have been observed up to now, can be used in a more comprehensive understanding of the genesis and structure of moral concepts. The results have links to existing models of moral judgment (e.g. Haidt, Kohlberg, Piaget; for more details see Johnson, 2014; Tramowsky et al., 2016). The study of moral conceptions seems to be particularly relevant because of the (almost) daily consumption of animal products and the multifaceted (also not private) consequences. The consideration of moral thoughts about SSIs in biology lessons also opens up the possibility to develop cross-curricular learning offers for the sustainable discourse on production and consumer behaviour.

Acknowledgment

References

Binngießer, J., Wilhelm, C. & Randler, C. (2013) Attitudes toward Animals among German Children and Adolescents. *Anthrozoös*, 26(3), 325-339.

Boerwinkel, D. J. & Waarlo, A. J. (2010) Genomics education for decision-making: research on socio-scientific learning and teaching. *Genomics Education for Decision-making*, 2, 15.

Brocos, P. & Jiménez-Aleixandre, M. P. (2016) *Analysis of justifications in pre-service teachers' socioscientific arguments about diets*. Paper presented at the 11th Conference of European Researchers in Didactics of Biology. Karlstad, 67.

Duit, R., Gropengießer, H., Kattmann, U., Komorek, M. & Parchmann, I. (2012) The Model of Educational Reconstruction – a Framework for Improving Teaching and Learning. In *Science education research and practice in Europe* (pp.13-37). Rotterdam: SensePublishers.

Gallese, V. & Lakoff, G. (2005) The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive neuropsychology, 22(3-4)*, 455-479.

Gerstenmaier, J. & Mandl, M. (1995) Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik*, *41*, 867-888.

Gropengießer, H. (2007) Theorie des erfahrungsbasierten Verstehens. In D. Krüger & H. Vogt (Eds.), *Theorien in der biologiedidaktischen Forschung* (pp.105-116). Berlin: Springer.

Gropengießer, H. (2008) Qualitative Inhaltsanalyse in der fachdidaktischen Lehr-Lernforschung. In P. Mayring & M. Gläser-Zikuda, (Eds.), *Die Praxis der Qualitativen Inhaltsanalyse* (pp.172-189). Weinheim: Beltz Verlag.

Groß, J. (2007) *Biologie verstehen: Wirkungen außerschulischer Lernorte*. Oldenburg: Didaktisches Zentrum.

Haidt, J. (2012) *The Righteous Mind*. New York: Pantheon.

Hößle, C. (2003) Der Embryo – Mensch von Anfang an? [Electronic version]. In S. Albrecht, J. Dierken, H. Freese & C. Hößle (Eds.), *Stammzellforschung – Debatte zwischen Ethik, Politik und Geschäft* (pp.43-75). Hamburg: University Press.

Johnson, M. (1987) *The Body in the Mind*. Chicago: University of Chicago Press.

Johnson, M. (2014) *Morality for Humans*. Chicago: University of Chicago Press.

Kattmann, U. (2013) Genes, race and culture. In M. Koegeler-Abdi & R. Parncutt (Eds.), *Interculturality: Practice meets research*. Newcastle upon Tyne.

KMK [Eds.] (2004) *Bildungsstandards im Fach Biologie für den Mittleren Bildungsabschluss* [Electronic version]. München: Luchterhand.

Kohlberg, L. (1976) *Moral stages and moralization: The cognitive-developmental approach. Moral development and behaviour: Theory, research, and social issues*, 31-53.

Komorek, M. & Duit, R. (2004) The teaching experiment as a powerful method to develop and evaluate teaching and learning sequences. *International Journal of Science Education*, *25*, 619-633.

Lakoff, G. (1987) *Women, fire, and dangerous things*. Chicago: University of Chicago Press.

Lakoff, G. & Johnson, M. (1999) *Philosophy In The Flesh*. New York: Basic Books.

Lakoff, G. & Johnson, M. (2003) *Metaphors we live by.* London: University of Chicago Press.

Meier, B.P., Hauser, D.J., Robinson, M.D., Friesen, C.K. & Schjeldahl, K. (2007) What's "up" with God? *Journal of Personality and Social Psychology*, *93*, 699-710.

Morin, O., Simonneaux, L., Simonneaux, J., Tytler, R. & Barraza, L. (2014) Developing and Using an S3R Model to Analyze Reasoning in Web Based Cross National Exchanges on Sustainability. *Science Education, 98(3)*, 517-542.

Piaget, J. (1997) *The moral judgement of the child*. Simon and Schuster.

Phillips, D. C. (Ed.) (2000) *Constructivism in Education: Opinions and Second Opinions on Controversial Issues*. University of Chicago Press.

Regan, T. (2001) *Defending Animal Rights*. University of Illinois Press.

Rudgren, S. N. C. & Rudgren, C. J. (2010) SEE-SEP: From a separate to a holistic view of socioscientific issues. In *Asia-Pacific Forum on Science Learning and Teaching*, Vol. *11(1)*, 1-24.

Sadler, T. D. & Zeidler, D. L. (2003) The morality of socioscientific issues: Construal and resolution of genetic engineering dilemmas. *Science Education, 88(1)*, 4-27.

Samuel, P. & Oliner, P. M. (1988) *The Altruistic Personality*. New York: Free Press.

Tramowsky, N., Paul, J. & Groß, J. (2016) Von Frauen, Männern und Schweinen – Moralvorstellungen zur Nutztierhaltung und zum Fleischkonsum im Biologieunterricht (pp.171-187). In M. Hammann & U. Gebhard [Eds.], *Lehr- und Lernforschung in der Biologiedidaktik.* Innsbruck: Studienverlag.

Wyhe, J. v. (Ed.) (2002) The Complete Work of Charles Darwin Online. Notebook on Transmutation. 1837. Retrieved October, 26, 2016 from http://darwin-online.org.uk/content/

frameset?itemID=CUL-DAR121.-&viewtype=side&pageseq=1.

Zabel, J. & Gropengießer, H. (2015) What can Narrative contribute to Students' Understanding of Scientific Concepts, e.g. Evolution Theory? *Journal of the European Teacher Education Network*, 10, 136-146.

Zeidler, D. L. & Sadler, T. L. (2008) The role of moral reasoning in argumentation: Conscience, character and care (pp. 201-216) In S. Erduran & M. Pilar Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research*. New York: Springer.

Zhong, C.B. & Liljenquist K. (2006) Washing away your sins. *Science*, 313, 1451-1452.

Chapter 2:
Students´ interest and motivation

CLIL Biology – teaching biology in a foreign language: the influence of classroom language on student motivation and acquisition of knowledge

Petra Duske[1] and Michael Ewig[2]

petra.duske@mail.de

Abstract

This study deals with the question of how far Content and Language Integrated Learning (CLIL) teaching in biology influences students' motivation and their acquisition of knowledge in biology. In a CLIL approach, a foreign language is used for the learning and teaching of both content and language. The quantitative study with pre-post-follow-up design and a multivariate analysis of the quantitative data found no significant difference between the students' (n=788) motivation towards biology or English as a Foreign Language (EFL), whether they were taught in English or in German. Whereas the CLIL students' motivation correlated with their motivation in the school subjects Biology and English, it did not correlate with their performance in these subjects. CLIL students showed at least similar results in all knowledge tests (on the topic of smell) as the students taught in German. No correlation was found between the students' growth of knowledge during the CLIL unit and their motivation or competence in English. Although CLIL seems to be equally suitable for girls and boys regarding their motivation towards biology, girls seemed to learn slightly more during a CLIL unit in biology than boys.

Keywords: acquisition of knowledge in Biology, bilingual teaching in Biology, content and language integrated learning (CLIL), students' motivation for biology and English as a foreign language,

---

[1] University of Education Weingarten, Germany

[2] University of Vechta, Germany

1. Introduction

Permafrost, oxygen, hydrogen - it all looks like science to me. But these terms actually have origins in Russian, Greek and French. Today, though, if a scientist is going to coin a new term, it's most likely in English. And if they are going to publish a new discovery, it is most definitely in English.
(Porzucki, 2014)

English has become the language of science. Scientific research needs to be communicated and discussed in English in order to be recognized nationally and internationally. Therefore, enabling students to take part in scientific discourse in English has become one of the aims of (advanced) scientific education today.

Content and Language Integrated Learning (CLIL) refers to a "dual-focused educational context in which an additional language" is used as a "medium in the teaching and learning of non-language content" (Marsh, 2002, p.15). Therefore, it is well-suited to be employed to teach biological sciences in English. Furthermore, biology is a communicative science by the extensive use of real objects and other meaningful media so it seems to be especially suitable for a CLIL approach (in English), since scientific findings and socially relevant topics are discussed in the *lingua franca* of science. CLIL has not only become a popular practice of immersion education (Graaff et al., 2007), but has also been implemented in most European education systems (Eurydice, 2006; Maljers et al., 2007; Pérez-Cañado, 2012). Although CLIL is a fusion of language and subject teaching, in class the emphasis is generally put on successfully communicating the subject content (KMK, 2014). Thus, it is surprising that CLIL research generally focuses on language aspects. While numerous studies from different countries have shown the benefits of CLIL teaching for language acquisition (Coyle, 2007; Klieme et al., 2006) there are few studies in content subject didactics and even fewer in biology didactics.

Motivation is one of the requirements of successful learning and therefore, it is important to take it into account in any school context. Deci & Ryan (1993) have formulated the "basic needs" in the self-determination theory, that need to be fulfilled for a person to act (and learn) successfully. They distinguish between different types of motivation and explain its influence on a person's actions. Due to the importance of motivation in the school context,

CLIL research has investigated the development of motivation in CLIL programmes. Some small scale studies have suggested a positive effect of CLIL on student motivation (Coyle, 2006; Lasagabaster, 2011).

Although the effects of CLIL teaching on the content subject concerned have not been systematically investigated, there are some studies considering the acquisition of knowledge showing ambiguous results (Genesee, 2004; Marsh, 2002; Seikkula-Leino, 2007; Stohler, 2006). There is a great demand for systematic research conveying the impact of CLIL programmes on content subjects (Coyle et al., 2010; García, 2009; Graaff et al., 2007; Pérez-Cañado, 2012). It is particularly important that this research is carried out by subject didactics, i.e. investigating CLIL biology from the viewpoint of biology didactics. Since investigating and improving teaching and learning biology is one of the main aims of biology didactics, it is crucial for biology didactics to enter the scientific discourse into the growing field of teaching and learning biology in English. This is particularly true since CLIL programmes are being promoted by most educational authorities in Europe (cf. Eurydice, 2006).

The key objective of this study therefore is to find out how far CLIL in biology influences

- the students' motivation towards biology or EFL, and
- the students' acquisition of knowledge in biology.

2. Research design and methods

To answer these questions, a quantitative study (n=788) with pre-post-follow-up design and a control group was conducted with 29 classes (grade 9) and 13 teachers at ten junior high schools (Realschulen) in southwest Germany. To find participants, schools using the CLIL approach in Biology classes were contacted and asked if they agreed to have (at least) two classes of grade 9 students taking part in the study the following year. Grade 9 was chosen because they should know enough English to be able to focus on content learning concerning the topic of smell and are not yet involved in preparing their final exams in grade 10. Both participating classes of each school needed to be taught by the same teacher; one class in German, one class in English. This was done to increase the validity of the results of the study by reducing the effects of different teachers.

The participating classes were randomly assigned to either the CLIL or the

corresponding German biology module. To control the profile of the participating classes and students, a general questionnaire regarding the social and educational background of the students was employed. The Biology module used dealt with the 'The world of smell' and comprised the following seven consecutive lessons:

1) Our senses: Stimulus – reaction diagram
2) Testing our sense of smell
3) What is going on in my body when I smell something?
4) Adaptation – "Everybody smells my perfume except me…!"
5) The sense of smell in the animal (and human) world
6) Use and abuse of olfaction (in our modern world)
7) Losing the sense of smell: Olfactory sensors?

The topic smell as one of the (human) senses was chosen because it is part of the German curriculum grade 9, but is generally not taught by most teachers. This was important to ensure that the topic had not been dealt with before because the study was conducted at the end of the school year. The teaching material was designed according to the following didactic concepts:

• Context based learning: various studies have shown that using context based teaching material improves the identification of students with the topic and therefore has positive effects on the learning process (Pilot et al., 2016)

• Cooperative learning: the principal steps of cooperative learning "think – pair – share" are especially suitable for content learning in a foreign language since it ensures the activity of a greater number of students and stresses the communicative aspect (Green & Green, 2008)

• CLIL methodology: "scaffolding" - the idea of scaffolding is to provide a variety of helpful material to promote language and content learning. The scaffolds are provided as long as needed and the aim is to reduce them as the learner advances (Gibbons, 2014).

• Other aspects of CLIL methodology include reorganization of the same content matter in a different form (oral, spoken, written, figure, cartoon, comment). This is necessary to be able to communicate about the subject matter in a foreign language but also supports subject learning.

Detailed teaching guidelines and teaching material, as identical as possible, were provided for both modules on 'The world of smell' in German and English respectively. This included detailed descriptions of each lesson, worksheets, and laboratory kits (including for example, olfactory samples) for group work and material to be used to illustrate aspects of smell and perception in class. All participating teachers were personally instructed about teaching and test procedures. Figure 1 shows the design of the study.

Pre-test	Teaching modules	Post-test	4 weeks	Follow-up-test
General questionnaire Knowledge test (German) Motivation test Biology English	7 lessons: **German** **CLIL** Control group: no lessons on topic	Knowledge test (German/English) Motivation test Module Evaluation of unit		Knowledge test (German/English)
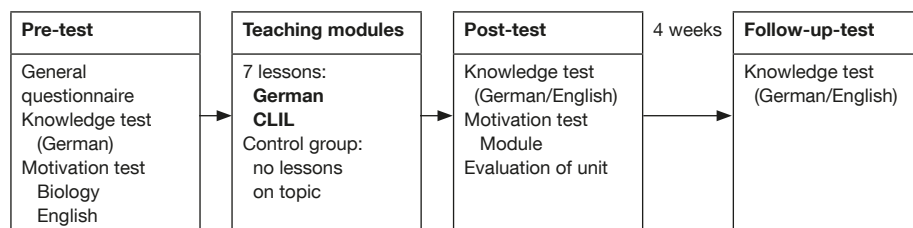
Figure 1 Design of the main study.

The main study started with the pre-test, comprising a general questionnaire, a German knowledge test on the topic of smell and perception, and a motivation test concerning the school subjects Biology and English. All students taking part in the Biology module (either in German or in English) did the post-test in the following lesson. It comprised a knowledge test in English and in German as well as the motivation test showing the influence of the Biology module. Four weeks later, the German and the English knowledge test was repeated to find out whether the learning effects were persistent.

While students taking part in the German or the CLIL module took part in the pre-, post- and follow-up tests, the control group only did the pre- and the post-test. They did not take part in the Biology module. To reduce the impact on the control group, they did not do the follow-up test. The control group was employed to look for and exclude pre-test effects. The general questionnaire comprised questions about the personal and educational background of the students and the students' school performance in different subjects. The motivation test comprised 9 scales with 50 items, which were adapted from a test by Seidel et al. (2003) investigating the motivation of students in physics. Agreement to the statements in the questionnaires could be marked in a four step format. The questionnaire comprised the following scales: Subject-specific self-concept, cognitive activity, quality of motivation,

motivational orientation, supporting factors in class, error culture/no blame culture. The items were recorded separately and used to form a motivation index, ranging from 1-4 (4 displaying the highest motivation).

The knowledge was developed according to the content of the Biology module on 'The world of smell'. The knowledge test as well as the motivation test and the teaching modules were tested and optimized in two preliminary studies. Before the intervention, the knowledge test was conducted in German. After the module it was both conducted in German and in English, because all students should be tested in the language of instruction. But it is only possible to compare the results of the tests if they use the same language. If students are doing the same test twice it is likely that this affects the outcome of the test. To control this fact, half of each class did the German knowledge test first and the other half did the English test first. The knowledge test consisted of 8 scales with 48 items. The scales included: sense organs, adequate stimulus, Brownian movement, information processing, adaptation, animal world, as well as use and abuse of the sense of smell. The knowledge test comprised Right-or-wrong-questions (12 items), Multiple choice questions with one correct answer and 3 distractors each (8 items), Matching tasks (13 items), and free format questions (labelling) (15 items). This mixed format was chosen to balance the advantages and disadvantages of each testing format. (Bühner, 2006).

For the multivariate analysis of the quantitative tests SPSS 21 has been used. The results of the knowledge test were encoded separately, then used to calculate the reliability of the test and later converted into a total score.

The independent variables 'classroom language', 'classroom context' (not dealt with in this paper) and 'sex' have been used as factors in the multivariate analysis. The dependent variables in this model are the total scores in the pre-, post- and follow-up tests and the results of the motivation test for the modules. The different covariates were tested for their significance. Due to significance and minimizing unwanted effects, the following covariates have been included in the multivariate analysis: 'score in pre-test', 'order of tests', 'performance in languages', 'performance in sciences', 'motivation in Biology', 'motivation in English'.

3. Findings

3.1 Control group

Students of the control group (n=70) had a similar mean score in the German knowledge pre-test (M=20.7 points; SD=6.65) (out of 48 possible points) as in the German knowledge post-test (M=21.0 points; SD=5.95) conducted 4 weeks later. Without any input, the mean score in the English knowledge post-test was significantly lower (M=13.7 points; SD=5.84). In order to allow t-tests, the knowledge test has an interval scale and the results have been positively tested on normal distribution (Kolmogorov-Smirnov, p>.05, and Q-Q-diagrams). T-tests for dependent samples applied to the mean scores of the German knowledge pre- and post-test of the control group (T=-.451; df=69) are not significant (p>.05) and show that there are no pre-test effects. Since there is a strong correlation (r=.615, p<.001) between the mean scores in the German knowledge pre- and post-test it can be assumed that the test score is reliable (Rasch et al., 2010). The mean score of the English knowledge post-test (M=13.7 points; SD=5.84) was lower than in the German knowledge pre-test. The differences (T=10.387; df=69) are highly significant (p<.001). Since the correlation between these two tests is very low (r=.593, p<.001), it is not only content knowledge that accounts for answering the English knowledge test. Applying the same knowledge test after 4 weeks did not cause any pre-test effects.

3.2 Descriptive findings

Figures 2 and 3 show the distribution of the sample according to different factors.
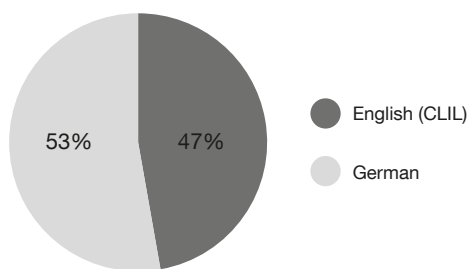
Figure 2 Distribution of sample on different modules.
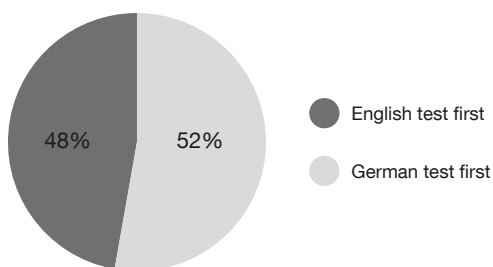
Figure 3 Distribution of sample on the order of tests.

Before the module, the motivation index for English (as a school subject) with a mean value of 2.4 (mean deviation: .617) was marginally lower than for Biology with a mean value of 2.5 (mean deviation: .612). In the knowledge pre-test, the students mean score was 21.3 out of a maximum of 48 points. Figure 4 shows the mean score of the students in the German knowledge pre-test who later did the knowledge post-test in different languages in a different order.
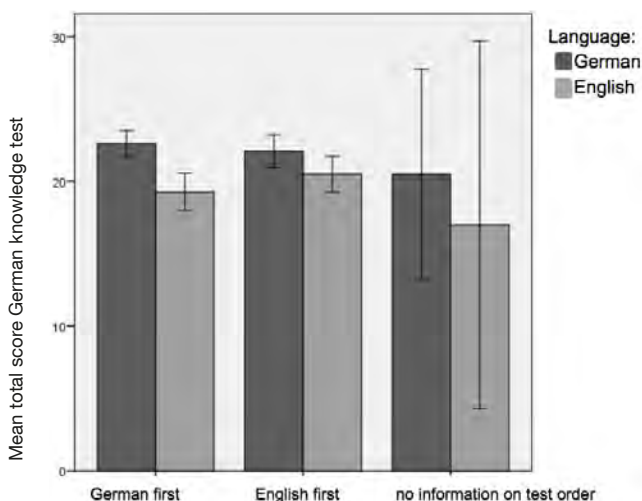
Figure 4 Distribution of sample on the different test order.

To control the influence of the test order, half of each class did the German test first and then the English test and vice versa. As expected, the students scored higher in the German than in the English test. Unexpectedly, doing

the English test after the German was not beneficial for the mean score. But due to the different mean scores in both groups, the order of the tests needs to be considered as a covariate in the multivariate analysis. The results for the Between-subject-effects of the corrected model of the multivariate analysis are shown in Table 1. They are significant (p<.001) and show that all dependent variables can be partly explained by the factors and the independent variables.

Table 1 Results of the Between-subject-effects of the corrected model of the multivariate analysis.

Origin	Dependent variable	Square sum Typ III	df	Mean of square	F	Sig.	Partial Eta-square
corrected model	Improvement post-test German	8393,361[a]	13	645,643	20,590	,000	,414
	Improvement follow-up-test Ger	9806,202[b]	13	754,323	25,381	,000	,465
	Improvement post-test English	12785,691[d]	13	983,515	24,319	,000	,455
	Improvement follow-up-test E	11179,864[e]	13	859,990	23,076	,000	,442
	Motivation- Index: CLIL	53,311[g]	13	4,101	18,455	,000	,388

Since eta-square (η^2) shows the explained variability of the results of the sample and is being used as a measure for the effect size, the effects described by this model are intermediate. (Rasch et al., 2010).

3.3 Student motivation

There is no significant difference between the students' motivation after the module taught in English (x=2.23; σ_x=0.04) and in German (x=2.39; σ_x=0.03) (p>.05). Taking a closer look at the CLIL students, we see that the CLIL students' motivation correlates with their motivation in the school subject Biology (r=.511; p<.001; n=181), even showing a great effect. The students motivation after the CLIL unit also correlates with their motivation in English (r=.442; p<.001; n=180). Unexpectedly, there is no significant correlation between the motivation-index after the CLIL unit and the students' performance in science (r=-.023, p>.05; n=181), or in English (r=-.047; p>.05; n=181). The slightly higher motivation-index of girls (x=2.35; σx =0.034) in comparison to boys (x=2.28; σx=0.036) taking part in the CLIL module is not significant (F=1.960; df=1; p>.05).

3.4 Acquisition of knowledge

CLIL students show similar or better results in the knowledge tests on 'The world of smell' as students taught in German. Figure 5 shows the results for

the German knowledge test (post-test: F=0,401; df=1; p>.05 / follow-up test: F=0.004; df=1; p>.05). Although the mean total score for CLIL students was lower in the pre-test, they did better in the post-test. But these differences are not significant.
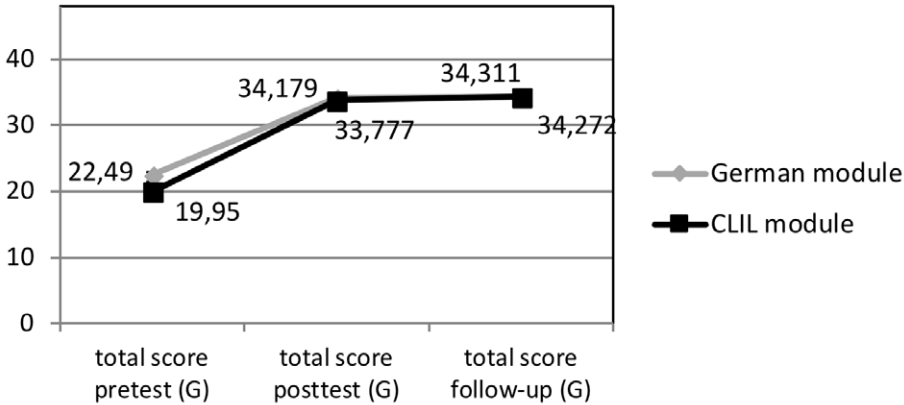
Figure 5 Mean score of German knowledge test in CLIL and German Biology module (max. 48 points).
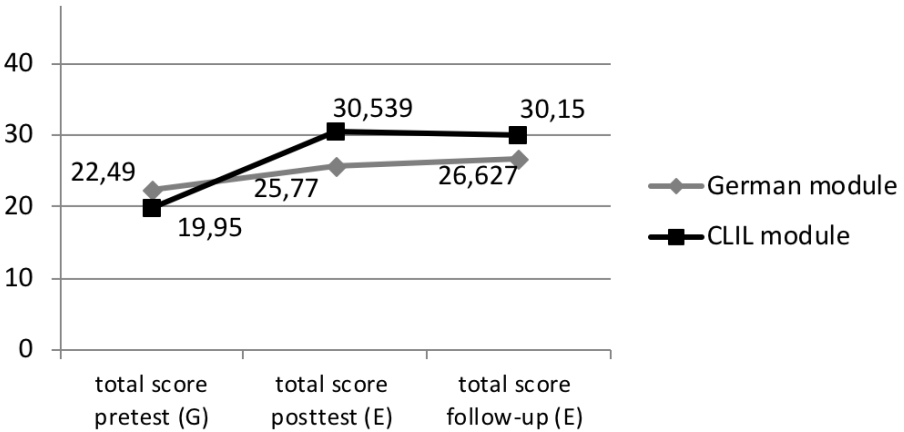
Figure 6 Mean score of English knowledge test in CLIL and German Biology module (max 48 points).

Figure 6 shows the results for the English knowledge test. As anticipated, the CLIL students displayed greater improvement towards the post-test (F=43.810; df=1; p<.001) and the follow-up test (F=25.945; df=1; p<.001) than the students taught in German. Taking a closer look at the CLIL students

only, it is surprising to see, that no correlation has been found between the students' acquisition of knowledge during the CLIL unit (in Biology) and their motivation for English as a school subject (r=.068; p=373; n=176) or their performance in English (r=-.057; p=.448; n=177). Taking gender issues into account we see that girls seem to learn slightly more during the CLIL as well as the German biology module than boys. The factor 'acquisition of knowledge' correlates with the factor 'sex' (r=-.213; p=.004; n=177). The reported effect for this correlation is small. This can also be seen in Figure 7, showing the mean score for boys and girls in the CLIL module.
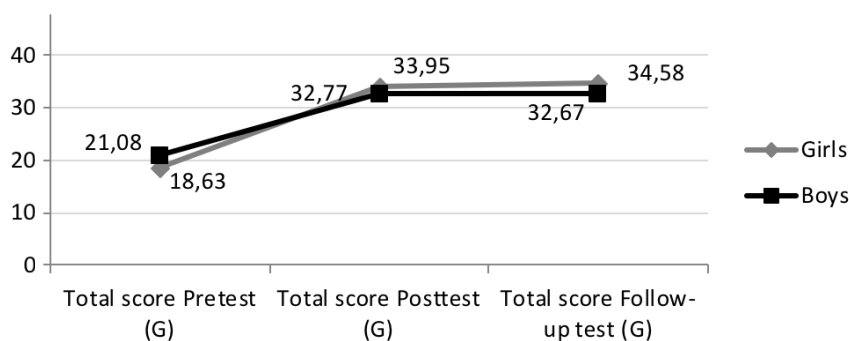
Figure 7 Mean score of knowledge test of boys and girls of CLIL group (max. 48 points).

4. Conclusion

The popularity of CLIL for students, parents and teachers in many European countries has often been explained by its effect on the students' motivation (Richter, 2004). This notion is supported by a number of empirical studies (e.g. Coyle, 2006; Eurydice, 2006; Lasagabaster, 2011; Oetter, 2005). Those studies were conducted in bilingual schools (or branches), where students attend additional language classes and always had one or several CLIL subjects. Some authors admit that, especially with younger or less advanced students, enjoyment and motivation might suffer due to the initial challenges of CLIL (Coyle et al, 2010; Yassin et al., 2009). Although this fairly short module was the first CLIL experience for most participating students and the study was conducted at German junior high school (*Realschulen*), there was no such effect. The results presented suggest that CLIL modules in biology are suitable

to sustain the students' motivation for Biology. Furthermore, CLIL classes seemed to be suitable to motivate both boys and girls in Biology.

Contrary to Yassin et al. (2009) and a preliminary study generating hypothesis for the present survey, no correlation has been found between the performance in English and the motivation in the CLIL module. The assumption that this was the case has been based on the theory of self-determination by Deci & Ryan (1993). Realizing oneself as being capable of achieving certain goals meets human basic needs, which is one of the requirements of motivation. One reason why this is not the case in the work presented could be the relative inaccuracy of school grades, which have been used to measure language performance. It is also possible that scaffolding, one of the methods employed in CLIL, has successfully minimized language problems (cf. Gibbons, 2014). As expected, there is no correlation between the motivation in the CLIL module and the performance in biology. This might be due to a relatively high motivation of most students in biology, often regardless of their performance in this subject. Unlike in Biology, motivation and performance correlated in other sciences (Schiefele et al., 1993). However, motivation in the CLIL unit is closely linked with motivation in Biology and in English, as motivation for different schools subjects has been found to correlate in other studies (Schiefele et al., 1993).

While critics fear that CLIL impairs the acquisition of knowledge in the content subject (Coyle et al., 2010; Stohler, 2006), this study shows that CLIL modules are suitable to impart subject knowledge in biology. On the contrary, other studies had found that CLIL students achieved higher academic standards (Badertscher & Bieri, 2009). However, those studies, again, took place in schools with CLIL branches where the students are chosen according to their academic performance, like a study in which CLIL students scored higher in PISA items in Biology than the German control group (Osterhage, 2007). Similar results are reported from Spain, where students for CLIL programmes are being preselected as well (Zarobe & Lasagabaster, 2010). Some authors report similar results for CLIL despite more complex requirements with the didactic structure of many CLIL arrangements (Badertscher & Bieri, 2009). This is not true for the study presented here, since the two modules on 'The world of smell' had the same structure. The CLIL methodology including scaffolding, reorganizing and communication of the content matter in various

ways, may lead to more student activity and engagement and may support persistent learning (Badertscher & Bieri, 2009; Coyle, 2010).

CLIL is traditionally found in more academic schools, but some authors suggest that it is also suitable for heterogeneous classes with less advanced students (Association for Language Learning, 2008; Coyle et al., 2010; Mehisto et al., 2008). Examples for this model are Singapore, Luxembourg or the Netherlands, where multilingual education has become standard. It has been put forward, that in heterogeneous multinational classes, CLIL can be a way of reducing disadvantages within the education system since the language of discourse is a foreign language for all students (Hallet, 2007). It allows students and teachers to take the perspective of a growing number of students who cannot speak their mother tongue at school. The need for teachers to spend more time on thinking about and explaining technical terms can also have positive effects on standard biology classes.

The small effect presented for the different acquisition of knowledge among boys and girls does not allow us to draw any firm conclusions. There is no evidence that CLIL disadvantages boys. This is important to provide suitable education for all students. Therefore, CLIL can be an effective and enjoyable way of learning both biology and an additional language. Unlike language classes, CLIL Biology enables students to talk about science in another language than their mother tongue (Kearsey & Turner, 1999; Baker, 2002). This is of particular importance since being able to discuss scientific matters in English is the basis of scientific discourse in a global community (cf. Graddol, 2004). Giving consideration to the dual-focused educational approach of CLIL, there is a need to strengthen the biological perspective in the CLIL classroom. By participating in the academic debate on CLIL, biology didactics can contribute to the development of innovations appropriate for modern education in the EU.

References

Association for Language Learning (2008) *Linking French and Geography – the Earth's structure, tectonic activity, volcanic formation and eruption: Subject Specific Support for the New Secondary Curriculum*. Leicester: Association for Language Learning.

Badertscher, H. & Bieri, T. (2009) *Wissenserwerb im Content and Language Integrated Learning: Empirische Befunde und Interpretationen*. Bern: Haupt.

Baker, A. (2002) Der Erwerb von akademischen Sprachfähigkeiten im Englischen durch bilingualen

Sachfachunterricht am Beispiel Biologie. In B. Krück, & K. Loeser, (Eds.), *Innovationen im Fremd-sprachenunterricht. 2 Fremdsprachen als Arbeitssprachen* (pp. 133-152). Frankfurt/Main: Lang.

Bühner, M. (2006) *Einführung in die Test- und Fragebogenkonstruktion* (2nd Edition). München: Pearson Studium.

Coyle, D. (2006) Content and language integrated learning: Motivating learners and teachers. *Scottish Languages Review* (13), 1-18.

Coyle, D. (2007) Content and Language Integrated Learning: Towards a Connected Research Agenda for CLIL Pedagogies. *International Journal of Bilingual Education and Bilingualism*, 10 (5), 543-562.

Coyle, D., Hood, P. & Marsh, D. (2010) *CLIL: Content and Language Integrated Learning.* Cambridge: Cambridge University Press.

Deci, E. & Ryan, R.M. (1993) Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik* 39 (2), 223-239.

Eurydice (2006) *Content and Language Integrated Learning (CLIL) at School in Europe.* http://www.indire.it/lucabas/lkmw_file/eurydice/CLIL_EN.pdf [30.11.2016].

García, O. (2009) *Bilingual Education in the 21st Century: A Global Perspective.* Malden: Wiley-Blackwell.

Genesee, F. (2004) What do we know about Bilingual Eduacation for Majority-Language Students? In T. Bhatia & W. Ritchie (Eds.), *The Handbook of Bilingualism* (pp. 547–576). Malden: Blackwell.

Gibbons, P. (2015) *Scaffolding language – scaffolding learning: Teaching English Language Learners in the Mainstream Classroom.* Portsmouth: Heinemann.

Graaff, R. de, Koopmana, J., Anikinaa, Y. & Westhoff, G. (2007) An Observation Tool for Effective L2 Pedagogy in Content and Language Integrated Learning (CLIL). *International Journal of Bilingual Education and Bilingualism*, 10 (5), 603-624.

Graddol, D. (2004) The Future of Language. *Science,* 303, 1329-1331.

Green, W. H. & Green, R. H. (2008) *Basic Moves of Teaching: Building on Cooperative Learning.* Bloomington: Trafford.

Hallet, W. (2007) Zwischen Bildungsstandards und Mehrsprachigkeit: Kompetenz-erwerb im Bilingualen Unterricht. In: P. Bosenius, J. Donnerstag & A. Rhode (Hrsg.), *Der bilinguale Unterricht Englisch aus der Sicht der Fachdidaktiken* (S. 17–35). Trier: Wissenschaftlicher Verlag.

Kearsey, J. & Turner, S. (1999) The value of bilingualism in pupils´ understanding of scientific language. *International Journal of Science Education*, 21(10), 1037-1050.

Klieme, E., Eichler, W. & Helmke, A. (2006) *Unterricht und Kompetenzerwerb in Englisch: Zentrale Befunde der Studie Deutsch-Englische Schülerleistungen International (DESI).* Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.

KMK - Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2014) *Konzepte für den bilingualen Unterricht: Er-fahrungsbericht und Vorschläge zur Weiterentwicklung.* Verfügbar unter: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2013/201_10_17-Konzepte-_bilingualer-_Unterricht.pdf [21.6.2014].

Lasagabaster, D. (2011) English achievement and student motivation in CLIL and EFL settings. *Innovation in Language Teaching and Learning*, 5 (1), 3-18.

Maljers, A., Marsh, D. & Wolff, D. (Eds.). (2007) *Windows on CLIL: Content and Language Integrated Learning in the European Spotlight.* The Hague: European Platform for Dutch Education.

Marsh, D. (Ed.) (2002) *CLIL / EMILE - The European dimension: Actions, Trends and Foresight Potential.* Jyväskylä: University of Jyväskyla.

Mehisto, P., Marsh, D. & Frigols, M. J. (2008) *Uncovering CLIL: Content and language integrated learning*

in bilingual and multilingual education. Oxford: Macmillan Education.

Oetter, P. (2005) Their mission is transmission: Thema AIDS im bilingualen Biologieunterricht. *Der fremdsprachliche Unterricht Englisch*, 39 (78), 26-29.

Osterhage, S. (2007) Sachfachkönnen (scientific literacy) bilingual und monolingual unterrichteter Biologieschüler: Ein Kompetenzvergleich. In: D. Caspari, W. Hallet, A. Wegner & W. Zydatiß (Hrsg.), *Bilingualer Unterricht macht Schule. Beiträge aus der Praxisforschung* (S. 41–50). Frankfurt am Main: Peter Lang.

Pérez-Cañado, M.L. (2012) CLIL research in Europe: Past, present and future. *International Journal of Bilingual Education and Bilingualism*, 15 (3), 315-341.

Pilot, A., Taconis, R. & den Brok, P. (Eds.) (2016) *Teachers Creating Context-Based Learning Environments in Science*. Rotterdam: Sense Publishers.

Porzucki, N. (2014) *Nobel Prize: How English beat German as language of science*, PRI's The World, BBC News Magazine. 12 October 2014. http://www.bbc.com/news/magazine-29543708 [29.11.2016]

Rasch, B., Friese, M., Hofmann, W. & Naumann, E. (2010*) Quantitative Methoden, Band 2: Einführung in die Statistik für Psychologen und Sozialwissenschaftler* (3. Edition) Heidelberg: Springer.

Richter, R. (2004) And now - in English, please. *Unterricht Biologie* (297/298), 4-12.

Schiefele, U., Krapp, A. & Schreyer, I. (1993) Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogischen Psychologie*, 10 (2), 120-148.

Seidel, T., Prenzel, M., Duit, R. & Lehrke, M. (Eds). (2003). *Technischer Bericht zur Videostudie: „Lehr-Lern-Prozesse im Physikunterricht"*. Kiel: IPN.

Seikkula-Leino, J. (2007) CLIL Learning: Achievement Levels and Affective Factors. *Language and Education*, 21 (4), 328-341.

Stohler, U. (2006) The acquisition of knowledge in bilingual learning: An empirical study on the role of language in content learning. *Vienna English Working Papers*, 15 (3), 41-46.

Stohler, U. (2006) The acquisition of knowledge in bilingual learning: An empirical study on the role of language in content learning. *Vienna English Working Papers*, 15 (3), 41-46.

Tejkalova, L. (2009) *Content and language integrated learning: Strategies and motivation in CLIL*. Saarbrücken: VDM.

Yassin, S.M., Marsh, D., Tek, O.E. & Ying, L.Y. (2009) Learners' Perceptions Towards The Teaching Of Science Through English In Malaysia: A Quantitative Analysis. *International CLIL Research Journal*, 1 (2), 54-69.

Zarobe, Y.R. de & Lasagabaster, D. (2010) CLIL in a Bilingual Community: The Basque Autonomous Community. In D. Lasagabaster Herrarte & Y. de Ruiz Zarobe (Eds.), *CLIL in Spain. Implementation, Results and Teacher Training (S. 12–29)*. Newcastle: Cambridge Scholars.

Motivational effects of structure and autonomy support in non-formal learning

Alexander Eckes[1], Detlef Urhahne[2] and Matthias Wilde[1]

alexander.eckes@uni-bielefeld.de

Abstract

Non-formal settings offer opportunities for self-determined learning. However, these learning settings often lack structure. Teachers accompanying their students in non-formal settings tend to use restrictive and controlling behaviour. Provisions of structure and autonomy-supportive instructions are regarded as beneficial for supporting intrinsic motivation. The aim of our study was to investigate the motivational effects of two types of teacher behaviour: autonomy-supportive versus controlling, and of two levels of structure: extensive versus basic. The sample consisted of 114 students (age: 11.37±0.59 years) from German schools of higher types of tracking. The students visited an exhibition at the university for four hours. We measured intrinsic motivation in a post-test using an adapted version of the Intrinsic Motivation Inventory with the subscales *interest / enjoyment*, *pressure / tension*, *perceived choice* and *perceived competence*. In accordance with the theory, results showed significant main effects of teacher behaviour on *perceived choice*. For structure, the significant main effect of *perceived choice* was contrary to the theoretical expectations. However, significant interaction effects for all subscales were found. The combination of controlling teacher behaviour and extensive structure seemed to be least supportive for students' intrinsic motivation.

Keywords: autonomy support, motivation, non-formal learning environment, Self-Determination Theory

---

[1] Bielefeld University, Germany

[2] Passau University, Germany

1. Introduction

Visits to non-formal settings, such as museums, can offer multifaceted learning environments and foster exploration. At the same time, these field trips are often affected by the "novel field trip phenomenon" (Falk et al., 1978). Students perceiving high levels of novelty may be disoriented and concentrate on non-relevant aspects of their surroundings (Falk & Balling, 1982). Furthermore, non-formal settings may be unstructured and provide an overabundance of choice. This could introduce demands on students that might lead to disorientation and missed learning opportunities (Falk & Balling, 1982; Gottfried, 1980; Griffin & Symington, 1997). In this context, it is hardly surprising that learning achievements during visits to non-formal settings reach a medium level at best (Koran et al., 1989). When investigating learning in non-formal settings, research proves ambiguous (Borun & Flexer, 1983; Griffin, 2004; Lewalter & Geyer, 2005). Rennie & McClafferty (1995) argued that this mixed picture arises from the visits' framework, the embeddedness of the visit in school lessons, the presentation of the visit and, in particular, the novelty of the setting. Falk & Dierking's (2000; 2016) Contextual Model of Learning (CMoL) describes museum learning, i.e. in museums, zoos, interactive exhibitions etc., as a holistic experience concerning the personal, sociocultural and physical contexts. Falk & Dierking (2000) attribute a number of components to each of the contexts. The personal context reflects a person's expectations, motivation and interests as well as their previous knowledge. The sociocultural context deals with mediation within visitor groups, and between guides and visitors. The physical context is concerned with advance organizers, orientation aides and the overall design of an exhibition (Falk & Storksdieck, 2005), as well as the preparation prior to a visit (Falk & Dierking, 2000; 2016).

Self-Determination Theory (SDT) explains how wellbeing, intrinsic motivation and motivated learning can arise in human beings (Deci & Ryan, 2002). Examining non-formal settings from a SDT-perspective might provide further insights into the learning process. In SDT the basic needs for autonomy, competence and relatedness are regarded as requirements for positive qualities of motivation to arise (Deci & Ryan, 2002). Fulfillment of the basic needs enhances students' intrinsic motivation and their sense of well-being, and leads to more engagement (Jang et al., 2009; Niemiec et al., 2010). According to cognitive evaluation theory, one of the sub-theories of SDT, autonomy

and competence are the most important prerequisites for intrinsic qualities of motivation to arise (Deci & Ryan, 1980). Autonomy support plays a key role in facilitating task-persistence when no external support is present (Connell & Wellborn, 1991), and can be fostered by autonomy-supportive instructions (Skinner et al., 2008). Autonomy support is characterized by non-controlling language, and giving students room to make decisions concerning their actions, while acknowledging students' task-related thoughts and perspectives. According to SDT, there are two opposing interpersonal styles, autonomy support and control, which lead to self-determined or to non-self-determined motivation, respectively (Vansteenkiste et al., 2012). The basic need for competence represents the innate tendency to seek new and challenging experiences, and to expand, discover and improve personal skills (Ryan & Deci, 2000). The need for competence is described as interacting meaningfully with one's surroundings and achieving highly valued goals (Deci et al., 1991; Ryan & Deci, 2002). This means that opportunities must be provided to practice personal skills and abilities, and to experience physical and psychological confirmation (Ryan & Deci, 2002).

When an out of school visit is viewed from a SDT perspective, teacher behaviour and the structure of the non-formal setting should be the particular focus. Concerning teacher behaviour, both interpersonal styles described by SDT are found in non-formal settings. According to Falk & Dierking (2016), educators must be mindful of the needs and goals of the visitors that coincide with their personal context, and which may be, in essence, similar to SDT's autonomy support. On the other hand, teachers visiting non-formal settings are often tense and afraid of losing students, or worry about being unable to answer students' questions (Dillon et al., 2006; Griffin & Symington, 1997). This often results in restrictive behaviour (Griffin & Symington, 1997), which students could perceive as controlling.

On the other hand, the structure of the non-formal setting and the resulting perceived competence of visitors are of significance. During their first encounter in a non-formal setting, learners might experience disorientation and insecurity, and the new location may be perceived as distracting (Falk & Balling, 1982; Gottfried, 1980; Griffin & Symington, 1997). Providing structure could help learners regulate their own learning process and regain feelings of competence. Structure can be viewed as a two-dimensional continuum, extending from a high degree of structure to a complete absence of structure

(Jang et al., 2010), the latter of which is referred to as chaos (Jang et al., 2010). Structure can be provided in various ways, e.g., in the form of worksheets (DeWitt & Storksdieck, 2008) and pre-visit preparations (Griffin, 1998). Another form of structure is spatial orientation, which provides students with a conceptual and spatial preview of a new location, e.g., a museum (Hiss, 2000). These provisions of structure may enable learners to maintain their focus and choose specific tasks or challenges that are well-suited to their level of ability (Griffin, 1998; Ryan & Deci, 2002). Thus, structure might support students' perception of competence (Dupont et al., 2014; Skinner, 1995; Skinner et al., 2008; Vansteenkiste et al., 2012).

2. Research question

In SDT and in non-formal settings the needs for autonomy and competence are of importance. Autonomy support can fulfil the need for autonomy (Skinner et al., 2008), and provision of structure can fulfil the need for competence (Dupont et al., 2014; Vansteenkiste et al., 2012). This is possible through developing perceived control and an internal locus of control, which reduces the sense of helplessness (Skinner & Belmont, 1993; Skinner et al., 2008). Hospel & Galand (2016) state that autonomy support and structure could be seen as complementary and mutually supportive dimensions (Jang et al., 2010; Sierens et al., 2009; Vansteenkiste et al., 2012). Combining the perspectives of CMoL and SDT, this study varied teacher behaviour (autonomy-supportive versus controlling) and the amount of structure given (extensive versus basic) in our exhibition, impacting the three contexts as well as the basic needs for autonomy and competence. Teacher behaviour presumably influences the components of the sociocultural context, as it directly affects the relationship between teacher and learner. It might be concerned only indirectly with the personal context, as we expect it to affect the learners' motivation. Structure, on the other hand, directly influences the components of the physical context and, presumably, indirectly influences the personal context as well. The given amount of structure might have an impact on learners' motivation, too. Consequently, our research question addresses the following main effects and their interactions: *How does teacher behaviour during the field trip, the structure of the exhibition, and the interaction between teacher behaviour and structure influence students' intrinsic motivation?*

3. Methods

The sample consisted of 114 (55% female) students aged 11.37±0.59 years from the highest type of tracking. All students visited an exhibition located in a German university for four hours, and attended a participatory exhibition thematically centred on the locomotor systems of humans and animals. Students worked in groups of three to four children. The exhibition was spread across three rooms, and each room offered up to six workstations. The three teachers in the exhibition ('guides') were biology students of advanced semesters.

The autonomy-supportive and controlling teacher behaviours were explained and clarified in great detail, and teachers were trained to behave according to theory, as either autonomy-supportive "A" or controlling "C". The autonomy-supportive teaching style (A) was characterized by specific verbal cues. Teacher-student interactions were governed by certain rules: instructions started with "can" or "may", and the imperative was not used. Teachers interrupted activities only when requested or if necessary due to security reasons. Students were encouraged to explore and work independently. Workstations were not assigned, and no time constraints were given. The controlling teaching style (C) was characterized by teachers actively moving about the room and occasionally looking over students' shoulders. Their communication was characterized by the use of imperatives and instructions using "must", "should", and "have to". Students were assigned to workstations according to a mandatory plan. Teachers periodically asked students what they were doing, or reminded them of time constraints and the importance of their work in relation to a written examination. In this study, we focused on a particular element of teachers' instructional style. Teachers were introduced to SDT, the use of the above-mentioned verbal cues and ways to use these behaviours effectively in practice. When students were asked whether teachers let them decide for themselves how to deal with tasks and workstations, and let them distribute their time freely, or whether they felt pressured, they only showed high perceived autonomy when facilitated by autonomy-supportive teachers. As three teachers were active during each school visit, the perceived autonomy was averaged. The internal consistency was Cronbach's $\alpha = .72$.

Regarding structure, students received either a basic level of structure "S", or an extensive amount of structure "S+". Independent of the treatment, the learning materials were identical regarding content. Basic structure (S) consist-

ed of a brief explanation of the working materials. Students in this treatment used basic worksheets with closed and open-ended questions regarding the exhibition. The extensive structure (S+) comprised all elements of the basic structure and contained additional elements. These additional elements were pre-visit preparations, where teachers provided a preview of the day and its contents, as well as clear expectations for students' work, structure-supportive instructions, and explanations of learning materials and their usage via examples taken directly from the worksheets (e.g. Falk & Dierking, 2000). Exhibits were marked by organizers and labels which corresponded to labels in the structured worksheets, operating as orientation aides. Worksheets used the same closed and open-ended questions. The rooms were each marked with an individual animal symbol (mascot) that was found on the corresponding worksheet, and was also used to guide students along pathways between the rooms. At the beginning of each room, students received minimal guidance that lasted three to four minutes to get acquainted with the exhibits. Both treatments were combined, resulting in four treatment conditions: AS and AS+, CS and CS+.

Being interested in the quality of motivation that arises through these treatments, we used an adapted version of the *Intrinsic Motivation Inventory* (IMI) (Ryan, 1982) with a five-point rating scale, ranging from 0 (*not at all true*) to 4 (*very true*) in the post-test. The assessed subscales were *interest/enjoyment*, *pressure/tension*, *perceived choice* and *perceived competence*. The *interest/enjoyment* subscale is considered the self-reporting measure of intrinsic motivation. The subscales *perceived choice* and *perceived competence* are positive predictors of intrinsic motivation, and *pressure/tension* is a negative predictor of intrinsic motivation. The internal consistencies of the subscales are shown in Table 1. All values were acceptable.

Table 1 Subscales of the adapted *Intrinsic Motivation Inventory*, reported as Cronbach's Alpha, number of items and example items.

Subscale	Cronbach's Alpha	Items	Example item
Interest/enjoyment	$\alpha = .81$	7	I enjoyed doing this activity very much.
Pressure/tension	$\alpha = .65$	5	I felt very tense while doing this activity.
Perceived choice	$\alpha = .75$	5	I did this activity because I wanted to.
Perceived competence	$\alpha = .76$	6	I think I am pretty good at this activity.

4. Results

The results of the impact of teacher behaviour and structure on students' intrinsic motivation are shown in Table 2. The perceived autonomy shows that when students received autonomy-supportive teacher behaviour, they perceived themselves as much more autonomy-supported. Teacher behaviour had an effect on intrinsic motivation that is generally consistent with theory. On the subscale *perceived choice*, we found a moderate positive main effect from autonomy-supportive teacher behaviour. Other subscales did not contribute to this overall picture. The effect of provision of structure was not consistent with theory. The only significant main effect was found in the subscale *perceived choice*. Most interesting were the interactions that were significant for each subscale. In the subscales *interest / enjoyment*, *pressure / tension* and *perceived choice* the effects were small to medium; in the subscale *perceived competence* the effect was medium to large. The results show that the predictors of intrinsic motivation were lowest in the treatment CS+. Students in this condition felt the least *interest / enjoyment*, *perceived choice*, the least *perceived competence*, and the highest *pressure / tension*.

Table 2 Means (M) and standard deviations (SD) for the perceived autonomy, as well as the subscales of the adapted intrinsic motivation inventory in the post-test, are shown for all combinations of autonomy-supportive (A) and controlling teacher behaviour (C) with different levels of structure, i.e., basic structure (S) and extensive structure (S+). Reported is a MANOVA with main effects of teacher behaviour and structure and the interactions of teacher behaviour x Structure in all subscales.

Subscale		S	S+	Teacher behaviour	Structure	Interaction
		M (±SD)	M (±SD)			
Perceived autonomy	A	2.96 (±0.62)	2.92 (±0.60)	$F(1,108) = 118,30$, $p < 0.001$, $\eta^2 = .523$	$F(1,108) = 0.85$, $p = $ n.s.	$F(1,108) = 0.43$, $p = $ n.s.
	C	1.78 (±0.49)	1.59 (±0.71)			
Interest/ enjoyment	A	2.42 (±0.78)	2.60 (±0.95)	$F(1,107) = 0.05$, $p = $ n.s.	$F(1,107) = 0.77$, $p = $ n.s.	$F(1,107) = 3.97$, $p < 0.05$, $\eta^2 = .036$
	C	2.70 (±0.70)	2.24 (±0.91)			
Pressure/ tension	A	1.39 (±0.84)	1.23 (±0.96)	$F(1,105) = 3.30$, $p = $ n.s.	$F(1,105) = 0.68$, $p = $ n.s.	$F(1,105) = 2.91$, $p < 0.05$, $\eta^2 = .027$
	C	1.41 (±1.04)	1.87 (±0.89)			
Perceived choice	A	2.33 (±0.76)	2.28 (±0.75)	$F(1,107) = 5.31$, $p < 0.05$, $\eta^2 = .047$	$F(1,107) = 4.33$, $p < 0.05$, $\eta^2 = .039$	$F(1, 107) = 2.86$, $p = 0.09$, $\eta^2 = .026$
	C	2.24 (±0.77)	1.68 (±0.87)			
Perceived competence	A	2.56 (±0.67)	2.79 (±0.72)	$F(1, 105) = 0.31$, $p = $ n.s.	$F(1,105) = 1.60$, $p = $ n.s.	$F(1, 105) = 8.13$, $p < 0.01$, $\eta^2 = .072$
	C	2.90 (±0.70)	2.29 (±0.93)			

5. Discussion and conclusion

The aim of our study was to examine the impact of teacher behaviour and structure on students' intrinsic motivation. In a non-formal setting, autonomy-supportive and controlling teacher behaviour, and basic and extensive structure were combined to construct a learning environment according to the CMoL (Falk & Dierking, 2000; 2016) and SDT (Deci & Ryan, 2002) theories. The study evaluated the quality of students' motivation while visiting an exhibition. The two factors, teacher behaviour and structure, were each varied in two levels. According to the theoretical assumptions, effects were expected as follows: autonomy-supportive teacher behaviour has beneficial effects on intrinsic motivation through the support of the basic need for autonomy, and (extensive) structure is beneficial for intrinsic motivation through the support of the basic need for competence. The analysis showed that the main effects for teacher behaviour and structure found cannot be regarded as major evidences for or against the theoretical assumptions, yet significant interactions suggest that both factors may be complementary and mutually supportive dimensions.

Regarding the main effect of autonomy-supportive teacher behaviour, this theory is supported by the significant effect of autonomy-supportive teacher behaviour for the subscale *perceived choice*. Stronger evidence was found in previous studies (Basten et al., 2014; Eckes & Wilde, 2016). The effect size is nevertheless near a medium effect range. Students of higher types of tracking may react differently to both teaching styles, either because they are used to stricter teaching styles or are already used to autonomy-supporting styles. A possibly beneficial effect of the treatment structure for students of higher types of tracking can only be examined by analysing the interactions of both factors.

The results suggest that autonomy support through teacher behaviour and competence support through structure did not have separate effects on students' motivation. As described, the effects of both factors overlap in reference to the three contexts of CMoL. Teacher behaviour presumably influences the components of the sociocultural and the personal context. Structure, on the other hand, influences the components of the physical and presumably the personal context. As we are assessing students' motivation situated in the personal context, an interaction between both factors is very reasonable.

As the significant interactions of all subscales suggest, the combination of both factors is most interesting. Perceptions of teacher behaviour and structure were dependent on one another. Two very different combinations seem to be effective for students' motivation. The first one was autonomy-supportive teacher behaviour combined with extensive structure (AS+). This combination achieved exactly what SDT states: Autonomy-supportive teacher behaviour fosters students' perceived autonomy (e.g. Grolnick & Ryan, 1987). Extensive structure might fulfil the students' need for competence (Connell & Wellborn, 1991). Both perceived autonomy and perceived competence facilitate intrinsic motivation (Deci & Ryan, 2002). Regarding the non-formal setting, the combination of autonomy support and extensive structure is likely to support students in their interaction with the effects of the settings' novelty, and may prevent disorientation or distraction.

The second combination that resulted in good qualities of motivation was unexpected. It was basic structure and controlling teacher behaviour (CS). In terms of the non-formal setting, the combination of control and basic structure should have led to distracted students that work on non-relevant aspects of the exhibition or try to orientate themselves. Furthermore, students' motivation was expected to be low. It is possible that the high achievers of our sample had a very good idea of how to learn well. They may have already known that a certain level of performance was expected of them. This implicit feeling of pressure might lead to very little tolerance of ambiguity which is enhanced, understandably, when less structure is given. Furthermore, students of young ages are dependent on external feedback and extrinsic reinforcement (Harter, 1981), which is what students often experience in school. Presumably, these students did not mind being told what to do in a controlling way, when they did not know what they were supposed to do otherwise. If these controlling instructions from the teachers gave them the external feedback and extrinsic reinforcement they needed to feel competent, it is no surprise that motivation was not impaired as long as the students felt competent (Harter & Jackson, 1992).

But what happens when students already know what to do and receive constant further instruction? Here, we need to focus another, less effective treatment combination: controlling teacher behaviour combined with extensive structure (CS+). On all subscales, this overload of instructions led to the

worst results regarding intrinsic motivation. When students already know what to do, due to a structured learning environment where they feel they understand everything, they do not appreciate being told repeatedly (and unnecessarily) what to do and how to do it again and again (Jang et al., 2010).

A limitation to this study is the characterization of the teachers' instructional style as it focused only on a specific part, yet it is this part in particular that has a tremendous impact on the perceived autonomy. Teacher behaviour as a whole is built from many different characteristics and beliefs, some of which are hard to change (Reeve & Cheon, 2016). For the focus of the motivating style, factors such as personality disposition, culture or beliefs are responsible for why teachers orient themselves towards autonomy support or control (Reeve et al., 2014; Taylor et al., 2009). Aeltermann, et al. (2016) state that proposed strategies have to be explained, their effectiveness shown and the necessary tools provided to put them into practice. To enable teachers in our study to recognize the operationalization, i.e., autonomy support and control as feasible teaching styles, teachers were introduced to SDT and trained in applying both behaviours throughout multiple training sessions. In relation to autonomy-supportive and controlling teacher behaviour, studies show that teacher-training works for teachers with pre-existing controlling, neutral, and autonomy-supportive styles (Cheon & Reeve, 2013; Cheon et al., 2012; Tessier et al., 2010). We cannot rule out the influences of teachers' characteristics as a whole on their autonomy-supportive and controlling behaviour. To minimize these effects teachers were instructed in each of the treatments, CS, CS+, AS and AS+. Additionally, students' perceived autonomy showed that teachers' teaching style was perceived as either autonomy-supportive or controlling.

In essence, for this high-achieving sample we may need to construct a learning environment that refers to all contexts of the Contextual Model of Learning, which are the sociocultural context (learner-teacher relationship), the physical context (structure) and the personal context (perception of autonomy and motivation), but the learning environment must be adjusted to a beneficial equilibrium. Controlling teacher behaviour and extensive structure could provide an excessive amount of instruction that undermines students' intrinsic motivation.

References

Aelterman, N., Vansteenkiste, M., Van Keer, H. & Haerens, L. (2016) Changing teachers' beliefs regarding autonomy support and structure: The role of experienced psychological need satisfaction in teacher training. *Psychology of Sport and Exercise, 23*, 64–72.

Baard, P.P., Deci, E.L. & Ryan, R.M. (2004) Instrinsic need satisfaction: A motivational basis of performance and well-being in two work settings. *Journal of Applied Social Psychology, 34*, 2045-2068.

Basten, M., Meyer-Ahrens, I., Fries, S. & Wilde, M. (2014) The Effects of Autonomy-supportive vs. Controlling Guidance on Learners' Motivational and Cognitive Achievement in a Structured Field Trip. *Science Education, 98*(6), 1033–1053.

Borun, M. & Flexer, B.K. (1983) Plants and pulleys: Studies of class visits to science museum exhibits. *Curator, 26*, 201–219.

Cheon, S.H. & Reeve, J. (2013) Do the benefits from autonomy-supportive training program endure? A one-year follow-up investigation. *Psychology of Sport and Exercise, 14*, 508–518.

Cheon, S.H., Reeve, J. & Moon, I.S. (2012) Experimentally based, longitudinally designed, teacher-focused intervention to help physical education teachers be more autonomy supportive toward their students. *Journal of Sport & Exercise Psychology, 34*, 365–396.

Connell, J.P. & Wellborn, J. G. (1991) Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In J. P. Connell, J. G. Wellborn, M. R. Gunnar & L. A. Sroufe (Eds.), *Self-processes and development*. Hillsdale, NJ: Lawrence Erlbaum.

Deci, E.L. & Ryan, R.M. (1980) The empirical exploration of intrinsic motivational processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol 13, pp. 39–80). New York, NY: Academic Press.

Deci, E.L. & Ryan, R.M. (2002) Self-determination research: Reflections and future directions. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 431–441). Rochester: University of Rochester.

Deci, E.L., Vallerand, R.J., Pelletier, L.G. & Ryan, R.M. (1991) Motivation and Education: The Self-Determination Perspective. *Educational Psychologist 26*(3&4), 325-346.

DeWitt, J. & Storksdieck, M. (2008) A short review of school field trips: Key findings from the past and implications for the future. *Visitor Studies, 11*(2), 181–197.

Dillon, J., Rickinson, M., Teamey, K., Morris, M., Choi, M. Y., Sanders, D. & Benefield, P. (2006) The value of outdoor learning: evidence from research in the UK and elsewhere. *School Science Review, 87*(320), 107-111.

Dupont, S., Galand, B., Nils, F. & Hospel, V. (2014) Social Context, Self-perceptions and Student Engagement: A SEM investigation of the self-system model of motivational development (SSM-MD). *Electronic Journal of Research in Educational Psychology, 12*, 5-32.

Eckes, A. & Wilde, M. (2016) The influence of structure and autonomy supportive teachers' behaviour on intrinsic motivation. *Electronic Proceedings of the 11th ESERA 2015 Conference. Science education research: Engaging learners for a sustainable future.* (p. 322–331). Helsinki, Finnland: University of Helsinki.

Falk, J.H. & Balling, J.D. (1982) The field trip milieu: Learning and behavior as a function of contextual events. *Journal of Educational Research, 76*(1), 22–28.

Falk, J.H. & Dierking, L.D. (2000) *Learning from museums: Visitor experience and the making of meaning*. Walnut Creek, CA: AltaMira Press.

Falk, J.H. & Dierking, L.D. (2016) *The museum experience revisited*. NY: Routledge.

Falk, J. & Storksdieck, M. (2005) Using the contextual model of learning to understand visitor learning from a science center exhibition. *Science Education, 89*(5), 744–778.

Falk, J.H., Martin W.W. & Balling, J.D. (1978) Environmental effects on learning: The outdoor field trip. *Science Education, 65*(3), 301–309.

Gottfried, J.L. (1980) Do children learn on school field trips? *Curator, 23*, 165–174.

Griffin, J. (1998) Learning science through practical experiences in museums. *International Journal of Science Education, 20*(6), 655–663.

Griffin, J. (2004) Research on students and museums: Looking more closely at the students in school groups. *Science Education, 88*, 59–70.

Griffin, J. & Symington, D. (1997) Moving from task-oriented to learning-oriented strategies on school excursions to museums. *Science Education, 81*(6), 763–779.

Grolnick, W.S. & Ryan, R.M. (1987) Autonomy in Children's learning: An Experimental and Individual Difference Investigation. *Journal of Personality and Social Psychology, 52*(5), 890–898.

Harter, S. (1981) A New Self-Report Scale of Intrinsic Versus Extrinsic Orientation in the Classroom: Motivational and Informational Components. *Developmental Psychology, 17*(3), 300–312.

Harter, S. & Jackson, B.K. (1992) Trait vs. Nontrait Conceptualizations of Intrinsic/Extrinsic Motivational Orientation. *Motivation and Emotion, 16*(3), 209–230.

Hiss, T. (2000) A Place for Learning. In J.H. Falk and L. D. Dierking (Eds.). *Learning from Museums: Visitor Experiences and the Making of Meaning*. Walnut Creek, CA: AltaMira Press.

Hospel, V. & Galand, B. (2016) Are both classroom autonomy support and structure equally important for students' engagement? A multilevel analysis. *Learning and Instruction, 41*, 1–10.

Jang, H., Reeve, J. & Deci, E.L. (2010) Engaging Students in Learning Activities: It Is Not Autonomy Support or Structure but Autonomy Support and Structure. *Journal of Educational Psychology, 102*(3), 588–600.

Jang, H., Reeve, J., Ryan, R.M. & Kim, A. (2009) Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically oriented Korean students? *Journal of Educational Psychology, 101*(3), 644–661.

Koran, J.J., Koran, M.L. & Ellis, J. (1989) Evaluating the effectiveness of field experiences: 1939-1989. *Visitor Behavior IV*, H. 2, 7–10.

Lewalter, D. & Geyer, C. (2005) Evaluation von Schulklassenbesuchen um Museum. *Zeitschrift für Pädagogik, 51*(6), 774–785.

Niemiec, C.P., Ryan, R.M. & Deci, E.L. (2010) Self-determination theory and the relation of autonomy to self-regulatory processes and personality development. In R. H. Hoyle (Ed.), *Handbook of personality and self-regulation* (pp. 169–191). Malden, MA: Blackwell Publishing Ltd.

Reeve, J. & Cheon, S.H. (2016) Teachers become more autonomy supportive after they believe it is easy to do. *Psychology of Sport and Exercise, 22*, 178–189.

Reeve, J., Vansteenkiste, M., Assor, A., Ahmad, I., Cheon, S.H., Jang, H. & Wang, C.J. (2014) The beliefs that underlie autonomy-supportive and controlling teaching: A multinational investigation. *Motivation and Emotion, 38*(1), 93–110.

Rennie, L.J. & McClafferty, T. (1995) Using visits to interactive science and technology centers, museums, aquaria, and zoos to promote learning in science. *Journal of Science Teacher Education* 6, 175–185.

Ryan, R.M. (1982) Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*, 450–461.

Ryan, R.M. & Deci, E.L. (2000) Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist 55*(1), 68–78.

Ryan, R.M. & Deci, E.L. (2002) Overview of self-determination theory: An organismic dialectical perspective. In E.L. Deci & R.M. Ryan (Eds.), *Handbook of self-determination research*. Rochester: University of Rochester.

Sierens, E., Vansteenkiste, M., Goossens, L., Soenens, B. & Dochy, F. (2009) The synergistic relationship of perceived autonomy support and structure in the prediction of self-regulated learning. *British Journal of Educational Psychology, 79*(1), 57–68.

Skinner, E.A. (1995) *Perceived control, motivation, & coping* (Vol. 8). Thousand Oaks, CA: Sage.

Skinner, E.A. & Belmont, M.J. (1993) Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85*, 571–581.

Skinner, E., Furrer, C., Marchand, G. & Kindermann, T. (2008) Engagement and Disaffection in the Classroom: Part of a Larger Motivational Dynamic? *Journal of Educational Psychology, 100*(4), 765–781.

Taylor, I., Ntoumanis, N. & Smith, B. (2009) The social context as a determinant of teacher motivational strategies in physical education. *Psychology of Sport and Exercise, 19*, 235–243.

Tessier, D., Sarrazin, P. & Ntoumanis, N. (2010) The effect of an intervention to improve newly qualified teachers' interpersonal style, students' motivation and psychological need satisfaction in sport-based physical education. *Contemporary Educational Psychology, 35*, 242–253.

Vansteenkiste, M., Sierens, E., Goossens, L., Soenens, B., Dochy, F., Mouratidis, A. & Beyers, W. (2012) Identifying configurations of perceived teacher autonomy support and structure: Associations with self-regulated learning, motivation and problem behavior. *Learning and Instruction, 22*(6), 431–439.

Intrinsic motivation in bilingual courses on molecular biology and bionics in an out-of-school laboratory

Annika Rodenhauser[1] and Angelika Preisfeld[2]

rodenhauser@uni-bremen.de

Abstract

Taking into account German students' deficiencies in scientific literacy as well as reading competence and the 'mother tongue plus two foreign languages' objective of the European Commission, bilingual courses on molecular biology (*Genetic Fingerprinting*, $n = 490$) and bionics (*A Glue from Snail Slime?!*, $n = 120$) were developed. Affective and cognitive factors were measured in a pre, post, follow-up test design, with intrinsic motivation being assessed by the *Kurzskala intrinsischer Motivation* (KIM) (Wilde et al., 2009). This study focuses on questions concerning the presence of intrinsic motivation under conditions where bilingual learning is combined with phases of practical experimentation. The intent was to investigate: i) the presence of intrinsic motivation in general, ii) the causes for the occurrence of pertinent aspects of intrinsic motivation (experiment-related or foreign language-related), and iii) possible influences of students' general interests and preferences relating to intrinsic motivation. Bilingual courses in an out-of-school laboratory can trigger intrinsic motivation in general, and all-rounders seem to profit from the bilingual courses most. Additionally, it has become evident that the perceived choice seems to be the component of intrinsic motivation showing the greatest correlation with content-related biological achievement.

Keywords: bilingual education, experimentation, intrinsic motivation, molecular biology, out-of-school lab

---

[1] University of Bremen, Germany

[2] University of Wuppertal, Germany

1. Introduction

Today's knowledge based society is getting increasingly globalized and demands scientific and foreign language competences from its members to be able to actively participate in social life. In comparison to these demands, shortcomings of these essential competences can be observed on different levels. On the one hand, practical science is scarcely taught in Germany (Euler, 2004). Thus, it is not surprising that PISA studies have assessed great deficiencies in scientific literacy of German students (Klieme et al., 2010). These conditions in schools are also believed to be responsible for students' general lack of interest in science and consequently for an absence of graduates. Different studies on out-of-school laboratories have proven positive influences of (monolingual) courses on students' attitudes towards science (e.g. Damerau, 2013; Engeln, 2004; Glowinski, 2007; Pawek, 2009), which raises hopes that out-of-school labs could counteract the lack of interest. On the other hand, the aim of the European Union's language policies (European Commission, 2012) is that every European citizen should be able to speak at least two languages in addition to their mother tongue.

As a means for reaching this goal, the commission recommends Content and Language Integrated Learning (CLIL) as an exposure to a foreign language that does not require extra time in the curriculum (European Commission, 2004). Taking into account these conditions, bilingual (German-English) courses on molecular biology and bionics for an out-of-school lab, combining practical experimentation with CLIL were developed and conducted. Findings concerning the acquisition of biological content knowledge and the development of biological self-concept have already been published (Rodenhauser, 2016; Rodenhauser & Preisfeld, 2015; 2016). It has been shown that cognitive achievement concerning biological content knowledge of students having participated in a bilingual course (English and German) does not differ significantly from the cognitive achievement of those that have participated in a monolingual course (German). Regarding affective aspects, it was found that bilingual courses have the potential to encourage scientifically oriented students in terms of foreign languages, and foreign language oriented students concerning science. But it is unclear whether the combination of scientific and foreign language aspects might have an impact on students' intrinsic motivation, and whether the positive influences that have been revealed for

monolingual courses also apply to bilingual courses.

Thus, this study aims to find out whether bilingual courses in out-of-school labs also have the potential to enhance students' intrinsic motivation and could consequently contribute to the reduction of students' general lack of interest in science. The central research question was the following: *Do students experience task-based intrinsic motivation through participation in the courses? And, if so, is this task-based intrinsic motivation triggered by the scientific or the foreign language aspects of the courses?*

1.1 Laboratory courses

The educational concept combining practical experimentation and CLIL fundamentals has been implemented in two courses in an out-of-school laboratory on a university campus. In the laboratory course *Genetic Fingerprinting* (Rodenhauser & Preisfeld, 2015; Damerau, 2013), students have to find a (virtual) criminal by using the technique of genetic fingerprinting. Practically, their task is to extract DNA from oral mucosa cells, perform PCR and gel electrophoresis afterwards and compare the bands of the amplified genes to identify the offender in the end. The general course procedure is contextualized in that students are confronted with the (virtual) criminal case. Following the concept of knowledge-based constructivism (Linn, 1990; Resnick & Hall, 1998), the criminal case, background information and methodological instructions are presented in a short pre-laboratory phase in which students are encouraged to reflect on and discuss laboratory methods needed to be able to convict the offender (Hodson, 1998; Lunetta, 1998). In this way, students are provided with knowledge with which they are able to construct their findings made during the experimental phase. Each student group performs a genetic fingerprint with one of three DNA samples (two from the suspects and one from the crime scene). First, the sample is prepared for the replication of the locus D1S80 (Budowle et al. 1991; Kasai et al. 1990) by PCR. During the PCR process, students extract their own DNA from oral mucosa cells to be able to reconstruct the extraction process. As soon as the amplified DNA sequences (D1S80) are available, a gel electrophoresis is carried out by the students. Afterwards, DNA bands in the gel can be observed under UV light. As it cannot be assumed that students develop an understanding about the course contents and methods simply by doing the experiments (Abd-El-Khalick et al., 2004), students' results and problems or questions they may have encountered are

discussed in a reflection phase at the end of the course.

In the laboratory course *A Glue from Snail Slime?!,* which addresses students from lower secondary level, students test the adhesive properties of snail slime in comparison to those of a standard glue stick (Buttemer, 2009; Rodenhauser & Preisfeld, 2016). This is done within the context of Bionics and Scientific Working. The course procedure is identical to that of the *Genetic Fingerprinting* course insofar as it starts with a short pre-laboratory phase and continues with an experimental and a reflection phase. The contextualisation consists of a presentation of an US-American scientist and his research project dealing with the construction of a 'robosnail' that moves on an artificial snail slime. In this context, students are invited to participate in this research project, which also constitutes the reason for needing to be able to present research results in English.

1.2 Self-Determination Theory

Self-Determination Theory (Deci & Ryan, 1993) is a motivational theory dealing with the occurrence of intrinsic motivation based on self-determination. It is assumed that three *basic needs* have to be fulfilled for an action to be perceived as self-determined. These are the needs for competence, autonomy and social relatedness. The need for competence is characterized by a persons' desire to meet given requirements. The desire for autonomy refers to the ambition to experience freedom. Satisfaction of the social relatedness need is perceived if a person feels connected to and accepted by persons in their social environment. As the concepts of intrinsic motivation, interest and Self-Determination Theory are strongly linked together (Krapp 1998), the occurrence or non-occurrence of a task-based intrinsic motivation can also hint at possibilities for the promotion of interests. Apart from that, the intention of the study was to gain an impression of correlations between different aspects of intrinsic motivation.

2. Research design and methodology

The quasi-experimental study was conducted with two experimental groups and one control group. One of them took part in a bilingual (English-German) course on molecular biology ($n = 198$) and the other group attended a monolingual (German) course ($n = 224$) with exactly the same content and

procedure. The control group ($n = 68$) did not take part in any of the courses. Data were collected in a pre, post, follow-up test design (course *A Glue from Snail Slime?!*: only bilingual experimental group; pre, post-test design, $n = 120$) with intrinsic motivation assessed using the *Kurzskala intrinsischer Motivation* (KIM) (Wilde et al., 2009) at the time of the post test. Students' average age was 16 *(M = 16.5; SD = 2.05)* and 58.5 % of the students were female. The subscales of the KIM are *interest/enjoyment* ($\alpha_{Bio} = 0.84$; $\alpha_{Engl.} = 0.93$), *perceived competence* ($\alpha = 0.8$), *perceived choice* ($\alpha = 0.79$) and *pressure/tension* ($\alpha_{Bio} = 0.63$; $\alpha_{Engl.} = 0.72$). To find out if a possible intrinsic motivation is caused by the biological or the foreign language components, the subscales *interest/enjoyment* and *pressure/tension* were formulated for both aspects (e.g. "The *experiments* were fun. / The *use of the foreign language* was fun.").

3. Findings

Analyses were conducted taking into account all bilingual courses to find out about the occurrence of intrinsic motivation in general and about the aspect of the educational concept that may have caused it.
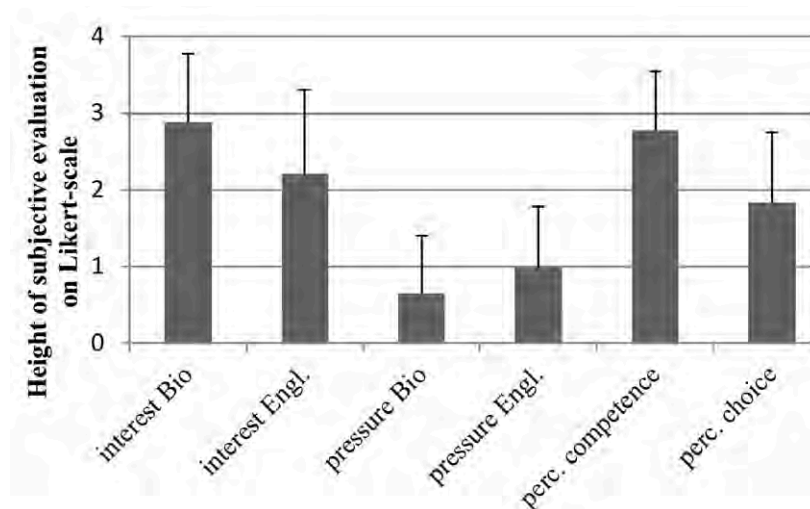
Figure 1 Results of the subjective assessment of the subscales of the KIM by participants of all bilingual courses (scale maximum = 4).

Figure 1 shows that the subscales *interest/enjoyment*$_{Bio}$ and *interest/enjoyment*$_{Engl.}$ are assessed as relatively high, whereas the interest in using the English

language is assessed as smaller than the one in the biological components. The values of the *perceived competence* scale indicate that course participants had the feeling of being up to the tasks of the courses. The subscales *pressure/ tension* that function as negative predictors of intrinsic motivation are assessed as relatively low, which indicates that the need for autonomy is met satisfactorily. Besides, the pressure caused by the use of the foreign language it is rated higher than the one caused by the biological components. The *perceived choice* was given an average evaluation.

Pearson correlations (taking into account all German and bilingual courses) revealed highly significant correlations between *perceived choice* and *interest/ enjoyment*$_{Bio}$ ($r = 0.62$; $p \leq 0.001$) as well as between *perceived competence* and *interest/enjoyment*$_{Bio}$ ($r = 0.48$; $p \leq 0.001$). Additionally, there is a highly significant correlation between *pressure/tension*$_{Bio}$ and *pressure/tension*$_{Engl.}$ *($r = 0.35$; $p \leq 0.001$)*. This indicates that the perceived choice constitutes the essential factor of the task-based intrinsic motivation in the courses. The correlation between the biological and foreign language pressure-triggering component suggests that students obviously do not differentiate between these. They seem to feel pressurized or not, regardless of the components exerting this pressure.

Table 1 Pearson-correlations (r), significances (2-sided) and sample size (N) of the KIM subscales (intrinsic motivation assessed during the post-test) for the course *Genetic Fingerprinting* (bilingual and monolingual courses). Correlations $r \geq |0.25|$ are displayed in light grey, correlations $r \geq |0.4|$ are displayed in dark grey.

		interest/en-joyment$_{Bio}$	interest/en-joyment$_{Engl.}$	pressure/tension$_{Bio}$	pressure/tension$_{Engl.}$	perceived competence	perceived choice
interest/en-joyment$_{Bio}$	r signif. N		0,28 $p \leq 0,001$ 198	-0,13 $p = 0,06$ 198	0,00 $p = 0,99$ 196	0,40 $p \leq 0,001$ 198	0,25 $p \leq 0,001$ 198
interest/en-joyment$_{Engl.}$	r signif. N	0,28 $p \leq 0,001$ 198		0,03 $p = 0,67$ 198	-0,36 $p \leq 0,001$ 196	0,12 $p = 0,08$ 198	0,02 $p = 0,77$ 198
pressure/tension$_{Bio}$	r signif. N	-0,13 $p = 0,06$ 198	0,03 $p = 0,67$ 198		0,27 $p \leq 0,001$ 196	-0,23 $p \leq 0,001$ 198	0,07 $p = 0,36$ 198
pressure/tension$_{Engl.}$	r signif. N	0,00 $p = 0,99$ 196	-0,36 $p \leq 0,001$ 196	0,27 $p \leq 0,001$ 196		-0,18 $p \leq 0,05$ 196	0,13 $p = 0,07$ 196
perceived competence	r signif. N	0,40 $p \leq 0,001$ 198	0,12 $p = 0,08$ 198	-0,23 $p \leq 0,001$ 198	-0,18 $p \leq 0,05$ 196		0,17 $p \leq 0,05$ 198
perceived choice	r signif. N	0,25 $p \leq 0,001$ 198	0,02 $p = 0,77$ 198	0,07 $p = 0,36$ 198	0,13 $p = 0,07$ 196	0,17 $p \leq 0,05$ 198	

For the course *Genetic Fingerprinting* (Table 1), the greatest correlation can be found between the subscales *interest/enjoyment*$_{Bio}$ and *perceived competence* ($r = 0.40$; $p \le 0.001$). Moreover, there is a negative correlation between the subscales *interest/enjoyment*$_{Bio}$ and *pressure/tension*$_{Engl.}$ ($r = -0.36$; $p \le 0.001$). This means that pronounced interest and enjoyment concerning foreign languages is associated with a low feeling of pressure concerning foreign language aspects.

Table 1 shows all correlations between the KIM subscales for the *Genetic Fingerprinting* course, and Table 2 shows the correlations for the *A Glue from Snail Slime?!* course. The correlation between *perceived choice* and *interest/enjoyment*$_{Bio}$ is much more distinct for the *A Glue from Snail Slime?!* course ($r = 0.46$; $p \le 0.001$), than for the *Genetic Fingerprinting* course ($r = 0.25$; $p \le 0.001$).

Table 2: Pearson-correlations (r), significances (2-sided) and sample size (N) of the KIM-subscales (intrinsic motivation assessed during the post test) for the course A Glue from Snail Slime?! (bilingual). Correlations r ≥ |0.25| are displayed in light grey, correlations r ≥ |0.4| are displayed in dark grey.

		interest/en-joyment$_{Bio}$	interest/en-joyment$_{Engl.}$	pressure/tension$_{Bio}$	pressure/tension$_{Engl.}$	perceived compet.	perceived choice
interest/en-joyment$_{Bio}$	r		0,38	-0,23	-0,11	0,42	0,46
	signif.		$p \le 0,001$	$p = 0,01$	$p = 0,25$	$p \le 0,001$	$p \le 0,001$
	N		119	118	119	120	118
interest/en-joyment$_{Engl.}$	r	0,38		-0,05	-0,35	0,22	0,18
	signif.	$p \le 0,001$		$p = 0,57$	$p \le 0,001$	$p = 0,02$	$p = 0,05$
	N	119		118	119	119	118
pressure/tension$_{Bio}$	r	-0,23	-0,05		0,55	-0,21	-0,08
	signif.	$p = 0,01$	$p = 0,57$		$p \le 0,001$	$p = 0,02$	$p = 0,38$
	N	118	118		118	118	117
pressure/tension$_{Engl.}$	r	-0,11	-0,35	0,55		-0,10	0,07
	signif.	$p = 0,25$	$p \le 0,001$	$p \le 0,001$		$p = 0,26$	$p = 0,46$
	N	119	119	118		119	118
perceived competence	r	0,42	0,22	-0,21	-0,10		0,36
	signif.	$p \le 0,001$	$p = 0,02$	$p = 0,02$	$p = 0,26$		$p \le 0,001$
	N	120	119	118	119		118
perceived choice	r	0,46	0,18	-0,08	0,07	0,36	
	signif.	$p \le 0,001$	$p = 0,05$	$p = 0,38$	$p = 0,46$	$p \le 0,001$	
	N	118	118	117	118	118	

Moreover, the correlation between *pressure/tension*$_{Engl.}$ and *pressure/tension*$_{Bio}$ was much greater for the course *A Glue from Snail Slime?!* ($r = 0.55$; $p \le 0.001$) than for the course *Genetic Fingerprinting* ($r = 0.27$; $p \le 0.001$). This indicates that younger students differentiate even less between pressure-triggering factors than the older ones. They appear to have faced a more general state of pressure. It was also found that post test scores of the cognitive test (includ-

ing all groups) correlate strongest with the subscale *perceived choice* ($r = -0.28$; $p \leq 0.001$). This means that students feeling more autonomous, score higher in the cognitive test.

To get an insight into students' general characteristics, biological and foreign language self-concept, interest in course topic (genetics/bionics), interest in subject (biology) and general attitudes towards foreign language learning were assessed in all groups and courses in the pre test. The scales were also used to conduct a hierarchical cluster analysis (complete linkage) to identify students uniting similar manifestations of these variables (Figure 2).
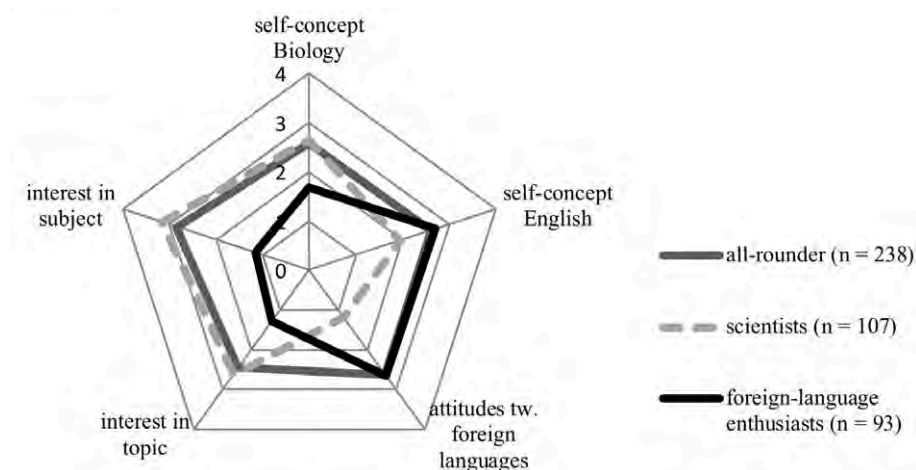
Figure 2 Visualization of means of the items of the affective variables for the groups having emerged from the cluster analysis (scale max.= 4).
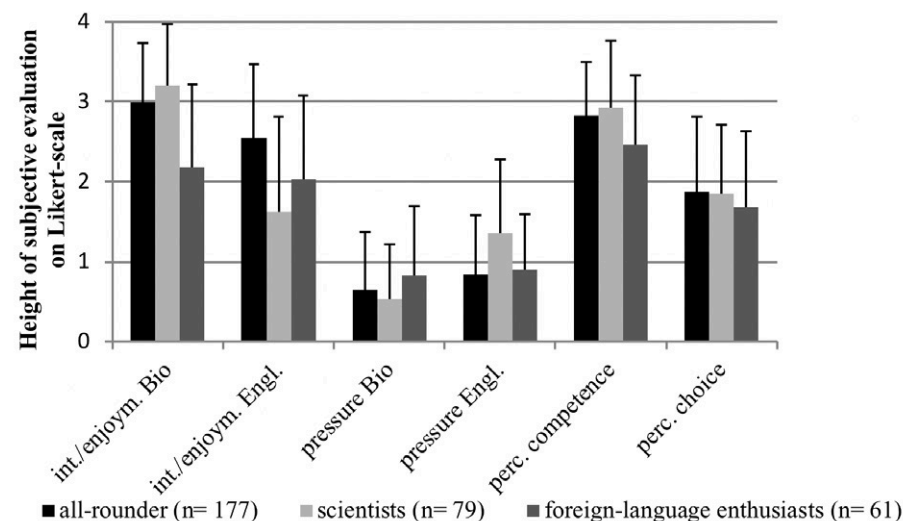
Figure 3 Results of the subjective assessment of the subscales of the KIM by the groups (scale max.= 4).

Three types of students could be identified as *all-rounders*, *scientists* and *foreign language enthusiasts* according to the patterns of manifestation of the different variables (Figure 2). In the next step, ANOVAs were used to find out about potential differences in intrinsic motivation for these groups.

Figure 3 Results of the subjective assessment of the subscales of the KIM by the groups (scale max.= 4).

Highly significant differences were found between members of the three groups for all subscales, except from *pressure*$_{Bio}$ ($F(2.314) = 2.56$; $p = 0.08$) and *perceived choice* ($F(2.314) = 34.87$; $p = 0.33$). To find out about differences between the groups, post-hoc tests were conducted. Table 3 shows that all-rounders and scientists differ in their interest in foreign languages and their feeling of being pressure-triggered by the foreign language aspects. All-rounders and foreign language enthusiasts as well as scientists and foreign language enthusiasts differ concerning their interest in biological course aspects.

Table 3 Results of post-hoc tests between the groups for subscales of the KIM (all participants of the courses Genetic Fingerprinting and A Glue from Snail Slime?!).

Subscale of KIM	all-rounder vs. scientists	all-rounder vs. f-l enthusiasts	scientists vs. f-l enthusiasts
interest$_{Bio}$	$p = 0.169$	$p \leq 0.001$	$p \leq 0.001$
interest$_{Engl.}$	$p \leq 0.001$	$p = 0.03$	$p = 0.06$
pressure$_{Bio}$	$p = 0.534$	$p = 0.298$	$p = 0.081$
pressure$_{Engl.}$	$p \leq 0.001$	$p = 0.913$	$p = 0.002$
perc. competence	$p = 0.614$	$p = 0.007$	$p = 0.002$
perc. choice	$p = 0.977$	$p = 0.359$	$p = 0.558$

The two laboratory courses were conceptualized for different grades and thus students of different ages. Therefore, *t* tests for independent samples were used to examine whether the different courses, and consequently students' ages, had an impact on students' intrinsic motivation. Figure 4 and Table 4 show the differences in intrinsic motivation (KIM subscales) between the courses.
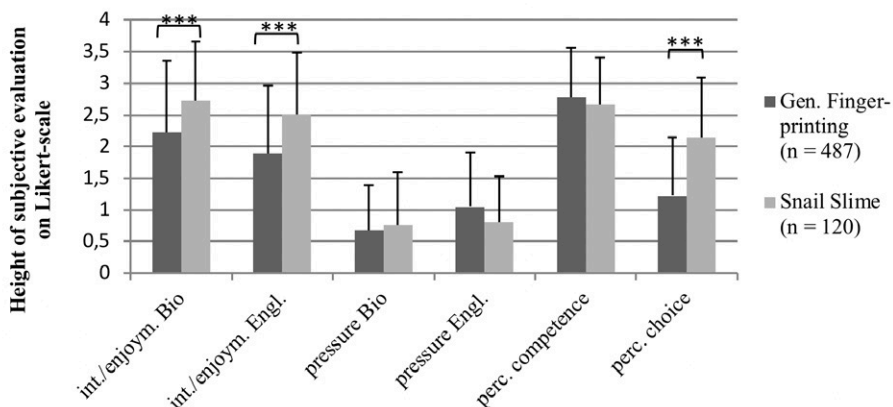
Figure 4 Results of the subjective assessment of the subscales of the KIM by all participants of the courses Genetic Fingerprinting and A Glue from Snail Slime?! (scale maximum = 4; ***: p ≤ 0.001).

The comparison between *interest/enjoyment$_{Bio}$* revealed a highly significant difference for both courses ($t(217.49) = $ -1.15; $p ≤ 0.001$; $w^2 = 0.001$). As can be seen in Figure 4, the younger students (*A Glue from Snail Slime?!*) showed a much higher interest and enjoyment relating to the biological content than the older participants (*Genetic Fingerprinting*).

Table 4 Results of the subjective assessment of the subscales of the KIM by all participants of the courses Genetic Fingerprinting and A Glue from Snail Slime?! with standard deviations (SD) und sample sizes (N).

time of testing	subscale	laboratory course	height of subjective evaluation on Likert scale	standard deviation (SD)	sample size (N)
	interest/enjoyment$_{Bio}$	Gen. Fing.	2.22	1.14	487
		Snails	2.73	0.93	120
	interest/ enjoyment$_{Engl}$	Gen. Fing.	1.90	1.07	282
		Snails	2.51	0.98	120
	pressure/tension$_{Bio}$	Gen. Fing.	0.67	0.72	486
		Snails	0.75	0.85	120
post test	*pressure/tension$_{Engl.}$*	Gen Fing.	1.05	0.86	221
		Snails	0.81	0.72	120
	perceived competence	Gen Fing.	2.78	0.79	316
		Snails	2.66	0.74	120
	perceived choice	Gen. Fing.	1.22	0.93	485
		Snails	2.15	0.94	120

Concerning *interest/enjoyment$_{Engl}$*, the *t* test revealed a similar finding ($t(399) = $ -5.33; $p ≤ 0.001$; $w^2 = 0.04$). In this case, the younger students also enjoyed

the activity significantly more than the older ones. On the basis of the effect size, it can be said that there is a relevant effect for the subscale *interest/enjoyment$_{Engl.}$* ($w^2 = 0.04$). Concerning the subscale *pressure/tension$_{Bio}$*, no significant difference between participants of the courses *Genetic Fingerprinting* and *A Glue from Snail Slime?!* can be found ($t(602) = $ -1.04; $p = 0.29$). Concerning the subscale *pressure/tension$_{Engl.}$*, participants of the course *Genetic Fingerprinting* felt significantly more pressurized than participants of the course *A Glue from Snail Slime?!* ($t(338) = 2.55$; $p = 0.01$; $w^2 = 0.01$). There was no significant difference between the courses for the subscale *perceived competence* $t(434) = 1.37$; $p= 0.17$).

Concerning the subscale *perceived* choice, there was a highly significant difference between the courses ($t (601) = $ -9.72; $p \leq 0.001$, $w^2 = 0.13$), which showed that participants of the course *A Glue from Snail Slime?!* felt considerably more elective (Figure 4 & Table 4). The effect size of $w^2 = 0.13$ hints at a substantial effect. Possible reasons for this will be discussed in the following section.

4. Discussion

Through the application of the scale of intrinsic motivation (KIM) (Wilde et al., 2009), the presence of task-based intrinsic motivation could be observed for both courses. By the use of an adjusted formulation of the items assessing the constructs of *interest/enjoyment* and *pressure/tension* concerning the biological and the foreign language aspects of the courses, an impression of the triggering factors could be gained. Intrinsic motivation was rather caused by the biological aspects of the courses.

For the groups having emerged from the cluster analysis (scientists, foreign language enthusiasts and all-rounders), highly significant differences in intrinsic motivation could be observed. Taking into account all KIM subscales, all-rounders show the highest amount of intrinsic motivation. This meets the expectations insofar that bilingual laboratory courses meet the biological as well as the foreign language preferences of these students. As expected, scientists were most enthusiastic about the biological aspects. Contrary to expectations, all-rounders, and not the foreign language enthusiasts, enjoyed the foreign language aspects of the courses most. Moreover, the foreign language enthusiasts felt the greatest pressure concerning the biological course aspects and the scientists concerning the foreign language aspects. Based on the high

correlations between the subscales assessing the pressure triggered by the biological and the foreign language course aspects, it can be concluded that participants do not seem to differentiate between the triggering factors, but experience a more general feeling of pressure.

Through the comparison of the courses *Genetic Fingerprinting* and *A Glue from Snail Slime?!*, it could be determined that participants of the course *A Glue from Snail Slime?!* had considerably more fun concerning the biological as well as the foreign language aspects of the course. Furthermore, participants of this course indicated to have perceived a significantly greater choice. This can be explained by the structure and procedure of the courses. Because of the topic and the associated procedure, the *Genetic Fingerprinting* course indeed has a less open structure. It was intended to structure the course procedure as open as possible, but because of the course's key goal of conducting a genetic fingerprint, where some experimental stages are compulsory, the openness was necessarily limited. In this context, Engeln (2004) did not find a direct correlation between the openness of a learning environment and other laboratory variables. But in the current study, a great correlation between the subscales *interest/enjoyment* and *perceived choice* was determined ($r = 0.62$). Hence, it can be assumed that the level of openness constitutes a crucial influence factor concerning the perception of the learning environment and thus, also of intrinsic motivation.

On the one hand, these findings support claims of concepts as Discovery Learning (Bruner, 1961) or Inquiry-based Learning (Schwab, 1966), which emphasize the activity of the learner and thus the need for learning environments being as open as possible (Bruner 1961; Steffe & Gale, 1995). On the other hand, it is often argued and even been proven that unguided lessons are not necessarily more effective than lessons focusing on instruction (Hattie, 2013; Kirschner et al., 2006). This ostensive conflict can also be found in Hatties' meta-analysis (2013). He states that Inquiry-based Learning promotes cognitive and affective aspects of learning. However, he also states that active and instructional lessons are more effective than unguided and moderated ones. But how can this conflict be resolved? Based on the results of the current study which are supported by neuroscientific findings on learning, it is assumed that it is not the actual openness of the learning environment, the choice given to students, but the perception of openness by the students that is of relevance

for the learning process. Thus, for one thing, the learning environment needs to be clearly structured, and learners need to have a basic knowledge on the topic that they can refer to during the learning process (Kirschner et al., 2006; Resnick & Hall, 1998).

However, there seem to be strong correlations between the perceived choice and students' interest as well as performance. This leads to the conclusion that the perception of the learning environment constitutes the crucial factor. Learners need to have the impression of being in an open learning environment which gives them opportunities to choose. This is also supported by Self-Determination Theory, which defines the need for autonomy as one of the three basic needs (Deci & Ryan 2000). Thus, the challenge seems not to be letting students solve problems on their own, without any kind of instruction, but to construct a structured and guided learning environment which the learners do not perceive as instructional. For reaching this goal, the concept of knowledge-based constructivism (Linn 1990; Resnick & Hall 1998) as well as the scaffolding method, going back to Vygotsky (1978), seem to be appropriate. Apart from this, these assumptions are supported by findings from neuroscience. The transportation of information from the amygdala to the prefrontal cortex (from the brain area responsible for encoding and information retrieval to the area storing information for a long time) is promoted by the confrontation with manageable challenges and the achievement of meaningful goals (Willis 2010). Such learning situations are also related to a high release of dopamine which is associated with pleasurable experiences. Consequently, when strategies are used which are associated with increasing dopamine release, the brain responds not only with pleasure, but also with increased focus, memory, and motivation (Storm & Tecott, 2005).

In the course of this study it was shown that the didactic concept of the bilingual out-of-school laboratory courses was not only fruitful in terms of knowledge acquisition and positive changes in self-concept (Rodenhauser & Preisfeld, 2015; 2016), but also in terms of intrinsic motivation, and thus as a link to the promotion of interests in science. Despite the structure of the courses, which in some parts was necessarily instructional, participants (especially for *A Glue from Snail Slime?!)* still seemed to feel they had choices. This led to a high intrinsic motivation. Furthermore, it has become evident that *perceived choice* seems to be the component of intrinsic motivation showing the greatest correlation with content-related biological achievement.

References

Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N.G., Mamlok-Naaman, R., & Hofstein, A. (2004) Inquiry in science education: International perspectives. I*nternational Journal of Science Education*, 88(3), 397-419.

Bruner, J. (1961) The Act of Discovery. *Harvard Educational Review,* (31), 21-32.

Budowle, B., Chakraborty, R., Giusti, A.M., Eisenberg, A.J. & Allen, R.C. (1991) Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. A*merican Journal of Human Genetics,* 48 (1), 137-144.

Buttemer, H.A. (2009) A glue from slug slime? *The Science Teacher,* 76 (4), 54-59.

Damerau, K. (2013) *Molekulare und Zell-Biologie im Schülerlabor - Fachliche Optimierung und Evaluation der Wirksamkeit im BeLL Bio (Bergisches Lehr-Lern-Labor Biologie)*. Wuppertal: Universitätsbibliothek Wuppertal.

Deci, E.L. & Ryan, R.M. (1993) Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2/93), 223-239.

Deci, E.L. & Ryan, R.M. (2000) The 'What' and 'Why' of goal pursuits: Human needs and the self-determination of behavior. In: *Psychological Inquiry* 11 (4), 227-268.

Engeln, K. (2004) *Schülerlabors: Authentische, aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken*. Berlin: Logos.

Euler, M. (2004) The role of experiments in the teaching and learning of physics. In: E. F. Redish und M. Vicentini (Eds.) *Research on physics education*. Amsterdam: IOS Press, 175-221.

European Commission (2004) P*romoting language learning and linguistic diversity. An action plan 2004-06*. Luxembourg: Office for Official Publications of the European Communities.

European Commission (2012) F*AQs on multilingualism and language learning. Brussels. Retrieved June 9, 2017*, from http://europa.eu/rapid/press-release_MEMO-12-703_en.pdf

Glowinski, I. (2007) *Schülerlabore im Themenbereich Molekularbiologie als Interesse fördernde Lernumgebungen*. Dissertation. Kiel: CAU Kiel.

Hattie, J. (2013) *Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von 'Visible Learning'*. Unter Mitarbeit von Wolfgang Beywl und Klaus Zierer. Baltmannsweiler: Schneider-Verlag Hohengehren.

Hodson, D. (1998) T*eaching and learning science. Towards a personalized approach* (1. publ). Buckingham, Philadelphia: Open Univ. Press.

Kasai, K., Nakamura, Y. & White, R. (1990) Amplification of a Variable Number of Tandem Repeats (VNTR) Locus (pMCT118) by the Polymerase Chain Reaction (PCR) and its Application to Forensic Science. *Journal of Forensic Sciences,* 35 (5), 1196-1200.

Kirschner, P.A., Sweller, J. & Clark, R.E. (2006) Why Minimal Guidance during Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist,* 41 (2), 75-86.

Klieme, E., Artelt, C., Hartig, J., et al. (2010) *PISA 2009: Bilanz nach einem Jahrzehnt*. Münster: Waxmann.

Krapp, A. (1998) Entwicklung und Förderung von Interessen im Unterricht. (Development and promotion of interest in instruction). *Psychologie in Erziehung und Unterricht*, 45 (3), 185-201.

Linn, M.C. (1990) Summary: Establishing a science and engineering of science education. In: Marjorie Gardner (Ed.) *Toward a scientific practice of science education*. Hillsdale, NJ: Erlbaum, 323-341.

Lunetta, V.N. (1998) The school science laboratory. Historical perspectives and contexts for contemporary teaching. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education. Part One*

(pp. 249–262). Dordrecht, Boston: Kluwer Academic.

Pawek, C. (2009) *Schülerlabore als interessefördernde außerschulische Lernumgebungen für Schülerinnen und Schüler aus der Mittel- und Oberstufe*. Dissertation. Kiel: CAU Kiel.

Resnick, L.B. & Hall, M.W. (1998) Learning organizations for sustainable education reform. *Daedalus*, 127 (4), 89-118.

Rodenhauser, A. & Preisfeld, A. (2015) Bilingual (German–English) Molecular Biology Courses in an Out-of-School Lab on a University Campus: Cognitive and Affective Evaluation. *International Journal of Environmental and Science Education*, 10(1), 99-110.

Rodenhauser, A. & Preisfeld, A. (2016) A Glue from Snail Slime?! – Umsetzung und Evaluation eines bilingualen Moduls für den Biologieunterricht. In B. Diehr, A. Preisfeld, & L. Schmelter (Eds.), *Bilingualen Unterricht weiterentwickeln und erforschen* (pp. 221-243). Frankfurt a.M.: Peter Lang.

Rodenhauser, A. (2016) *Bilinguale biologische Schülerlaborkurse. Konzeption und Durchführung sowie Evaluation der kognitiven und affektiven Wirksamkeit*. Dissertation. Wuppertal: Universitätsbibliothek Wuppertal. (urn:nbn:de:hbz:468-20161108-105143-7)

Schwab, J. (1966) *The Teaching of Science*. Cambridge, MA: Harvard University Press.

Steffe, L.P. & Gale, J. (1995) *Constructivism in education*. Hillsdale, NJ: Lawrence Erlbaum.

Storm, E.E. & Tecott, L.H. (2005) Social circuits: peptidergic regulation of mammalian social behavior. *Neuron*, 47 (4), 483-486. DOI: 10.1016/j.neuron.2005.08.004.

Vygotsky, L.S. (1978) *Mind in society: the development of higher psychological processes*. Cambridge, Mass.: Harvard University Press.

Wilde, M., Bätz, K., Kovaleva, A. & Urhahne, D. (2009) Überprüfung einer Kurzskala intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 31-45.

Willis, J. (2010) Current Impact of Neuroscience in Teaching and Learning. In: David A. Sousa (Ed.): *Mind, Brain, and Education. Neuroscience Implications for the Classroom*. Bloomington, In: Solution Tree Press (Leading edge, 6), 28-45.

The role of interesting and motivating contexts in the assessment of content knowledge and decision-making

Mariella Roesler[1], Nicole Wellnitz[1] and Jürgen Mayer[1]

mariella.roesler@uni-kassel.de

Abstract

Students' interest and motivation differ depending on the contexts in which biological content is embedded. In test situations, interest and motivation have an effect on students' test performance, but the extent to which they have an influence in tests with biological content and contexts is still unclear. The aim of this study is to analyze the influence of interest and motivation on performance in written tests with biological content and contexts. For this purpose, a competency test with 128 items was developed to measure Grade 10 students' ($N = 1543$) competencies in *content knowledge* and *decision-making*. The biological contents (e. g. *enzymology*) were embedded in four contexts (*health*, *environment*, *technology*, *natural resources*). In order to assess interest and motivation, standardized questionnaires ($N_{items} = 116$) were adapted and deployed. The results indicate that there are no differences in average task difficulty across contexts. Nevertheless, students' interest and motivation vary depending on the context and content of a task. The *health* context turned out to be particularly interesting and engaging. Moreover, interest and motivation explain 18% of performance variance. The study concludes that test performance does not always reflect cognitive competence.

Keywords: assessment, competencies, context, interest, motivation

---

[1] Institute for Biology Education, University of Kassel, Germany

1. Introduction

The embedding of content in educationally relevant contexts is becoming more and more important within the framework of school education. However, it has been shown that students' interest and motivation differ depending on the context in which biological content is embedded (Drechsel et al., 2011; Stief et al., 2012). This fact is important for performance measurements because students' test performance is affected by their interest and motivation (Krapp et al., 1992; Wise & DeMars, 2005). The extent to which interest and motivation influence performance on tests with biological content and contexts is still unclear. The aim of this study is to analyze the influence of interest and motivation on performance in written tests with various contexts. These results will allow us to describe the amount of performance variation through cognitive and motivational determinants.

2. Theoretical framework

Within the framework of school education, a context orientation can be observed internationally in the form of context-based approaches. In this way, science competencies in, for example, the *use of content knowledge* and *decision-making* can be encouraged and tested in various contexts so as to enable students to use their knowledge in a variety of life situations. Competencies in the use of content knowledge are an important component of scientific literacy. But in our society, students do not just need (biological) content knowledge, they also need the ability to make decisions about socio-scientific issues, which is another component of scientific literacy (Sadler & Zeidler, 2004). Context-based approaches seek to increase students' interest and motivation, to clarify the relevance of content and to encourage students' ability to transfer subject-specific concepts into a variety of life situations (Bennett et al., 2007; Gilbert, 2006; Wierdsma et al., 2016).

Nevertheless, there is no generally accepted definition of the word *context* (van Oers, 1998). Following the definition of the *Programme for International Student Assessment* (PISA), we see contexts as "life situations that involve science and technology" (Drechsel et al., 2011, p. 77), and are relevant to students' present and future lives (Bennett et al., 2007). A context weaves together content from everyday life and subject-specific content (Haugwitz, 2009). With respect to context, three levels can be distinguished: the idiocultures, the

classroom situation, and the tasks students deal with. On the task level, the task represents a problem and the context forms the storyline of the problem (Finkelstein, 2005). Consequently, a context-based task has two structures: a surface structure (real world) represented by the context, and a deep structure (world of subject-specific models) represented by the content of the task (Löffler & Kauertz, 2014). The content *neurons and synapses* can, for example, be embedded in the context of *drugs*. A resulting context-based task can be about how drugs affect synapses. It turns out that not all contexts are equally interesting and motivating for students (Drechsel et al., 2011; Stief et al., 2012). The context of *diseases and epidemics* is, for example, more interesting for students than the *environment* or *agriculture* contexts (Holstermann & Bögeholz, 2007). The fact that contexts are not equally interesting and motivating is of particular importance for performance measurements that are used to measure competencies, which can be defined as context-specific cognitive dispositions (Klieme et al., 2008). The particular importance of interesting and motivating contexts results from the fact that students' test performance is affected by their interest and motivation (Krapp et al., 1992; Wise & DeMars, 2005).

According to the person-object theory of interest, interest can be described as an interaction between a person and an object. Interest can be further subdivided into individual and situational interest. Individual interest is a specific personal disposition. This interest already exists, and students bring it into a test situation. Situational interest is caused primarily by external factors – for example, in a test situation, by how interesting the test tasks are (Krapp, 2002). Empirical findings from a meta-analysis indicate that the relationship between interest and achievement has been interdisciplinarily confirmed ($r = .30$) (Schiefele et al., 1993). This relationship was also confirmed with regard to biology ($r = .16$) (Schiefele et al., 1993) and to the natural sciences ($r = .26$) in the PISA 2006 science assessment (Schwantner, 2009), where interest in contexts was measured via embedded interest items (Drechsel et al., 2011).

Test motivation describes willingness to engage in working on test items (Baumert & Demmrich, 2001) and can be described with reference to expectancy-value theory. This theory states that a person´s motivation arises from their personal expectations and values: expectations about how successfully a task will be completed as well as the value of the expected result and its consequences (Eccles & Wigfield, 2002). Empirical findings from a meta-analysis indicate that motivation and test performance are positively correlated, with

effect sizes between $r = .23$ and $r = .38$ (Wise & DeMars, 2005). For example, motivation could predict 8 to 12% of test performance in the subject of mathematics (Asseburg, 2011) and 14 to 21% of test performance in an undergraduate psychology class (Sundre & Kitsantas, 2004).

3. Key objectives

Previous empirical findings indicate that interest and the motivational stimulation of contexts have an influence on students' test performance. Previous research findings on the interestingness of contexts stem mainly from interest studies or studies on teaching-learning research referring to context-based teaching models. The role of contexts in performance tests has not yet been sufficiently investigated. Thus, the aim of this study is to analyze the following three questions:

Q1 Are there differences in difficulty between tasks (*content knowledge* & *decision-making*) in different contexts?

Q2 Are there differences in interest and motivation as a result of context and content?

Q3 Do interest and motivation have a significant influence on students' performance and hence competence in tests with biological content and contexts?

4. Research design and method

In order to answer the research questions, 1543 Grade 10 students (47% male) worked on a paper-and-pencil test, which included a competency test with 128 items and questionnaires with 116 items. The competency test was used to measure students' test performance in science (*content knowledge* and *decision-making*). The items were embedded in the four contexts of *health, environment, technology* and *natural resources* (OECD, 2006; Schreiner & Sjøberg, 2004). The items for the competency test were constructed to be equal in number for both competencies and distributed among the four contexts. During the construction of the items, the task characteristics *format*, *cognitive processes* and *complexity* were systematically controlled because we wanted items from all contexts to be equally difficult on average. To control the format of the items, we used both single-choice and extended response items. Eight

items (task) for a given context (four items on *content knowledge* and four items on *decision-making*) were based on a common stem (stimulus text) (Figure 1). Additionally, four items had a competence-specific item stem. Each item consisted of a question or an operationalized work order and either the answer options (single-choice) or placeholders (extended response). We generated approximately the same number of single-choice and extended response items in every context.

Stimulus Text: Dairy Products
No matter whether it is milk, yogurt, cheese or curd, dairy products are part of a proper breakfast. Most people know where milk comes from, but few people know the details of how other dairy products are produced from milk.

Item Stem: Content Knowledge	**Item Stem: Decision-Making**
Untreated milk can be purchased labeled as raw milk. It contains many vitamins and lactic acid bacteria. [...] They make raw milk rapidly sour by fermenting milk sugar (lactose) into lactic acid. The lactose is cleaved by the bacterial enzyme β-galactosidase into the two sugars glucose and galactose. [...] This process is used in the production of cheese made from sour milk [...].	Cheese made from sweet milk is produced with the help of rennet. Rennet is an enzyme mixture. It is obtained from the stomachs of calves, which are still fed with milk. To collect the rennet, the calves must be slaughtered. [...] Due to the high consumption of cheese, microbially generated rennet is used nowadays, too. [...] It is not necessary to specify the origin of the rennet on cheese packaging.
Item 1 of 4 (cognitive process: select)	**Item 1 of 4 (cognitive process: select)**
People who suffer from lactose intolerance cannot digest lactose. Nevertheless, they have no problems eating cheese made from sour milk.	Sarah is vegetarian and has decided to abstain from cheese made from sweet milk.
Explain the effects on lactose during the production of cheese made from sour milk.	Explain what reason could have led to Sarah´s decision.

Questionnaire **Item 1 of 13**	**Questionnaire** **Item 2 of 13**
I think the tasks about *dairy products* are interesting.	I think the content *use of enzymes* is interesting.
I strongly agree ☐ ☐ ☐ ☐ ☐ I strongly disagree	I strongly agree ☐ ☐ ☐ ☐ ☐ I strongly disagree

Figure 1 Item examples for the technology context, task context: dairy products (extended response) and item examples for the questionnaires.

To vary item difficulties and impose different requirements on students, the cognitive processes required to solve the items were varied systematically. The easiest task (*select*) required students to choose given information. The cognitive process *organize* meant that given information had to be structured. In order to solve the hardest tasks (*integrate*), students needed to combine several pieces of given information (Walpuski et al., 2011). We constructed the same number of items requiring students to select, organize and integrate information for each context. The intent was to construct the same number

of easy and difficult items for each context. In addition, we controlled the complexity of the items. We aimed to keep the scope and crosslinking of information that must be processed in the items constant. Each context included at least 32 items: 16 to assess *biological content knowledge* and 16 to assess competencies in *decision-making*. Eight items with the same context formed a task (Figure 1).

To assess the motivational variables interest and motivation, standardized questionnaires were adapted and deployed (Figure 1, Table 1). The interest questionnaire included the two scales *situational interest* and *individual interest*. The motivation questionnaire likewise included two scales: *expectancy* and *value*. Each motivational factor was captured via a Likert scale (5-point or 6-point) with the poles *strongly disagree* to *strongly agree* or *not good at all* to *very good* (Figure 1). The 128 items of the competency test were distributed among 32 test booklets, which were linked with anchor items. Due to the high number of items, a multi-matrix design was chosen. Each booklet contained eight items from two different contexts. Furthermore, six items for *interest*, three items for *expectancy*, and four items for *value* were transferred from the questionnaires into the embedded design after each set of four items for a competence to capture the task-related influence of interest and motivation on test performance (Figure 1, Table 1). Tasks for the competency test were embedded in contexts, and the questions about interest and motivation referred to the embedded tasks. A multi-matrix design was also chosen for the construct *situational interest* because it was captured at several different levels: the context of a task, the task itself and the content of the task. The words in the situational interest items printed in italics (Table 1) differed accordingly. Test motivation was always captured with respect to the task embedded in one of the four contexts. To capture individual interest with regard to the subject of biology, five items were inserted once at the beginning of each test booklet.

Table 1 Scales for the motivational variables (*embedded).

	Scale	Example Item	Level	N	Reference
Interest	**Individual Interest**	The things we learn in biology are interesting to me.	Subject	5	Kölbach, 2011
	Situational Interest*	I think the field technology is interesting.	Context	8	Haugwitz, 2009; van Vorst, 2013
		I think the tasks about dairy products are interesting.	Task	32	
		I think the content use of enzymes is interesting.	Content	64	
Motivation	**Expectancy***	What do you expect, how well did you complete these tasks?	Task	3	Boekaerts, 2002
	Value*	Doing well on these tasks is important to me.	Task	4	Boekaerts, 2002; Sundre, 2007

Each student answered on average 7.68 ($SD = 0.47$) items from the competency test. They had 90 minutes to complete the test and answer the questionnaires. Each item on the competency test was solved by 96.44 ($SD = 12.33$) students on average. Answers on the competency test were analyzed using a coding scheme of either full credit or no credit ($\kappa = 0.81$). Data analysis of the competency test and the questionnaires about interest and motivation are based on the item response theory. The software ConQuest® (Wu et al., 2007) was used to estimate the item parameters and person parameters. Further evaluations were based on classical test theory using the software SPSS and Mplus® (Muthén & Muthén, 2012).

Test quality was analyzed first. A Rasch analysis of the competency test showed that the items are suitable for capturing students' competencies. The competency test showed a good item fit ($0.80 < \text{wMNSQ} < 1.20$; $T \le 1.96$; $r \ge .25$) for 123 of 128 items (96%). Five items did not show a good item fit and were eliminated from the item pool. The person reliabilities were within an acceptable range (content knowledge: $\text{rel}_{\text{EAP/PV}} = .58$; decision-making: $\text{rel}_{\text{EAP/PV}} = .55$) because each student only solved about eight items on average. The item reliability is satisfying with .96 because each item was solved by almost 100 students on average. The results of the Rasch analysis of the questionnaires show that the person reliabilities for the four scales are satisfying ($.83 \le \text{rel}_{\text{EAP/PV}} \le .91$). The item reliabilities are satisfying as well ($.95 \le \text{rel}_{\text{item}} \le .99$). Person and item reliabilities can be interpreted just like Cronbach´s alpha. In the next step, analyses of variance (ANOVA, t-test), correlations and influences (regressions) were conducted in order to investigate the research questions.

5. Findings

5.1 Test performance

To answer the question "Are there differences in difficulty between tasks (*content knowledge* & *decision-making*) in different contexts?" (Q1), average item difficulties were compared across contexts. Figure 2 indicates the average item difficulty of each context. High values indicate difficult items and low values indicate easy items. Analyses of variance (ANOVA) show that the items in the four contexts do not differ in terms of difficulty ($F(3, 119) = 1.59$, $p = .196$). There are no statistically significant differences in average item difficulty across contexts.
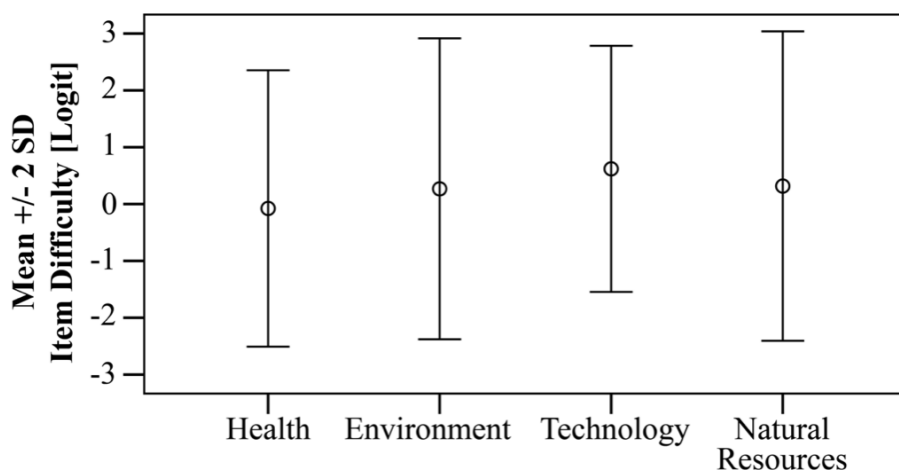
Figure 2 Average item difficulty of each context.

5.2 Differences in interest and motivation

Concerning the question of whether there are differences in interest and motivation as a result of context and content (Q2), the analyses show that students' situational interest ($F(3, 60) = 120.03$, $p < .001$) does vary depending on context. Figure 3 shows the average situational interest in each context, with high values indicating high interest and low values indicating low interest.
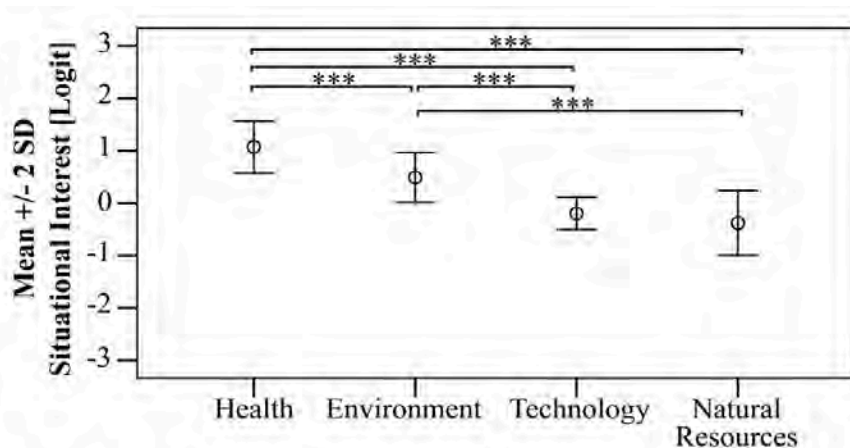
Figure 3 Situational interest in each context (*** p < .001).

The everyday life contexts of *health* and *environment* turned out to be particularly interesting. Students express rather low interest in contexts such as *technology* and *natural resources*.

Regarding interest in the contents (Q2) (Figure 4) embedded in the four contexts, it can be seen that the contexts of *health* and *environment* are significantly more interesting than their embedded contents ($F(7, 50.76) = 64.79$, $p < .001$). In addition, the content *human biology* is more interesting than the other content.
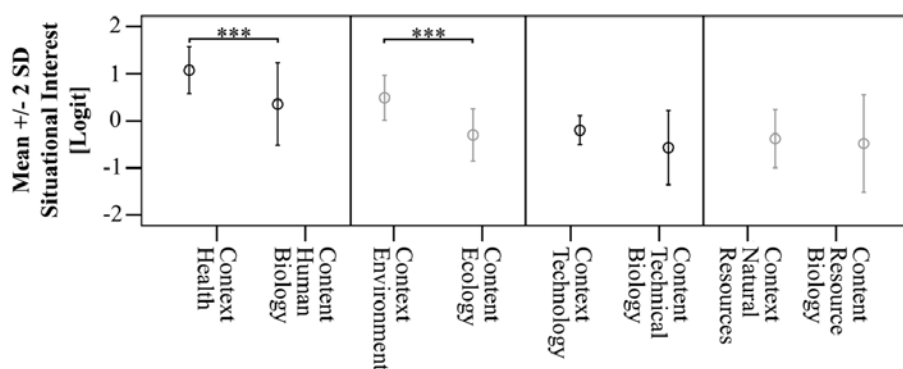
Figure 4 Situational interest in each context and in the embedded contents (*** p < .001).

Moreover, the analyses show that students' motivation varies depending on the context. In Figure 5 (expectancy) and Figure 6 (value), the motiva-

tional stimulation of the contexts is displayed. The expectancy component $(F(3, 92) = 16.21, p < .001)$ and the value component $(F(3, 124) = 17.06, p < .001)$ of motivation differ significantly depending on the context. Figure 6 shows the same sequence of contexts as in Figure 3. The context of *health* is, for example, more motivating than the other contexts.
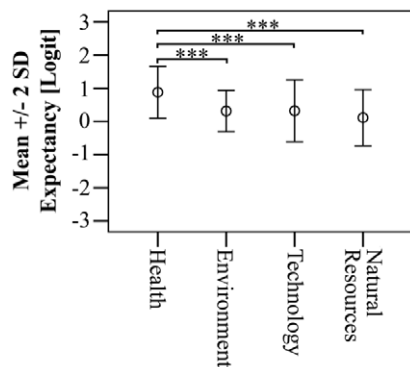
Figure 5 Motivational stimulation of the contexts (Expectancy) (*** p < .001).
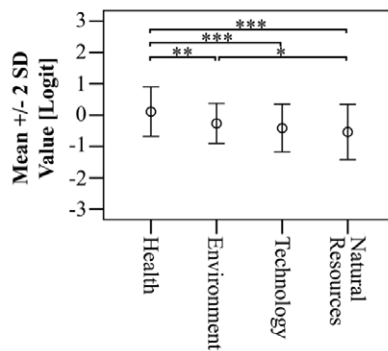
Figure 6 Motivational stimulation of the contexts (Value) (* p < .05; ** p < .01; *** p < .001).

5.3 Influences of interest and motivation

Regarding the question whether interest and motivation have a significant influence on students' performance, and hence competence, in tests with biological content and contexts (Q3), all contexts were summarized to measure the science competence. It turns out that high interest and high motivation are associated with high competence in science (Table 2).

Table 2 Correlations between competence and the motivational variables (*** p < .001).

	Interest		Motivation	
	Individual	Situational	Expectancy	Value
Science Competence	.18***	.28***	.40***	.27***

There are significant correlations between science competence and individual (.18) and situational interest (.28), as well as motivation (expectancy: .40, value: .27). The highest correlation exists between science competence and the expectancy component of motivation (Table 2). The structural equation model (Figure 7) shows the influence of interest and motivation on students'

science competence. All values represent standardized coefficients. The fit indices indicate a good fit of the model. The RMSEA is below 0.05, the CFI value is above the limit of 0.95, and the χ^2-test (χ^2/df) is, with 0.89, clearly below 2.50.

It can be seen that situational interest, expectations about how successfully a task will be completed, and the value of the expected result and its consequences (Eccles & Wigfield, 2002), are positively correlated ($r = .38$, $r = .57$, $r = .63$), and that interest and motivation have a significant influence on competence. Regarding motivation, only expectation of success has an influence ($\beta = .37$), but this influence is higher than the influence of situational interest ($\beta = .14$). Individual interest has no direct influence on competence ($\beta = .05$, $p = .065$) but does have a direct influence on situational interest ($\beta = .39$), expectancy ($\beta = .27$) and value ($\beta = .33$). The indirect influence of individual interest on competence ($\beta = .13$, $p < .001$) is mediated by situational interest and the expectancy component of motivation. Individual and situational interest and the expectancy component of motivation together explain 18% of the variance in the students' science competence.
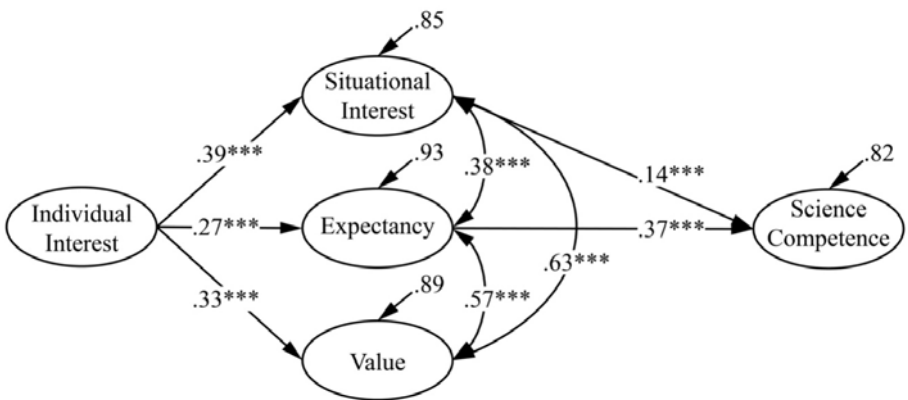
($\chi2 = 2.67$, df = 3, RMSEA (CI 90%) = .00 (.00-.04), CFI = 1)

Figure 7 Influence of interest and motivation on science competence (*** $p < .001$).

6. Conclusions and implications

The results of this study indicate that context has no influence on the difficulty of correctly completing a task (Q1) (see Werner et al., 2014). All contexts were equally difficult on average. This result was expected because the same

number of easy and difficult items were constructed for each context. For this reason, the influences of the motivational variables on test performance and hence science competence could be clearly analyzed.

Consistent with other studies (Drechsel et al., 2011; Holstermann & Bögeholz, 2007), students' interest differed depending on the context and content of a task. In accordance with Gräber (1995), we have shown that contexts from everyday life are particularly interesting. The *health* context is more interesting for students than the *environment* context (see Elster, 2007). There is less interest in the *natural resources* context (see Holstermann & Bögeholz, 2007). It can be assumed that interest in a given context affects interest in that context's embedded content. The content *human biology* is, for example, embedded in the most interesting context of *health* and is more interesting than the other contents. In line with Stief et al.'s (2012) findings for the subject of chemistry, this study has shown that students' motivation also differs across various biological contexts (Q2). Although the items from all contexts were equally difficult on average, they were not equally interesting and motivating. This means that in this study, interest and motivation were not affected by the difficulty of tasks. Rather, the motivational stimulation of test tasks in performance measurements is caused by the context and should be systematically controlled to increase the validity of (cognitive) performance measurements.

Moreover, students' competence was associated with interest and motivation (Q3), as expected (see Schiefele et al., 1993; Schwantner, 2009). Interest and motivation had a significant influence on students' competence, in this case science competence. This influence was based on the situational motivational factors (situational interest, expectancy, value). The personal motivational factor (individual interest) had no direct influence. The situational-motivational variables *situational interest* and *expectancy* together explained 18% of performance variance. The expectancy component of *motivation* had the greatest influence on science competence ($\beta_{Expectancy} = .37$). This result is comparable with results from the subject of mathematics, where the expectancy component of motivation had nearly the same and also the greatest influence among the motivational variables ($\beta_{Expectancy} = .29 - .36$) (Asseburg, 2011). Consistent with a study on the subject of mathematics (Asseburg, 2011) and a study of an undergraduate psychology class (Sundre & Kitsantas, 2004), about 80% of the test result can be explained through cognitive variables. A person's

competence as well as fluid and crystalline intelligence can probably explain the majority of the 80% of performance variance (see Kampa, 2012). Future studies need to examine whether the results are the same when different contexts are chosen.

The results of this study demonstrate that test results do not reflect cognitive competence when interest and motivation are low because test performance is influenced by interest and motivation. This study clarifies the motivational effect on subject-specific test performance and describes performance differences among students in the subject of biology in a more differentiated way because it includes cognitive and motivational determinants. Moreover, the results of this study have implications for future research and for teaching practice. With regard to future research, conclusions can be drawn for the design of competence tests as well as carrying out competence measurements. The influence of contexts can be taken into account in task creation for future measurements by systematically controlling the contexts. This means that contexts should be chosen systematically, not arbitrarily, when designing performance tests. This would help avoid between-subjects distortions, especially in comparative competence measurements. In addition, the results of this study suggest an embedded assessment of motivational variables for future competence measurements. These aspects would allow the evaluation of cognitive competencies to be improved. For purposes of comparability, interdisciplinary performance measurements should be analyzed to determine which proportion of performance differences are based on differences in cognitive factors and which proportion depend on motivational differences between subjects. With respect to school, the findings can be used in specific instructional models and the development of learning environments. Findings about the interestingness of different contexts can be used to increase students' interest in tests or teaching situations. Thus, even less interesting topics can be made more stimulating by embedding them in a context that is interesting for students (Gräber, 1995; Holstermann & Bögeholz, 2007).

As situational interest has a higher influence on competence than individual interest, the interest of the students should be awakened in every topic in school. The situational interest which influences the competence is strongly correlated with the value component of motivation. For this reason, the value of a topic or, regarding tasks, the value of the result and its consequences

should be clarified to students. By increasing the students' interest and motivation, their performance can also be improved.

Acknowledgements

References

Asseburg, R. (2011) *Leistungsbereitschaft in Testsituationen: Motivation zur Bearbeitung adaptiver und nicht-adaptiver Leistungstests* [*The willingness to work in test situations: motivation to work on adaptive and non-adaptive performance tests*]. Wissenschaftliche Beiträge aus dem Tectum Verlag Reihe Psychologie: Vol. 24. Marburg: Tectum.

Baumert, J. & Demmrich, A. (2001) Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education*, 16(3), 441–462.

Bennett, J. Lubben, F., & Hogarth, S. (2007) Bringing Science to Life: A Synthesis of the Research Evidence on the Effects of Context-Based and STS Approaches to Science Teaching. *Science Education*, 91(3), 347–370.

Boekaerts, M. (2002) The On-Line Motivation Questionnaire: A self-report instrument to assess students' context sensitivity. *New Directions in Measures and Methods*, 12, 77–120.

Drechsel, B., Carstensen, C.H. & Prenzel, M. (2011) The Role of Content and Context in PISA Interest Scales – A study of the embedded Interest Items in the PISA 2006 Science Assessment. *International Journal of Science Education*, 33(1), 73–95.

Eccles, J.S. & Wigfield, A. (2002) Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132.

Elster, D. (2007) Student interests - the German and Austrian ROSE survey. *Journal of Biological Education*, 42(1), 5–11.

Finkelstein, N. (2005) Learning Physics in Context: A study of student learning about electricity and magnetism. *International Journal of Science Education*, 27(10), 1187–1209.

Gilbert, J.K. (2006) On the nature of "context" in chemical education. *International Journal of Science Education*, 28(9), 957–976.

Gräber, W. (1995) Schülerinteressen und deren Berücksichtigung im STS-Unterricht: Ergebnisse einer empirischen Studie zum Chemieunterricht [Student interests and their consideration in STS teaching: Results of an empirical study on chemistry teaching]. *Empirische Pädagogik*, 9, 221–238.

Haugwitz, M. (2009) *Kontextorientiertes Lernen und Concept Mapping im Fach Biologie: Eine experimentelle Untersuchung zum Einfluss auf Interesse und Leistung unter Berücksichtigung von Moderationseffekten individueller Voraussetzungen beim kooperativen Lernen* [*Context-oriented learning and concept mapping in biology: An experimental investigation of the influence on interest and achievement considering moderation*

effect of individual preconditions during cooperative learning]. Dissertation: Universität Duisburg-Essen.

Holstermann, N. & Bögeholz, S. (2007) Interesse von Jungen und Mädchen an naturwissenschaftlichen Themen am Ende der Sekundarstufe I [Interest of boys and girls in scientific topics at the end of secondary education]. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 71–86.

Kampa, N. (2012) *Aspekte der Validierung eines Tests zur Kompetenz in Biologie: Eine Studie zur Kompetenz in Biologie und ihren Teildimensionen Konzept- und Prozesswissen* [*Aspects of the validation of a test for competence in biology: A study on competence in biology and its partial dimensions conceptual and process knowledge*]. Dissertation: Humboldt Universität zu Berlin.

Klieme, E., Hartig, J. & Rauch, D. (2008) The Concept of Competence in Educational Contexts. In J. Hartig & D. Klieme E.: Leutner (Eds.), *Assessment of Competencies in Educational Contexts* (pp. 3–22). Göttingen: Hogrefe.

Kölbach, E. (2011) *Kontexteinflüsse beim Lernen mit Lösungsbeispielen* [*Influence of contexts on learning with example solutions*]. Berlin: Logos.

Krapp, A. (2002) Structural and dynamic aspects of interest development: theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, 12, 383–409.

Krapp, A., Schiefele, U. & Winteler, A. (1992) Interest as a predictor of academic achievement: a meta-analysis of research. In K. A. Renninger & Hidi, S. & Krapp, A. (Eds.), *The role of interest in learning and development* (pp. 183–196). Hillsdale, NJ: Erlbaum.

Löffler, P. & Kauertz, A. (2014) Applying physics models in context-based tasks in physics education. In C. Constantinou, N. Papadouris, & A. Hadjigeorgiou (Eds.), *E-Book Proceedings of the ESERA 2013 Conference*. Science Education Research For Evidence-based Teaching and Coherence in Learning (pp. 171–179). Retrieved from www.esera.org/media/documents/ ESERA_eBook_2013.pdf

Muthén, L.K. & Muthén, B.O. (2012) *Mplus: Statistical analysis with latent variables: User's guide*. Los Angeles: Muthén & Muthén.

OECD. (2006) *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006.* Paris: OECD Publishing.

Sadler, T.D., & Zeidler, D.L. (2004) The Morality of Socioscientific Issues: Construal and Resolution of Genetic Engineering Dilemmas. *Science Education*, 88(1), 4–27.

Schiefele, U., Krapp, A. & Schreyer, I. (1993) Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung [Meta-analysis of the correlation between interest and achievement in school]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25(2), 120–148.

Schreiner, C. & Sjøberg, S. (2004) *Sowing the Seeds of ROSE: Background, Rationale, Questionnaire Development and Data Collection for ROSE (The Relevance of Science Education) - A Comparative Study of Students' Views of Science and Science Education*. Retrieved from roseproject.no/key-documents/ key-docs/ad0404-sowing-rose.pdf

Schwantner, U. (2009) Die Motivation der Jugendlichen in Naturwissenschaft [The motivation of young people in science]. In C. Schreiner & U. Schwantner (Eds.), PISA 2006. Österreichischer *Expertenbericht zum Naturwissenschafts- Schwerpunkt* (pp. 266–282). Graz: Leykam.

Stief, K., Thillmann, H. & Sumfleth, E. (2012) Homework motivation and achievement in chemistry education. In C. Bruguière, A. Tiberghien, & P. Clément (Eds.), *Proceedings of the ESERA 2011 Conference*. Science learning and Citizenship (pp. 118–122).

Sundre, D. L. (2007) *The Student Opinion Scale (SOS): A measure of examinee motivation. Test Manual. Harrisonbourg (VA)*: The Center for Assessment & Research Studies.

Sundre, D. L. & Kitsantas, A. (2004) An exploration of the psychology of the examinee: Can examinee self-regulation and test-taking motivation predict consequential and non- consequential test performance? *Contemporary Educational Psychology*, 29, 6–26.

van Oers, B. (1998) From context to contextualizing. *Learning and Instruction*, 8(6), 473–488.

van Vorst, H. (2013) *Kontextmerkmale und ihr Einfluss auf das Schülerinteresse im Fach Chemie* [*Characteristics of contexts and their influence on students' interest in chemistry*]. Berlin: Logos.

Walpuski, M., Ropohl, M. & Sumfleth, E. (2011) Students' knowledge about chemical reactions - development and analysis of standard-based test items. *Chemistry Education Research and Practice*, 12, 174–183.

Werner, M., Schwanewedel, J. & Mayer, J. (2014) Does the context make a difference? Students' ability in decision-making and the influence of contexts. In C. Constantinou, N. Papadouris, & A. Hadjigeorgiou (Eds.), E-*Book Proceedings of the ESERA 2013 Conference*. Science Education Research For Evidence-based Teaching and Coherence in Learning (pp. 1350–1358). Nicosia, Cyprus. Retrieved from www.esera.org/media/documents/ESERA_ eBook_2013.pdf

Wierdsma, M., Boersma, K.T., Knippels, M.-C. & van Oers, B. (2016) Developing the ability to recontextualise cellular respiration: an explorative study in recontextualising biological concepts. *International Journal of Science Education*, 38(15), 2388-2413.

Wise, S.L. & DeMars, C.E. (2005) Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, 10, 1–17.

Wu, M.L., Adams, R.J., Wilson, M.R. & Haldane, S.A. (2007) *Acer ConQuest. Version 2.0. Generalised Item Response Modelling Software*. Camberwell: Australian Council for Educational Research.

Chapter 3:
Scientific thinking, nature of science and argumentation

The contribution of epistemological beliefs to informal reasoning regarding health socio-scientific issues

Andreani Baytelman[1], Kalypso Iordanou[2] and Costas P. Constantinou[1]

baytel@ucy.ac.cy

Abstract

The purpose of this study was to investigate the contribution of pre-service primary teachers' epistemological beliefs to their informal reasoning, while trying to manage health socio-scientific issues (health-SSIs). We used dual-process theories to represent the underlying cognitive process. These theories postulate two distinct processes of reasoning: heuristic/spontaneous processes and analytic/reflective processes. These compete for the control of the response constructed by participants during reasoning tasks. Pre-service primary teachers (N=240) were asked to construct (during spontaneous and analytical thinking) different types of supportive arguments, counterarguments and rebuttals after they had read three different scenarios on health-SSIs. The quantity and the quality of the different types of arguments, constructed during participants' informal reasoning, were employed as indicators for the quality of their informal reasoning. To assess participants' epistemological beliefs, we used a multidimensional perspective. The results indicated that relatively sophisticated epistemological beliefs, especially about the structure of knowledge, positively predicted the number and quality of supportive arguments, counterarguments and rebuttals related to health-SSIs that the pre-service primary teachers constructed during their informal reasoning. We discuss the significance and health educational implications of these findings.

Keywords: epistemological beliefs, health education, health socio-scientific issues, informal reasoning, pre-service primary teachers

1. Introduction

One of the greatest challenges facing educational systems in the coming years is to improve higher-order cognitive skills (Sadler & Zeidler, 2005a). Competence in higher-order cognitive skills such as informal reasoning (IR) is increasingly required for participation in a knowledge society, and for the negotiation and resolution of socio-scientific issues (SSIs) - societal dilemmas with conceptual, procedural, or technological links to science (Bell & Lederman, 2003; Kolstø, 2001; Sadler, 2004). Science education researchers have become interested in examining several factors, such as epistemological beliefs (EBs), values, desires and expectations as potential contributors to IR (Bråten & Strømsø, 2010; Sadler & Zeidler, 2005a; Wu & Tsai, 2011). For example, Wu and Tsai (2011) demonstrated that high school students' EBs relate directly to their IR.

The purpose of the current study was to expand the understanding of the contribution of learners' EBs to their IR by focusing on the relation between pre-service primary teachers' EBs and their IR, while trying to manage health socio-scientific issues (health-SSIs). In this study, we use the term health-SSIs to refer to SSIs in health contexts (Ratcliffe & Grace, 2003). Our rationale for the choice of SSIs in health contexts is the following: The last few decades, SSIs related to, for example, usage of vaccines, vitamins, high voltage lines, diets, treatment of cancer, avian influenza, virus A influenza, cell phones, genetically modified organisms (GMOs), etc., have attracted increasing attention in the developed world. This is because they are related to health, are complex, open-ended issues and because the media offer a huge amount of conflicting scientific and pseudoscientific information on these issues, which require analysis and evaluation for their management and for decision-making. Additionally, health-SSIs are not commonly dealt with in the SSI literature (Lee, 2012). The motivation to focus on pre-service teachers comes from limited research and curriculum development in the area of health education, everyday science, SSIs and teacher development, and the fact that this area can potentially engage students with scientific practices and promote health literacy.

In the present study, EBs are defined as beliefs about the nature of knowledge and the process of knowing (Bråten & Strømsø, 2010; Hofer & Pintrich, 1997; Strømso et al., 2011). IR involves reasoning about causes and consequences and about advantages and disadvantages, or pros and cons, of partic-

ular propositions or decision alternatives. It underlies attitudes and opinions, involves ill-structured problems that have no definite solution, and often involves inductive (rather than deductive) reasoning (Zohar & Nemet, 2002). In this study, we used dual-process theories to represent the underlying cognitive process of IR. These theories postulate two distinct processes of reasoning - heuristic/spontaneous processes and analytic/reflective processes, which compete for the control of the response constructed by participants during reasoning tasks. Researchers argue that SSIs are ideal candidates for the application of IR, because SSIs are complex, open-ended, often contentious dilemmas, with no definitive answers, that can give rise to multiple solutions or arguments (Kuhn, 1993; Sadler, 2004; Sadler & Zeidler, 2005a; Sadler & Zeidler, 2005b).

By carrying out this investigation, we hoped to understand better the contribution of pre-service primary teachers' EBs to their IR, and explore ways to prepare future teachers to improve the quality of their IR regarding health-SSIs, and their ability to foster development of these practices among their future students. Given these aims, we investigated the relationships between pre-service primary teachers' EBs and the quality of their IR during heuristic/spontaneous processes and analytic/reflective processes, in relation to health-SSIs. Specifically, we set out to answer the following two research questions:

1. Is there a relationship between pre-service primary teachers' epistemological beliefs and the quantity of arguments that they construct on health-SSIs during their informal reasoning?
2. Is there a relationship between pre-service primary teachers' epistemological beliefs and the quality of arguments that they construct on health-SSIs during their informal reasoning?

In the current study, quantity and quality of the different types of supportive arguments, counterarguments and rebuttals on health-SSIs, constructed by the pre-service primary teachers during their IR, were employed as indicators for the quality of their IR (Sadler & Zeidler, 2004; 2005a; 2005b). We employed not only supportive arguments, but also counterarguments and rebuttals because according to Kuhn (1992; 1993) the more participants construct counterarguments and rebuttals the higher quality of IR they have.

1.1 Health socio-scientific issues and informal reasoning

The meaning of health has undergone fundamental changes over the last cen-

tury, as marked by a shift from the traditional medical model, which focuses on the physical health of individuals, to the biopsychosocial model which posits that health is influenced by a multitude of biological, psychological, and sociological factors (Lee, 2012). Many health issues can be considered SSIs (Ratcliffe & Grace, 2003), because are open-ended, ill-structured problems, typically contentious and subject to multiple perspectives and solutions (Sadler & Zeidler, 2004; 2005b). Additionally, engaging with SSIs is complex and difficult, involving forming opinions, making decisions at personal or societal level, critical analysis of media reports (from where most of our daily information emanates), evaluation of claims, consideration of values, and ethical and moral reasoning and may require some understanding of probability and risk. Therefore, the negotiation and resolution of SSIs have been generally characterized by the process of informal reasoning, rather than formal reasoning (Sadler, 2004).

According to Sadler & Zeidler (2005a), IR as a construct subsumes the cognitive and affective processes that contribute to the resolution of complex issues, and it is a thinking process that leads to the construction and evaluation of arguments, without relying only on the rules of formal logic, but also on various other factors, such as epistemological beliefs, values, intuition, emotion, desires, expectations, prejudices, feelings, etc. IR as cognitive process assumes importance when information is less accessible, or when the problems are more open-ended, complex, or ill-structured, such as SSIs.

The cognitive mechanism of IR can be explained by the dual process theories (Evans, 2003; 2008), whose principle is based on the existence of two separate cognitive systems: System I and System II. The System I corresponds to spontaneous thinking, while System II is the expression of analytical (algorithmic and reflective) and abstract hypothetical thinking. (Evans, 2003; 2008; Evans & Holmes, 2005). The cognitive Systems I and II may jointly or separately from each one, guide informal reasoning of individuals (Evans, 2003; 2008). System I operates under the existing cognitive structures, prior knowledge, personal and epistemological beliefs, prejudices, feelings, values and experiences of the individual, and related heuristics strategies. The function of System II is directly influenced by the general intelligence and working memory abilities of the individual, their concentration span and ability to prevent the action of System I, the strength of the bias of each individual and the at-

tention time (Evans, 2008; Berrouillet, 2011). Researchers (Evans & Stanovich, 2013; Kahneman, 2011) argue that System 1 and System II processes interact in decision making, and intervention by System II does not necessarily override a System I response. System II checks out the initial intuition with deeper processing. Specifically, Kahneman (2011) argues that fast processing requiring little resource must be combined with another kind of processing that is slow, effortful, and resource intensive. Additionally, the researchers argue that even when intuitive responses are quickly prompted, participants take some time attempting to justify them, making associative connections between the stimuli they encounter and their underlying cognitive structures.

Several researchers argue that IR is expressed through construction of arguments and argumentation and can be investigated and evaluated through argumentation skills (Kuhn, 1992; 1993; Sadler & Zeidler, 2005a; 2005b; Wu & Tsai, 2011; Zeidler, 2004). According to Sadler and Zeidler (2004), "Individuals can express informal reasoning through argumentation (Driver et al., 2000). However, informal reasoning and argumentation represent unique constructs. Informal reasoning refers to the cognitive and affective processes involved in the negotiation of complex issues and the formation or adoption of a position. Argumentation refers to the expression of informal reasoning…."

1.2 Epistemological beliefs and informal reasoning

Epistemological beliefs (EBs) generally refer to individuals' beliefs about the nature of knowledge and the process of knowing (Bråten & Strømsø, 2010; Hofer & Bendixen, 2012; Hofer & Pintrich, 1997). Although multiple theoretical models of EBs have been developed, two overarching groups of models can be identified: those that examine EBs from a developmental perspective, and those that explore EBs from a multidimensional perspective. The primary goal of developmental models (Kuhn et al., 2000) is to explain the development of EBs, proposing that they evolve through stages, whereas the models of EBs that examine those beliefs from a multidimensional perspective focus primarily on the nature and characteristics of beliefs (Schommer, 1990).

According to the multidimensional approach of EBs, researchers have argued that EBs concern multiple dimensions (Hofer & Pintrich, 1997; Schommer, 1990). Schommer described EBs as a system of more or less independent epistemological dimensions, conceptualized as beliefs about the certainty (related with the stability of knowledge), simplicity (related with the structure

of knowledge), and source of knowledge, as well as beliefs about the speed and ability of knowledge acquisition. While the dimensions of certainty, simplicity and source in Schommer's conceptualization fall under the more generally accepted definition of EBs, known as beliefs about the nature of knowledge (certainty, simplicity) and knowing (source) (Hofer & Pintrich, 1997), the dimensions speed and ability have been controversial, because they mainly concern beliefs about learning (speed) and intelligence (ability). Hofer & Pintrich (1997) argued that EBs should be defined more purely, with two dimensions concerning the nature of knowledge (what one believes knowledge is) and two dimensions concerning the nature or process of knowing (how one comes to know). According to Hofer & Pintrich (1997), the two dimensions concerning the nature of knowledge are: (a) Simplicity of knowledge, ranging from the belief that knowledge consists of an accumulation of more or less isolated facts to the belief that knowledge consists of highly interrelated concepts, and (b) Certainty of knowledge, ranging from the belief that knowledge is absolute and unchanging to the belief that knowledge is tentative and evolving. The two dimensions concerning the nature of knowing are: (c) Source of knowledge, ranging from the conception that knowledge originates outside the self and resides in external authority, from which it may be transmitted, to the conception that knowledge is actively constructed by the person in interaction with others, and (d) Justification for knowing, ranging from justification of knowledge claims through observation and authority, or on the basis of what feels right, to the use of rules of inquiry and the evaluation and integration of different sources (Hofer & Pintrich, 1997). Accordingly, Hofer and Pintrich's model differs from Schommer's by omitting the nature of learning factors and adding another nature of knowing factor, justification. Some years later, Conley et al. (2004) suggested a new dimension of EBs, the dimension of Development of knowledge, which is related with the nature of the development of knowledge.

Weinstock & Cronin (2003) investigated empirically the contribution of EBs to IR by involving prospective jurors. Their findings showed that EBs underlie specific juror-reasoning skills and the overall production of arguments. Wu & Tsai (2011) investigated the relationship among 68 high school students' EBs, conceptual understanding regarding nuclear power usage, and their IR regarding this issue, which is expressed through argumentation. Re-

sults indicated that students' beliefs about the justification of scientific knowledge were significantly correlated with the quality of their arguments. In this study, we sought to extend current understandings of the relationship between EBs and IR by examining more closely the possible relationships between pre-service primary teachers' EBs and the number and quality of the different types of supportive arguments, counterarguments and rebuttals constructed during their IR (during heuristic/spontaneous processes and analytic/reflective processes) regarding health-SSIs.

2. Research design and method

2.1 Participants

The participants in this study were 240 pre-service primary teachers (university students of the Faculty of Social Sciences and Education) at a public university in Europe. There were 220 female participants and 20 males, with an overall mean age of 21 (SD = 1,5). Participation in the research was optional and all data were treated anonymously and confidentially.

2.2 Data collection

For this study, we used three different health-SSIs. The first health-SSI dilemma was on usage versus non-usage of vaccines against a new flu virus. The second health-SSI dilemma was on consumption of bottled water versus tap water. The third health-SSI dilemma was on usage of underground versus overhead high voltage lines in residential areas. Our rationale for the choice of these health-SSIs was the following: (a) Over the last decade, health-SSIs such as usage of vaccines, high voltage lines and consumption of bottled versus tap water have attracted increasing attention in the country of the participants, and they were familiar with these SSIs. (b) The participants of this study had already learned about vaccines, drinking water and high voltage lines in their science classes at the secondary school and they had basic prior scientific knowledge regarding these issues. For each health-SSI, we developed a scenario and leaflet with conflicting scientific and not scientific information from different sources such as official websites of relevant government agencies, articles in scientific and professional journals and magazines, articles by laypeople and non-experts, newspaper articles, and teaching resources. The sequence of the information of different positions of the leaflets was reviewed by three

researchers to make sure that the sequence of the information could not affect the results of the study.

We examined EBs from a multidimensional perspective, because we wanted to focus on the nature and characteristics of EBs. More specifically, we used the Dimensions of Epistemological Beliefs toward Science (DEBS) Instrument (Baytelman & Constantinou, 2016), which is based on the multi-dimensional perspective of EBs. The 30-item DEBS Instrument captures three dimensions concerning knowledge (certainty, simplicity and development of knowledge), and two dimensions concerning knowing (source and justification of knowledge). Each dimension consisted of six items and the items were rated on a four-point Likert-scale, ranging from strongly disagree (1) to strongly agree (4). High scores on this measure represent more sophisticated beliefs, while low scores represent less sophisticated beliefs. In the present research, the reliability estimate (Cronbach's a) for the scores on the measure based on the Certainty factor was 0.75, on the Simplicity factor was 0.68, on the Source factor was 0.76, on the Justification factor was 0.76, and on the Development factor was 0.77. The reliability estimates were within the acceptable range for measures developed and used for research purposes (Nunnally, 1978).

The participants' IR processes were assessed by open-ended questionnaires, developed by Wu & Tsai (2011), with some modifications. More specifically, we used three different open-ended questionnaires (one for each health-SSI) requested the construction of different types of supportive arguments, coun-terarguments and rebuttals during informal reasoning (by spontaneous and analytical thinking) (Wu & Tsai, 2011) on each health-SSI scenario. We posed different types of arguments, because in this way we could assess participants' abilities to corroborate information from different sources and to reason about an issue from different perspectives (Barzilai & Eshet-Alkalai, 2015) (See examples of questions of the open-ended questionnaires, and examples of students' arguments in the Appendix).

The variability of arguments was considered in terms of social, economic, ethical, ecological and scientific aspects. Because we were interested in the participants' IR that underlies argumentation, and not argumentation per se, we asked participants to construct specific argument structures, which allowed them to reveal evidence of unarticulated reasoning and implicit justification. In order to measure both, spontaneous and analytical thinking, the condi-

tions required for the operation in each case, as described in the literature, were ensured. Analytical thinking requires adequate time for the evaluation of spontaneous response, clear work instructions and advice for the use of analytical thinking and evaluation of spontaneous response (Barrouillet, 2011), and related information from different sources. These conditions are not required for spontaneous thought (Barrouillet, 2011).

For each participant we computed the quantity and quality of the different types of supportive arguments, counterarguments and rebuttals constructed during spontaneous and analytical thinking. An argument was considered valid, if it involved the presentation of a claim and the legitimacy of that claim was improved through justification. The quality of arguments was determined based on a scoring scheme developed for the purposes of the present study, based on the coding scheme developed by Sadler and Fowler (2006). The quality of arguments was assessed according to the number and accuracy of the pieces of evidence that were provided to support a claim. Thus, arguments that included many pieces of valid and reliable evidence were considered as stronger than arguments with only one piece of evidence (Sadler & Fowler, 2006). The quality of arguments produced was scored 0-4. Inter-rater reliability was conducted by the first author and an independent judge using 30% of the tests of each SSI questionnaire. Reliability was estimated at 94% (Cohens' k = .94), with all disagreements resolved after discussion between the two coders. Following this, the first author coded the remaining tests.

2.3 Procedure

Each participant took part in three sessions. In the first session (20 minutes), the EBs measure was administered. In the second session (20 minutes), participants received the three health-SSI scenarios and the three open-ended questionnaires. After they had read the three different scenarios, were asked to make intuitive/spontaneous judgments regarding the SSI-dilemmas and construct spontaneous different types of supportive arguments, as well as counterarguments and rebuttals on each health-SSI to justify their judgments. All participants completed the tasks in the same order. In the third session, leaflets with conflicting information for each health-SSI scenario were handed out, asking the participants to check their intuitive judgments and spontaneous responses using analytical thinking. After reading the conflicting information, the questionnaires for investigation of analytical thinking were handed out to

the participants. The participants were allowed 40-60 min to complete them.

To answer the research questions of this study, Pearson correlation analyses were carried. Specifically, correlations between scores for each EB dimension (certainty, simplicity, source, justification, development) and scores for the total number and quality of the supportive arguments, counterarguments and rebuttals constructed during IR (spontaneous and analytical thinking together) on the three health-SSIs. Second, multiple regression analyses were carried with the number and the quality of supportive arguments, counterarguments and rebuttals constructed on the health-SSIs of the study as dependent variables. In each analysis, scores for each epistemological dimension were the predictor variables.

3. Results

Table 1 presents descriptive statistics (means, standard deviations, minimum and maximum scores) for all measured variables. Participants' scores on the EBs measures suggested relatively sophisticated beliefs about certainty, justification and development of knowledge, and less sophisticated beliefs about simplicity and source of knowledge. Moreover, data confirmed that participants constructed more and higher quality supportive arguments than counterarguments and rebuttals during their IR (spontaneous and analytical thinking).

These findings indicated that participants' quality of IR was moderate, because according to Kuhn (1992; 1993) the more participants construct counterarguments and rebuttals the higher quality of IR they have.

Table 1 Descriptive statistics for all variables of the study.

Variable	M	SD	Min	Max
Certainty of knowledge	3.01	.37	2	3.88
Simplicity of knowledge	2.64	.36	1.67	3.50
Source of knowledge	2.62	.42	1.50	3.83
Justification of knowledge	3.19	.31	2.43	4.00
Development of knowledge	3.41	.39	2.33	4.00
Number of supportive arguments	7.72	2.29	.00	15.00
Number of counterarguments	6.73	2.38	.00	11.00
Number of rebuttals	6.54	2.43	.00	12.00
Quality of arguments	14.60	6.82	.00	31.00
Quality of counterarguments	12.43	6.21	.00	28.00
Quality of rebuttals	12.17	6.32	.00	31.00

Table 2 displays Pearson correlations among participants' epistemological dimensions and quantity and quality of their supportive arguments, counterarguments and rebuttals constructed during their IR. The Pearson correlations indicated significant positive correlations among simplicity beliefs and number and quality of supportive arguments, counterarguments and rebuttals. Source beliefs were positively and significantly related only to the number of arguments. Certainty beliefs and justification beliefs were positively and significantly related only to the quality of supportive arguments. Development beliefs were not correlated with any of the dependent variables. These findings indicated that participants' simplicity beliefs (beliefs about the structure of knowledge) were positively and significantly related to the quantity and quality of their IR.

Table 2 Pearson correlations among epistemological dimensions and number and quality of arguments.

Variable	1	2	3	4	5
1. Certainty of knowledge	-				
2. Simplicity of knowledge	-.031	-			
3. Source of knowledge	-.064	.271**	-		
4. Justification of knowledge	.275**	.021	.039	-	
5. Development of knowledge	.513**	.043	-.116	.282**	-
6. Number of supportive arguments	.096	.168*	.154*	.072	.036
7. Number of counterarguments	.017	.199**	.142*	.029	-.074
8. Number of rebuttals	.013	.177**	.144*	.078	-.085
9. Quality of supportive arguments	.137*	.205**	.077	.130*	.084
10. Quality of counterarguments	.097	.248**	.093	.083	.042
11. Quality of rebuttals	.054	.233**	.097	.129	.040

Note. *$p < .05$, **$p < .01$, ***$p < .001$.

Tables 3 and 4 present the results of multiple regression analyses for participants' epistemological dimensions predicting the number and quality of supportive arguments, counterarguments and rebuttals constructed during their IR (heuristic/spontaneous processes and analytic/reflective processes).

Table 3 Results of multiple regression analysis for epistemological dimensions predicting participants' number of arguments.

Predictor	Number of supportive arguments		Number of counterarguments		Number of rebuttals	
	B(SE)	β	B(SE)	B	B(SE)	B
Certainty of knowledge	-.04 (.39)	.01	-.09 (.41)	-.01	-.02(.42)	.00
Simplicity of knowledge	.36 (.42)	.05*	.92 (.45)	.14*	.90(.45)	.13*
Source of knowledge	.56 (.37)	.11	.44 (.40)	.08	.30(.40)	.05
Justification of knowledge	.14 (.46)	.02	.13 (.49)	.02	.19(.49)	.09
Development of knowledge	.20 (.36)	.04	-.43 (.38)	-.07	-.50(.38)	-.08

Note. *p < .05, **p < .01, ***p < .001.

The results of the multiple regression analyses for epistemological dimensions predicting number of supportive arguments, counterarguments and rebuttals showed that only the epistemological dimension of simplicity of knowledge explained a statistically significant amount of variance in the number of supportive arguments (R^2 =.19, $F_{(9,228)}$=6.27, p<.001), the number of counterarguments (R^2 =.15, $F_{(9,228)}$=4.03, p<.001) and the number of rebuttals (R^2 =.17, $F_{(9,228)}$=5.26, p<.001) produced. The other dimensions did not explain a statistically significant amount of variances.

This indicates that the more participants believed that the structure of knowledge is not simple and knowledge consists of interrelated concepts and ideas rather than isolated facts (relatively sophisticated simplicity beliefs) the more likely they were to construct more supportive arguments, counterarguments and rebuttals during their IR on health-SSIs.

Table 4 Results of multiple regression analysis for epistemological dimensions predicting participants' quality of arguments.

Predictor	Quality of supportive arguments		Quality of counterarguments		Quality of rebuttals	
	B(SE)	β	B(SE)	B	B(SE)	B
Certainty of knowledge	1.56 (1.41)	.08	1.70(1.25)	.10	1.28(1.25)	.07
Simplicity of knowledge	3.16(1.27)	.16**	4.71(1.13)	.27***	3.93(1.13)	.22**
Source of knowledge	-.16 (1.10)	-.09	-1.56(.99)	-1.02	-1.55(.99)	-.10
Justification of knowledge	-.36 (1.47)	-.02	-.64(1.30)	-.03	-.23(1.31)	-.01
Development of knowledge	-.58 (1.48)	-.03	-1.70(1.31)	-.09	-1.53(1.32)	-.09

Note. *p < .05, **p < .01, ***p < .001.

The results of the multiple regression analyses for variables predicting quality of arguments, counterarguments and rebuttals showed that only the epistemological dimension of simplicity of knowledge explained a statistically significant amount of variance in the quality of supportive arguments (R^2 =.17, $F(9,227)=5.00$, $p<.001$), of counterarguments (R^2 =.20, $F(9,227)=6.01$, $p<.001$) and of rebuttals (R^2 =.20, $F(9,227)=6.28$, $p<.001$) produced. The other dimensions did not explain a statistically significant amount of variances. This indicates that the more participants believed that the structure of knowledge is not simple and knowledge consists of interrelated concepts and ideas rather than isolated facts (relatively sophisticated simplicity beliefs) the more likely they were to construct of higher quality supportive arguments, counterarguments and rebuttals on health-SSIs.

4. Discussion and conclusions

We investigated whether pre-service primary teachers' epistemological beliefs could contribute significantly to the quality of their informal reasoning during heuristic/spontaneous processes and analytic/reflective processes on health-SSIs. We used dual-process theories to represent the underlying cognitive process. In the current study, quantity and quality of the different types of supportive arguments, counterarguments and rebuttals on health-SSIs, were employed as indicators for the quality of participants' IR (Kuhn, 1992; 1993; Sadler & Zeidler, 2005a; Wu & Tsai, 2011).

The results indicated that the participants who exhibited relatively sophisticated beliefs about the structure of knowledge - simplicity beliefs - constructed greater quantity and quality of supportive arguments, counterarguments and rebuttals on controversial health-SSIs during their IR, than participants who held less sophisticated beliefs about the structure of knowledge. Our findings are in line with previous studies carried out by other researchers. For example, Bell & Linn (2000) found that students who viewed science as dynamic, complex and constantly changing, tended to create more complex and integrated arguments. On the other hand, Nussbaum & Bendixen (2003) demonstrated that students who believe that knowledge is simple, certain and unchanging, reported that arguments were anxiety-promoting and thus they tended to avoid them.

A possible explanation for our findings is as follows: Participants who view knowledge as a collection of simple and discrete pieces of factual information (relatively naïve simplicity beliefs), rather than as a complex system of organized theoretical concepts and ideas (relatively sophisticated simplicity beliefs), may tend not to look for more information from multiple sources, not to take into account conflicting viewpoints and conflicting information, not to consider alternative views, and not to recognize different aspects of a health-SSI for decision making. In contrast, participants with relatively sophisticated beliefs about the structure of knowledge - simplicity epistemological beliefs - take into account conflicting viewpoints and conflicting information of a health-SSI, consider alternative views, recognize different aspects of a health-issue, use more evidence overall, select different information sources and not only sources consistent with their views, construct more and higher quality supportive arguments, counterarguments and rebuttals, during their IR for decision making on health-SSIs.

The findings of this study have implications on health education. A main implication leads to highlighting the importance of fostering pre-service primary teachers' epistemological beliefs, especially beliefs about the structure of knowledge (simplicity dimension) by dealing with health-SSIs. Interventions providing students with opportunities to deal with controversial health-SSIs, for which there is no single right answer (Sadler & Zeidler, 2005a; Wu & Tsai, 2011), to check out the spontaneous responses with deeper processing (analytical thinking) and to clarify and evaluate probability, risks, thoughts and arguments, appears to be a promising pathway for supporting students to develop EBs and IR on health-SSIs. In this way, they would have practice collecting conflicting information from multiple sources and different domains, incorporating reflective thinking about the nature of knowledge and knowing, analyzing and weighing up conflicting evidence, making claims, constructing supportive arguments, counterarguments and rebuttals and making decisions at personal or societal level. By doing this, we can prepare future teachers to improve their epistemological beliefs and the quality of their IR regarding health-SSIs, and their ability to foster development of these practices among their future students. This is very important in a developed world, because people often have to manage health-SSIs, and additionally they have to cope with alternative treatments and solutions, differing interpretations of symp-

toms, or even several interpretations of the underlying pathological explanations of illnesses and risks related to health-SSIs, at personal and/or societal level.

Finally, there are some limitations to the current study that may provide impetus to further work in this area. First, although the health-SSIs addressed in the current study are of international applicability, we cannot impute generalizability for our results based on a relative small sample consisted of 93% females. Second, because we used only questionnaire data in the current study, we could not probe participants' responses to items as with in-depth interviews. Future studies should take a closer look at the interplay between epistemological beliefs and informal reasoning on health-SSIs.

References

Barrouillet, P. (2011) Dual-process theories of reasoning: The test of development. *Developmental Review, 31*, 151-179.

Barzilai, S. & Eshet-Alkalai, Y. (2015) The role of epistemic perspectives in comprehension of multiple author viewpoints. *Learning and Instruction, 36*, 86–103.

Baytelman, A. & Constantinou, C. (2016) Development and validation of an instrument to measure epistemological beliefs for science. In J. Lavonen, K. Juuti, J. Lampiselkä, A. Uitto & K. Hahl (Eds.), *Electronic Proceedings of the ESERA 2015 Conference*. Science education research: Helsinki, Finland: University of Helsinki. ISBN 978-951-51-1541-6.

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

Bell, P. & Linn, M.C. (2000) Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education, 22*, 797–817.

Bråten, I. & Strømso, H. (2010) Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes, 47*, 1-31.

Conley, M., Pintrich, P., Vekiri, I. & Harrison, D. (2004) Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology, 29*, 186–204.

Evans, J. St. B.T. (2003) In two minds: Dual-process accounts of reasoning. Trends in *Cognitive Sciences, 7*, 454-459.

Evans, J. St. B. T. (2008) Dual- Processing accounts of reasoning. Judgment and Social Cognition. *Annual Review of Psychology, 59*, 255-278.

Evans, J. St. B. T. & Curtis-Holmes J. (2005) Rapid Responding Increases Belief Bias: Evidence For The Dual-Process Theory Of Reasoning. *Thinking & Reasoning, 11*, 382–389.

Evans, J. St. B. T. & Stanovich, K. E. (2013) Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*, 223-241, 263-271.

Hofer, B. K. & Pintrich, P. R. (1997) The development of epistemological theories: beliefs about knowledge and knowing their relation to learning. *Review of Educational Research, 67*, 88-140.

Hofer, B. K. & Bendixen, L. D. (2012) Personal epistemology: Theory, research, and future directions. In K. Harris, (Ed.), *Handbook of educational psychology*, pp. 225–254. Washington, DC: American

Psychological Association.

Kahneman, D. (2011) *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kolstø, S. D. (2001) Scientific literacy for citizenship: Tools for dealing with the science dimension of controversial socio-scientific issues. *Science Education, 85*, 291–310.

Kuhn, D. (1992) Thinking as argument. *Harvard Educational Review*, 62, 155-178.

Kuhn, D. (1993) Connecting scientific and informal reasoning. *Journal of Developmental Psychology, 39*, 74-103.

Kuhn, D., Cheney, R. & Weinstock, M. (2000) The development of epistemological understanding. *Cognitive Development, 15*, 309–328.

Lee, Y.C. (2012) Socio-Scientific Issues in Health Contexts: Treading a rugged terrain, International *Journal of Science Education, 34*, 459–483.

Nunnally, J. C. (1978) *Psychometric theory*. New York: McGraw Hill.

Nussbaum, E.M. & Bendixen, L.M. (2003) Approaching and avoiding arguments: The role of epistemological beliefs, need for cognition, and extraverted personality traits. *Contemporary Educational Psychology, 28*, 573–595.

Ratcliffe, M. & Grace, M. (2003) *Science Education for Citizenship*, Maidenhead: Open University Press.

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

Sadler, T., & Fowler, S. (2006) A threshold model of content knowledge transfer for socioscientific argumentation. *Science Education, 90*, 986-1004.

Sadler, T.D. & Zeidler, D.L. (2004) The Morality of Socioscientific Issues. Construal and Resolution of Genetic Engineering Dilemmas. *Science Education, 88*, 4-27.

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

Sadler, T. D. & Zeidler, D. L. (2005b) The significance of content knowledge for informal reasoning regarding socioscientific issues: Applying genetic knowledge to genetic engineering issues. *Science Education, 89*, 71–93.

Schommer, M. (1990) Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*, 498–504.

Strømso, H.I., Braten, I. & Britt, M.A (2011) Do students' beliefs about knowledge and knowing predict their judgement of texts' trustworthiness? *Educational Psychology, 31*, 177-206.

Weinstock, M. & Cronin, M.A. (2003) The everyday production of knowledge: Individual differences in epistemological understanding and juror-reasoning skill. *Applied Cognitive Psychology, 17*, 161–181.

Wu, Y. & Tsai, C. (2011) High School Students' Informal Reasoning Regarding a Socio-scientific Issue, with Relation to Scientific Epistemological Beliefs and Cognitive Structures. *International Journal of Science Education, 33*, 371-400.

Zohar, A. & Nemet, F. (2002) Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching, 39*, 35-62.

Appendix

Examples of questions from the open-ended questionnaire for the construction of arguments, as well as examples of students' supportive arguments, counterarguments and rebuttals.

Health-SSI example: consumption of bottled water versus tap water

The city's municipal authorities and representatives of the State Laboratory argue that the tap water is suitable for consumption and there is no danger from consuming it. They also affirmed that tap water costs a tiny amount compared to bottled water. Another group of experts dealing with health issues, as well as importers and retailers of bottled water argue that the only safe drinking water is bottled water. Moreover, they argue that although it costs slightly more than the tap water, its advantages are so numerous that buying it is the best choice.

Example 1:

PART A

1. Are you in favour of bottled water or tap water? Underline as: <u>Bottled water</u> / Tap water

2. If you want to convince a friend about your position, which arguments will you propose to convince him/her?

I prefer bottled water because in tap water there are organic pollutants which react with chlorine used for the chlorination of tap water and produce additional harmful substances for our bodies and health.

PART B

If somebody holds an opposite position from you on this issue, which arguments may he/she have?

Toxic substances migrating from water plastic bottles, such as diphenyl A, which are more dangerous for our bodies than chlorine, causing serious health problems such as cancer.

PART C

According to the arguments you have mentioned in Part B, can you write down your opposite ideas to justify your position?

The quantity of toxic substances migrating from plastic bottles is extremely low and not

dangerous, in contrast to the high amount of chlorine which is added to tap water.

Example 2:
PART A
1. Are you in favour of bottled water or tap water? Underline as: Bottled water / Tap water
2. If you want to convince a friend about your position, which arguments will you propose to convince him/her?
I prefer tap water because high quantities of petroleum are used in the manufacture of plastic bottles and it is dangerous for our health and pollutes the environment.
PART B
If somebody holds an opposite position from you on this issue, which arguments may he/she have?
Plastic bottles can be recycled and so there is no problem with pollution.
PART C
According to the argument you have mentioned in Part B, can you write down your opposite ideas to justify your position?
Although plastic can be recycled, the thoughtless and unnecessary consumption of petroleum pollutes the environment and harms health.

"They implant this chip and control everyone." 'Misuse of science' as a central frame in students' discourse on neuroscientific research

Alexander Bergmann[1] and Jörg Zabel[1]

alexander.bergmann@uni-leipzig.de

Abstract

This study investigates the students' perspective on the ethical, social, and legal dimension of neuroscience by involving the students in an everyday-like discourse. In a thought experiment, eight small groups of German grade 9 and 10 grammar school students (13-16 years old) had to decide as members of an ethical commission whether three artificial neuroscientific research proposals should be funded. Analysing the resulting discussions by the Documentary Method revealed that 'misuse of science' is a central frame the students use when discussing science and its relation to society. The article discusses how to deal with this challenging belief both in science education research and science teaching.

Keywords: discourse, documentary method, neuroscience, misuse of science, students' perspective

---

[1] University of Leipzig, Germany

1. Introduction

Neuroscience is expected to be one of the key scientific disciplines of the 21st century. It raises important ethical, social, and legal questions, e.g. about adequate treatment strategies for neurodegenerative diseases, or the development of neuroenhancers to increase human cognitive abilities. Fostering students' competence to participate in public discourse on these and other socio-scientific issues is a major goal of science education (Zeidler et al., 2005; Zeidler & Nichols 2009). This encompasses sufficient knowledge of science content and about science, as well as the ability to 'undertake socio-scientific decision-making' (Holbrook & Rannikmae, 2009, p. 283).

In order to reach this goal and to contribute to a realistic public view on neuroscience and its relationship to society, we emphasise the potential of discursive settings in the science classroom. Discourse is 'central to the way communities collectively construct norms and expectations, define common knowledge for the group, build affiliations, frame disciplinary knowledge, and invite or limit participation' (Kelly, 2014, p. 321). This discourse should involve the ethical, legal, and social aspects of neuroscience. Furthermore, it must address students' individual perspectives on these aspects. This challenges science teachers, because discursive teaching settings on bioethical topics usually are perceived as much more open-ended and vivid (Alfs, 2012). To know how their students frame neuroscientific topics and how they reason on them in everyday life may support teachers in coping with the 'threatening' open-endedness and vividness of bioethical discourse in the science classroom.

Therefore, we investigated how students perceive neuroscientific research and how they reason about its ethical dimension in an everyday-like discourse. A special focus was kept on the learners' beliefs and everyday myths (Gebhard, 2007) related to neuroscience.

To involve students in discourse about neuroscientific research, a thought experiment was designed. As members of an ethical commission, small groups (N=8) of German 9th and 10th grade students autonomously had to decide whether three neuroscientific research projects should be funded. The students' discussions were videotaped and analysed by means of the Documentary Method (Bohnsack, 2003).

In the students' discussions, misuse of science was a central and recurring

topic. Using two example passages from the participants' discussions, we describe in detail how students perceive the potential misuse of neuroscience. Furthermore, we compare it to the students' perspective on genetic engineering, and discuss the consequences of our findings for science teaching.

2. Theoretical background

In recent years, discourse in the science classroom has become an important area in science education research (e.g. Kelly, 2014, 2007; Driver et al., 2000). We here combine the discursive approach with Haidt's social-intuitionist model of moral reasoning (2001). Also, the framework suggested here assumes that students' beliefs (Jones & Carter 2007; Pajares, 1992) and everyday myths (Gebhard, 2007) have significant influence on their decision-making processes.

2.1 Analysing bioethical discourse in the science classrooms

As language in use (Jaworski & Coupland, 1999), discourse in general is 'constructed among people in some context, with some history, projections of future actions, and ideological commitments' (Kelly, 2014, p. 322). Discourse participants engage in a collaborative and dynamic process of co-construction of knowledge; they do not only interact but 'interthink' (Mercer, 2004). Engaging students in discourse in the science classroom helps them to learn the scientific language and practices, but also to act as competent members of a science-based society and to engage in decision-making processes. The public discourse on genetic engineering during the last decades illustrates how difficult it is to reach this goal. Exaggerated media representations, lobbyism, and misleading publicly shared myths cause irrational fears and beliefs about genetics. Limiting the focus on the content matter may lead to students who mistrust scientific endeavour in general (Gebhard & Mielke, 2001).

So far, a number of studies have examined the role of argumentation and explanation, particularly the abilities of students to build thorough arguments in various scientific settings (e.g. Osborne, 2010; Grace, 2009; Driver et al., 2000). Other works focused on the ethical and societal dimensions of science and scientific results. Socioscientific issues transcend 'pure' scientific argumentation but also cover value-related moral and ethical reasoning (e.g. Walker & Zeidler, 2007; Albe 2008). While many authors draw upon a mostly rationalist and individualist perspective when it comes to decision-making in science

education (e.g. Hößle, 2001; Reitschert, 2009; Bögeholz et al., 2014), we suggest this perspective to be inadequate to describe students' everyday ethical discourse.

Why? In his social-intuitionist model of moral reasoning Haidt (2001) suggests that moral judgement and reasoning are significantly influenced by moral intuitions. The latter can be defined as 'sudden appearance in consciousness of a moral judgment, including an affective valence (good-bad, like-dislike), without any conscious awareness of having gone through steps of searching, weighing evidence, or inferring a conclusion' (Haidt, 2001, p. 818). These intuitions form moral judgements that refer to a 'set of virtues held to be obligatory by a culture or subculture' (Haidt, 2001, p. 817). Only if persons are required to justify their moral judgements externally, they engage in an effortful process to search for arguments supporting the already made judgement. The social part of Haidt's model suggests that moral judgements are sensible for social persuasion, and are moderated by social context or group norms.

Consequently, when analysing the students' perspective on the social, legal and ethical dimension of neuroscience, a research setting is preferable that allows them to autonomously engage in moral reasoning together with their peers. The way their discourse unfolds relies heavily on the groups' former experience and their sociocultural context, their knowledge, values, attitudes and beliefs about science in general, and about neuroscience in specific. The analysis of autonomous discourse of student groups makes it possible to derive a set of age- and milieu-specific beliefs about neuroscience and its ethical dimension.

2.2 Students' beliefs and everyday myths

Whereas the notion of 'knowledge' refers to scientifically accepted constructs, beliefs are usually unjustified or only justified by individual experience. They are 'naïve' conceptions, subsumed in the knowledge construct and viewed as an integral part of schema (Jones & Carter, 2007). As cognitive constructs they also differ from attitudes, which are a persons' 'predisposition to respond positively or negatively to things, people, places, events, or ideas' (Simpson et al., 1994, p. 212). As individual representations of the physical and social reality, beliefs influence our perception and are crucial for the way we structure our actions and behaviour, e.g. engaging in discourse on neuroscience (Pajares,

1992; Rokeach, 1968). Everyday myths, in contrast to beliefs, rather characterize the affective dimension of the students' perspective on biological topics. They can be found in symbolically charged biographical notions and stories (Gebhard, 2007, p. 120). This approach originates in cultural philosophic and cultural psychological theory (Boesch, 1980; Cassirer, 1972) but also refers to research on students' conceptions (e.g. Duit, 1992). Everyday myths comprise aspects of the self-image, conception of man and a worldview, they reflect the values, interests and behaviours of a person (Combe & Gebhard, 2007, p. 63). Therefore, everyday myths are likely to play an important role also in students' discourse on neuroscience.

2.3 Students' thinking about the ambivalence of scientific endeavour

With every step towards a more detailed scientific description of our brain and its function, new ethical issues arise. Neuroethics concerns itself with philosophical and ethical dimensions as well as societal impacts of these current developments (Hildt, 2012). It reflects on the ethical implications of brain research and clinical practice, and discusses the consequences for society and the conceptualisation of man (Roskies, 2002). As neuroscience allegedly deals with the core of human existence, it draws lot of attention in the public discourse and the media. Exaggerations, pseudoscientific explanations and neuromyths (Dekker et al., 2012; Herculano-Houzel, 2002) sometimes lead to an irrational perspective on neuroscience, e.g. the idea to 'hack' other people's brains.

Regarding the complexity of the topic, the influence on the public discourse, and the trend to mystification, both neuroscience and genetics have a lot in common. But in contrast to neuroscience, genetics has been an established topic in science curricula for nearly three decades, and much research has been carried out on the student's perspective. Our study may help to understand the students' perspective on neuroscience.

Several authors analysed attitudes, knowledge, and beliefs concerning genetic engineering, genetic screening or genetically modified food (Gebhard et al., 1994; Baalmann et al., 1998; Lewis & Kattmann, 2004). In an exploratory study on students' everyday myths about genetics, German 11th grade students were asked to reflect on the relationship between man, nature, and genetics in group discussions (Gebhard & Mielke, 2001). In summary, the students perceived nature as a normative instance which should guide human action.

'What is natural is good' was one of the most influential everyday myths in their discussions. In consequence, they perceived mankind as counterpart of nature, not legitimated to intervene in natural processes.

Within the discussions they also debated on the 'ambivalence of scientific endeavour'. Human curiosity led to problematic knowledge, which could be used for good and dangerous purposes at the same time. E.g., selective reproduction techniques and genetic screening offer 'godlike abilities', threatening human individuality and excluding the naturally born (Gebhard & Mielke, 2001). The students also expressed the wish for genetic engineering to be ethically controlled. Overall, they were concerned about the future developments in the field of genetic engineering.

3. Methods and research design

The current investigation was carried out to answer the following questions: (1) What beliefs do students hold about neuroscience? and (2) how do these beliefs influence the students' discourse on neuroscientific research?

To answer these questions, a thought experiment was designed. Small groups of German grammar school students (grade 9 and 10) autonomously had to decide on the funding of imaginary neuroscientific research projects as members of an ethical commission. We adopted the group setting as a standard procedure from previous studies on students' perspective on other ethical relevant topics, for example on genetic engineering (Gebhard & Mielke, 2001). A modification was made concerning the role of the researcher. In contrast to guided group discussions in other research settings, our group process was not interrupted by the researchers, thereby providing a more everyday-like, 'natural' discussion within the student groups.

The thought experiment was presented to the students with a manual, containing (1) an introductory text, (2) three research proposals and (3) a task sheet. The introduction set up the context of the thought experiment. In each group, up to four students were to act as members of an ethical commission. The commission's task was to evaluate the ethical dimension of three fictional research proposals from leading neuroscientific researchers and to decide if the research should be funded or not. The proposals were to be evaluated independently from another. Figure 1 illustrates Proposal I and the general structure of the research proposals.

Neurodegenerative Diseases and Brain-Computer-Interfaces *Principle investigator / Head of research group: Professor Ramón y Cajal*	**title and research group**
Our research group's main goal is to investigate the structure, function and development of the human brain. Furthermore, we are developing technological devices to fight symptoms of multiple neurodegenerative diseases such as Alzheimer's or Parkinson's. Our latest invention is a fingertip-sized computer chip that is implanted directly into the brain, a so-called Brain-Computer-Interface. This digital brain-'pacemaker' measures the patient's brain activity and supports damaged brain areas via micro-electrical impulses. Clinical trials with rats, great apes, and diseased people show that symptoms like tremor can be treated successfully.	**former research**
Further funding of our research will help to improve the function of the Brain-Computer-Interface, to minimize the surgical risks, to fix issues of non-compatibility, and to restore our patients' sensomotoric abilities.	**recent research**
In the future, it may also become possible to control machines via Brain-Computer-Interfaces, as well as to apply this technology to other areas such as road traffic, aviation, and heavy industry.	**speculative moment**

Figure 1 Proposal I and general structure of the research proposals.

All proposals contained a short description of the former research, a recent research goal and, in the last paragraph, some speculative elements about what might be possible in the future. The proposals were concerned with the use of Brain-Computer-Interfaces, neuroimaging techniques and neuropharmaceuticals, as these topics are central within the neuroethical and philosophical discourse on neuroscience and at the same time have a broad reception in the public. Introductory text and research proposal were formulated in a factual and objective tone, in order not to explicitly lead the students into questioning the use and misuse of science (Figure 1).

The participants were given 60 minutes to discuss all research proposals, and they were required to find a unanimous decision. They were also asked to prepare a brief report to a scientific committee (the researcher), providing reasons for their decision. The data of eight groups were collected at a grammar school in Ilmenau (Thuringia, Germany) in December 2015. All students participated voluntarily. The groups consisted of peers, which also contributed to a more everyday-like discourse. The group discussions and the reports to the scientific committee were videotaped, transcribed, and analysed by means of the Documentary Method (Bohnsack, 2003). This is a method for the analysis of group discussions, already well-established in ethnography. It focuses on the specific manner in which different groups discuss the same topic. As a

result, so-called *frames* are reconstructed and explicitly formulated. A frame is an expression of an orientation or belief, collectively shared throughout several groups that belong to the same socio-cultural context.

To reconstruct the students' frames concerning neuroscience, we used a three-step model (Przyborski, 2004):

the group discussion data were divided into topic-specific passages (Formulating Interpretation).

the content of each passage and the participants' interactions were analysed as to their discourse patterns (Reflecting Interpretation).

thematically identical passages were systematically compared within and across groups to reconstruct central frames guiding the students' discourse on neuroscientific research (Frame Description).

Constantly comparing passages from different groups widened the scope of the interpreter's perspective. Discussing each interpretative step within the research group provided a discursive validation of the analysis results.

4. Results

The goal of this investigation was to analyse how students perceive neuroscientific research, with a special focus kept on the ethical dimension of their perspective. In fact, we were able to reconstruct a broad set of frames documented in the data. Our assumption is that these frames in discourse indicate central beliefs of the respective sociocultural group.

In this article, we focus on a frame we call 'misuse of science', which occurred frequently in several groups. As it was abundant in the students' discussions, it seems to be important for their perception of neuroscience in general.

In the following, we present and overview on the dimensions of 'misuse of science'. We illustrate this frame with two passages from the group discussions. Both were translated from German, slightly abbreviated and edited for better understanding. The transcripts also indicate restarts of sentences ('/'), interruptions through other speakers ('//…//'), and comments made by the interpreter.

4.1 Overview on the dimensions of 'misuse of science'

All groups eventually agree that there is a risk of misuse in their respective proposal. Table 1 provides a detailed overview of 'misuse of science'. It illustrates who misuses the proposed research, and for what purpose.

Table 1 Overview on possible misuse scenarios of Brain-Computer-Interfaces, neuroimaging techniques, and neuropharmaceuticals, as perceived by the students.

Misuse by …	Brain-Computer-Interfaces	Neuroimaging techniques	Neuropharmaceuticals
Individuals	• Hacking other people's brains and manipulating them • Using technology for terror attacks (e.g. plane crashes and bombing)		• Increasing one's abilities and (mental) capacity
Military	• Developing new weapons and super soldiers (Cyborgs)	• Mindreading in military interrogations	• Developing power-enhancing drugs for soldiers
Politicians		• Spying out and eliminating oppositionists • Monitoring and controlling the population	• Modifying moral thinking and values of people (e.g. national socialists) • Tranquilizing oppositionists
Companies	• Using chips to monitor and control consumer behaviour		• Increasing the productivity of workers by keeping them awake and productive

The detailed analysis revealed further aspects of the frame 'misuse of science'. First of all, it can be characterized by its ubiquity. Nearly every (sub) topic the students talked about in their eyes implied at least some danger of misuse. Another facet of 'misuse of science' refers to the question who is misusing it and who are the possible victims. Within the students' discourse, usually there is a notion of a hierarchical scale amongst society linked to misuse. Politicians, rich people, companies or military leaders tend to be the malevolent agents who dominate and manipulate the poor, or generate individual advantages in other ways. Thereby, the gap between rich and poor will deepen even further, because the poor cannot make use of any of the new developments, e.g. enhancement via brain-computer interfaces or neuropharmaceuticals. Neuroscientific research therefore not only leads to misuse, but also fosters inequity and contributes to segregation. As a matter of fact, some of the groups extended this line of thought and expected our society to completely decay, if not mankind to be eliminated.

Interestingly, in the students' view the scientists themselves are not actively involved in this process at all. Although they do the research and derive results from it, they do not control the subsequent application of these results. From the students' perspective, the only way for the researchers to intervene is before the research is done. Afterwards, there is no chance for them to prevent the misuse. In the long term, the misuse of science cannot be prevented, it will definitely happen in the future.

The following two examples illustrate how the students use the frame 'misuse of science' to structure their debate on Proposal I.

4.2 Example I – Group 6 comments on Proposal I

In this example, Group 6 comments on the use and misuse of brain-computer interfaces. The group consists of three male tenth-grade students aged 15 to 16 years. They discussed in a very calm manner, with comparatively long turns and very little interruptions by other speakers. From the very beginning, all group members appear to have a joint perspective on this Proposal, see Figure 2.

1	T:	In my opinion this proposal aims not only at curing diseases. I believe this will turn
2		into the complete opposite/ it will be misused for the exact opposite value. That is a
3		military purpose. For it is commonly known that research for military purposes
4		always tends to be generously funded.
5		
6	L:	Hum ok. The fundamental idea of this proposal is good. They want to cure diseases
7		like Alzheimer's, Parkinson or Dementia. But the risk of misuse is quiet big. I agree,
8		they will misuse the BCIs for military purpose.
9		
10	B:	Yes of course. It is well known: the USA wants to build super soldiers.
11	T:	Ok, the next question is …

Figure 2 Group 6 – Passage 'risk of misuse'.

Student T suggests that Proposal I contains not only benefits, but also can be misused 'for the exact opposite value', namely a military purpose. Student T assumes that military purposes are generally well funded, and that scientific research is exploited for these purposes. He doesn't specify what kind of exploitation or misuse he means. Student L elaborates the positive aspects of Proposal I briefly, but also approves that there is a big risk of misuse for military applications. Student B exemplifies this with reference to the project of US-American 'super-soldiers', which he regards to be 'well-known'. Student T concludes by affirming this example, and suggests moving on to another

topic. By validating, elaborating and exemplifying student A's proposition, the group demonstrates its agreement on the validity of the proposed content. In the eyes of this group, 'misuse of science' is a fact beyond any doubt.

4.3 Example II – Group 2 comments on Proposal I

In this example, Group 2 comments on the use and misuse of brain-computer interfaces and neuroimaging techniques. The group consists of four ninth-grade students aged 14 to 15 years, three of them female (Le, La, K) and one male (P). The passage is located in the first quarter of the whole discussion. In comparison to other groups, there were more interruptions, and turn-taking was much more dynamical. Two opposing factions appear in the passage. Whereas LE and K emphasise the benefits of Proposal I, P and LA tend to reject it, see Figure 3.

1	**Le:**	I like the idea of using neuroimaging techniques in criminal proceedings.
2	**P:**	(provokingly) Well, you want to punish someone before he has committed a crime
3		just because he thought about it?
4	**La:**	(supporting P) Precisely! It is the only thing//K: The question is// that I do not like
5		at this proposal.
6	**K:**	No, only if someone has committed a crime, they do something with his brain.
7	**La:**	But our thoughts should/ nobody should be able to read our thoughts//K: Yes//
8		aren't our thought the last and only things that remain free for us human beings?
9	**P:**	And you would test everybody to prevent/ read the mind of everybody?
10	**Le:**	No, just of those, who attracted attention with criminal behaviour.
11	**La:**	Don't you think, some rich people would misuse this technique? There is no way
12		back, if they develop it. They won't do this only in criminal proceedings. It will be
13		used everywhere.
14	**Le:**	It is the same with the computer chips in proposal one. They implant this chip and
15		control everyone.
16	**La:**	Yes.
17	**Le:**	Possibly.

Figure 3 Group 2 – Passage 'risk of misuse'.

The passage falls into two subsequent sections, line 1-10 and line 11-17, see Figure 3. In the first section, Le proposes that the use of neuroimaging techniques has several benefits for crime prosecution, without further exemplification. P asks a provocative question, as in his opinion, this proposal threatens the freedom of thoughts. While La is supporting P's antithetical reference, K elaborates Le's first idea of using neuroimaging techniques in criminal trails.

In her opinion, they should only be used in case someone already has committed a crime before. Le and K don't realize or maybe ignore the ethical concerns about the invasion of privacy and the presumption of innocence. Instead, they refer to a simple 'technological pragmatism', which basically implies that once technological developments are developed, they should be used. In opposition, La and K neglect the benefits and instead point out the danger of mind-reading.

The oppositional setting changes instantly when La introduces the possibility of misuse of science (line 11-13). Rich people would use this technique not only in criminal proceedings, but also in the public. Although she doesn't specify what sort of misuse she expects, the group instantly and entirely agrees with her concerns. In this passage, the notion of 'misuse of science' suddenly neutralises the arguments in favour of the proposal that had been previously uttered. La's concerns, however fuzzy they remain, dominate over the pragmatic arguments and the perspective of benefits for crime prosecution that had influenced the discourse before.

5. Discussion and conclusion

In summary, the frame 'misuse of science' appears as powerful and essential common ground in students' discussions about neuroscience, as far as our data indicate. It is interesting to compare these results to former research on the students' perspective on genetics (Gebhard & Mielke, 2001). Both genetics and neuroscience raise questions about the normative reflection and regulation of the process of scientific endeavour. Whereas the threat to nature is the strongest concern students hold against genetic engineering, they rather worry about the societal dimension when it comes to neuroscience. In their eyes, neuroscientific research is a peril for a 'natural' social equilibrium. In their perspective, the recent arrangement of society should be conserved. Neuroscientific research threatens this arrangement by providing opportunities to manipulate and exploit the weaker, the poorer, and those with less political power. A decisive difference between the students' perception of genetics and neuroscience is the degree of their respective risks. Whereas genetic engineering 'only' opens the way to selective reproduction, neuroscience may allow mass oppression and ubiquitous misuse.

We found that misuse is a central and ubiquitous belief that many students

hold about neuroscience, combined with a rather pessimistic view of scientists' influence. This raises several questions. Is this belief limited to neuroscience, or do students expect science in general to be misused for manipulation and oppression? Is there a relation between beliefs about misuse of science and the raising distrust of science in public, e.g. about vaccines and climate change? And if so, what are the consequences for science teaching? To answer these questions, further research is necessary, including studies on the students' perspectives on neuroscience, but also other scientific topics. Taking into account the recent developments in synthetic biology (e.g. CRISPR/Cas), maybe also the students' perspective on genetics should be revisited.

On a more practical level, we suggest to collectively reflect and discuss the students' beliefs about misuse of science in the classroom (Dittmer & Gebhard, 2012). As 'misuse of science' significantly influences students' decision-making processes and view of science, it cannot be ignored in science teaching. We suggest implementing historical case studies of science misuse in science lessons and thereby discussing about the responsibility of scientists, research associations or policymakers. Science teaching on current findings in bioscience should integrate topics such as precautionary principles, dissemination of scientific results, responsibility in science, and how policymaking, science and society interact. Science teachers should discuss the existing efforts to bridge the gap between science and society (e.g. Responsible Research & Innovation, Science with and for Society, Citizen Science). We want future generations to be alert and to prevent any misuse of science, however for this very reason they need to develop a more balanced and appropriate view of scientific progress. Pointing out the social embeddedness of all scientific endeavours may help the students to do so.

References

Albe V. (2008) When Scientific Knowledge, Daily Life Experience, Epistemological and Social Considerations Intersect: Students' Argumentation in Group Discussions on a Socio-Scientific Issue. *Research in Science Education*, 38, 67-90.

Alfs, N. (2012) *Ethisches Bewerten fördern. Eine qualitative Untersuchung zum fachdidaktischen Wissen von Biologielehrkräften zum Kompetenzbereich „Bewertung".* Hamburg: Dr. Kova .

Baalmann, W., Frerichs, V., Gropengießer, H., & Kattmann, U. (1999) Das Modell der Didaktischen Rekonstruktion. Untersuchungen in den Bereichen „Genetik" und „Evolution". In R. Duit & J. Mayer (Eds.), *Studien zur naturwissenschaftsdidaktischen Lern-und Interessenforschung* (pp. 82-92). Kiel: IPN.

Boesch, E.E. (1980) *Kultur und Handlung: Einführung in die Kulturpsychologie*. Bergen/Stuttgart/Wien: H. Huber.

Bögeholz, S., Böhm, M., Eggert, S., & Barkmann, J. (2014) Education for Sustainable Development in German Science Education: Past-Present-Future. *Eurasia Journal of Mathematics, Science and Technology Education*, 10(4), 231-248.

Bohnsack, R. (2003) *Rekonstruktive Sozialforschung: Einführung in qualitative Methoden*. Opladen: Leske & Budrich.

Cassirer, E. (1972) *An essay on man: An introduction to a philosophy of human culture*. Yale: Yale University Press.

Combe, A., & Gebhard, U. (2007) *Sinn und Erfahrung: Zum Verständnis fachlicher Lernprozesse in der Schule*. Opladen & Farmington Hills: Verlag Barbara Budrich.

Dekker, S., Lee, N.C., Howard-Jones, P., & Jolles, J. (2012) Neuromyths in education: Prevalence and predictors of misconceptions among teachers. *Frontiers in Psychology*, 429(3), 1-8.

Dittmer, A., & Gebhard, U. (2012) Stichwort Bewertungskompetenz: Ethik im naturwissenschaftlichen Unterricht aus sozial-intuitionistischer Perspektive. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 81-98.

Driver, R., Newton, P. & Osborne, J. (2000) Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287–312.

Duit, R. (1992) Forschungen zur Bedeutung vorunterrichtlicher Vorstellungen für das Erlernen der Naturwissenschaften. In K. Riquarts, W. Dierks, R. Duit, G. Eulenfeld, H. Henning, & H. Storck (Eds.), *Naturwissenschaftliche Bildung in der Bundesrepublik Deutschland*, Bd. IV, (pp. 47-84). Kiel: IPN.

Gebhard U., & Mielke R. (2001) Selbstkonzeptrelevante Aspekte der Gentechnik. In H. Mummendey (Ed.), *Bielefelder Arbeiten zur Sozialpsychologie. Nr. 200. Psychologische Forschungsberichte*. Universität Bielefeld.

Gebhard, U. (2007) Intuitive Vorstellungen bei Denk-und Lernprozessen: Der Ansatz „Alltagsphantasien ". In D. Krüger & H. Vogt (Eds.), *Theorien in der biologiedidaktischen Forschung* (pp. 117-128). Berlin/Heidelberg: Springer.

Gebhard, U., Feldmann, K., & Bremekamp, E. (1994) Hoffnungen und Ängste. Vorstellungen von Jugendlichen zur Gentechnik und Fortpflanzungsmedizin. In E. Bremekamp (Ed.), *Faszination Gentechnik und Fortpflanzungsmedizin. Informationen und Materialien für die politische Jugendbildung* (pp. 11-25). Bad Heilbrunn: Klinkhardt.

Grace, M. (2009) Developing High Quality Decision-Making About Biological Conservation in a Normal Classroom Setting. *International Journal of Science Education*, 31(4), 551-570.

Haidt, J. (2001) The emotional dog and its rational tail: a social intuitionist approach to moral judgement. *Psychological Review*, 108(4), 814-834.

Herculano-Houzel, S. (2002) Do You Know Your Brain? A Survey on Public Neuroscience Literacy at the Closing of the Decade of the Brain. *Neuroscientist*, 8(2), 98–110.

Hildt, E. (2012) *Neuroethik*. München/Basel: Ernst Reinhardt Verlag.

Holbrook, J., & Rannikmäe, M. (2009) The Meaning of Scientific Literacy. *International Journal of Environmental and Science Education*, 4(3), 275-288.

Hößle, C. (2001) *Moralische Urteilsfähigkeit. Eine Interventionsstudie zur moralischen Urteilsfähigkeit von Schülern zum Thema Gentechnik*. Innsbruck: Studienverlag.

Jaworski, A., Coupland, N. (1999) *The discourse reader*. Routledge.

Jones, M. G., & Carter, G. (2007) Science teacher attitudes and beliefs. In S.K. Abell & N.G. Lederman

(Eds.), *Handbook of research on science education* (pp.1067-1104). London: LEA Publishers.

Kelly, G. J. (2007) Discourse in science classrooms. In S.K. Abell & N.G. Lederman (Eds.), *Handbook of research on science education* (pp. 443-469). London: LEA Publishers.

Kelly, G. J. (2014) Discourse practices in science learning and teaching. In S.K. Abell & N.G. Lederman (Eds.), *Handbook of research on science education*, 2, (pp. 321-336). London/New York: Routledge.

Lewis, J., & Kattmann, U. (2004) Traits, genes, particles and information: re visiting students' understandings of genetics. *International Journal of Science Education*, 26(2), 195-206.

Mercer, N. (2004) Sociocultural discourse analysis. *Journal of applied linguistics*, 1(2), 137-168.

Osborne, J. (2010) Arguing to Learn in Science: The Role of Collaborative, Critical Discourse. *Science*, 328, 463-466.

Pajares, M. F. (1992) Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332.

Przyborski, A. (2004) *Gesprächsanalyse und dokumentarische Methode: qualitative Auswertung von Gesprächen, Gruppendiskussionen und anderen Diskursen*. Berlin/Heidelberg: Springer.

Reitschert, K. (2009) *Ethisches Bewerten im Biologieunterricht. Eine qualitative Untersuchung zur Strukturierung und Ausdifferenzierung von Bewertungskompetenz in bioethischen Sachverhalten bei Schülern der Sekundarstufe I*. Hamburg: Verlag Dr. Kova .

Rokeach, M. (1968) *Beliefs, attitudes, and values: A theory of organization and change*. San Francisco: Jossey-Bass.

Roskies, A. (2002) Neuroethics for the new millennium. *Neuron*, 35(1), 21-23.

Simpson, R. D., Koballa, T. R., Oliver, J. S., & Crawley, F. E. (1994) Research on the affective dimension of science learning. In D.L. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 211-234). New York: Macmillan.

Walker, K. A., & Zeidler, D. L. (2007) Promoting discourse about socioscientific issues through scaffolded inquiry. *International Journal of Science Education*, 29(11), 1387-1410.

Zeidler D.L., & Nichols B.H. (2009) Socio-Scientific Issues: Theory and Practice. *Journal of Elementary Science Education*, 21(2), 49-58.

Zeidler D.L., Sadler T.D., Simmons M.L., & Howes E.V. (2005) *Beyond STS: A Research-Based Framework for Socioscientific Issues Education*. Wiley Periodicals, 357-368.

The impact of peer discussions on students' arguments when addressing socio-scientific issues in biotechnology

Birgitta Berne[1]

birgitta.berne@telia.com

Abstract

This study reports on an intervention in science education in which 14-15 year old students discussed different SSIs in peer groups. Prior to these discussions the students were provided with different perspectives of the issues and given time to search for further information themselves. The analysis of students' arguments suggests the students formed their initial arguments during their preparations for the discussions. Furthermore, throughout the peer discussions they elaborated on the initial arguments and supported each other to intertwine ethical considerations or content knowledge with the arguments, or to revise their arguments. Thus, the students seemed to focus on diverse aspects of the different issues. When discussing a general DNA-register they focused on the virtue ethics aspects, when discussing GMOs they focused on longer term consequences, and when discussing issues in connection to stem cells they focused on the content knowledge. Moreover, the recurrent peer discussions appeared to support the students to advance in their argumentation process in that they responded to claims by questioning, evaluating and revising each other's claims. As such, the study provides valuable indications concerning the importance of introducing prepared peer-discussions about different socio-scientific issues into the teaching of science in schools.

Keywords: biotechnology, learning progression, scientific literate, socio-scientific issues, teacher researcher

---

[1] Lindåsskolan, Billdal, Sweden

1. Introduction

One of the aims for science education in Sweden is that students should become informed reflective citizens able to use scientific content knowledge, as well as ethical aspects, as part of an argument when taking a stand (The National Agency of Education – Skolverket, 2000; 2011). In other words, the students should become scientifically literate (SL). Citizens of today have to make personal and ethical decisions about a range of scientific issues based on information available through the press and other media (Osborne et al., 2004). Hence, to be able to make decisions about, for example genetic screening, stem cell research and genetically modified foods, students need to understand the core ideas in genetics (Duncan et al., 2009).

In order to promote SL for citizenship the international literature suggests that students benefit from having the opportunity to discuss controversial socio-scientific issues (SSIs) (e.g. Dawson, 2010; Jones et al., 2007; Sadler, 2011; Simonneaux, 2013). SSIs handle content-transcending topics such as science as a social process, limitations of science and values in science (Kolstø, 2001). These issues are in line with the ethical and social aspects of science emphasised in the Swedish curriculum (Skolverket, 2000, 2011), since they integrate science education and character education. Hence, as argued by Zeidler & Keefer (2003), for being able to make informed decisions students have to understand the connections inherent in socio-scientific issues. Citizens and future citizens need to be able to understand the issues and the implications of the decisions made by individuals and agencies, but they also need to participate and influence these decisions and, as such, this need has implications for science education (Berkowitz & Simmons, 2003). However, as expressed by Zeidler et al., (2002) if students are expected to make rational and informed decisions about a science and technology permeated society, they must have had the opportunity to experience decision- making.

In their review of the state of science education in Europe, Osborne & Dillon (2008) found little work that involved students constructing an argument. They reported that contemporary school science education in Europe offered few opportunities for a pedagogical approach centred on deep understanding that requires space for students to discuss, to think critically and to consider others' points of view. Studies that focus on the changes that individuals and groups undergo when they have the possibility to use argumentation for the construction of knowledge are also rare (Evagorou & Osborne, 2013). Thus, as

made explicit by Evagorou & Osborne (2013), there is still a need to use "an SSI context to explore the characteristics of collaborative argumentation, and explore the impact of the process on the product" (p. 214).

Evagorou & Osborne's study (2013) focused on two dyads of students discussing one specific SSI. Despite their study, a gap remains to be explored concerning what happens if students get the possibility to have recurrent discussions handling different SSIs within a lesson period. In response to such gaps in the evidence a study was designed in accordance with the recommendations from Sadler (2011), in which empirical evidence from classroom-based studies of SSI implementation are seen as specifically relevant to such evidential gaps. The research questions that guided the current study was:

What kind of progression in the argumentation process can be found during students' recurrent discussions about different SSIs?

2. Methodology

The curriculum context for this action research study was a science module in genetics and biotechnology in which the first goal was that students should gain scientific content knowledge in order to justify claims with appropriate evidence and reasoning (Berland & McNeill, 2010). Secondly the goal was to change the classroom norm, which was that just two or sometimes three students were taking part in 'whole class discussions'. The intention was to make the classroom talk into a more dialogic discourse where ideas were not just brought into the classroom, but all ideas where questioned and evaluated, as recommended by Ford (2008).

In order to identify the progression in students argumentation process an intervention was designed based on previous research on argumentation, SSI and science education (e.g. Jiménez-Aleixandre & Erduran, 2008; Jorde, 2007; Sadler & Zeidler, 2005). As different students might find various SSIs interesting, a single discussion might not be enough for making conclusions concerning students' argumentation. Furthermore, as stressed by McNeill and Krajcik (2009), recurrent discussions are critical for determining whether, and how, students engage in argumentation changes. Hence the teaching intervention drew upon contemporary discussions about different topics of biotechnology: DNA-registers, genetically modified organisms (GMOs) and the growing of stem cells (including the issues of cloning and 'designer babies').

2.1 Participants

All 20 students in the class, 9 boys and 11 girls in grade 9 (14–15 years of age) took part in the study. They attended a small school located on the west coast of Sweden where social and economic conditions are relatively favourable. Covering all students was of importance to the study, since the goal was to explore the possibilities of collaborative argumentation in discussions among students interacting in an everyday classroom situation.

2.2 Theoretical lens

The theoretical lens for this study was that learning involves a passage from the social context to individual understanding, which means that new ideas must be rehearsed between people and talked through (Vygotsky, 1935/1978). Hence the teaching was based on students discussing SSIs in connection to biotechnology in peer groups.

2.3 The intervention

To start, the students were taught genetics. Six month later peer discussions were arranged during science teaching periods in the students' classroom. The issues were delivered one after the other, one issue a week. First the students were presented with different views of the issues in teacher-led lessons, online programmes, movies, articles or books. As homework the students were to read short texts in their biology textbook and more advanced texts in a book by Brändén (2002). Furthermore, the students were asked to search for more information from the Internet and to write down their own arguments for, or against, the issue at hand, before discussing with their peers. Two days later, the students met their peers for discussions. They were given a catalyst to start the discussion and were divided into groups of 5, in total 4 groups. To ensure that there would be talk in all the groups, talkative and quiet students were placed in each group, and so were boys and girls. The students themselves recorded the discussions. My earlier experience, from a teacher's perspective as well as from my pilot study, was that students could capture situations you as a teacher or a researcher would never be invited to engage in.

    After the first discussion the students were given explicit instructions on how to construct a good argument according to the Toulmin (1958) model, they were instructed to connect data to claims and to use warrants, backings, qualifiers and rebuttals. As pointed out by for example Duschl & Osborne

(2002), Dori et al. (2003), and Jiménez-Aleixandre & Erduran (2008), the students benefit from getting explicit practise in argumentation, for the students to gain content knowledge at the same time as they develop a strategy for how to construct an argument.

After the second discussion the students were informed about how the Toulmin's model could be extended to capture all the aspects of the Swedish grading system. Thus students were encouraged to connect *scientific* data to claims and they were also encouraged to examine the person posing an argument and analyse the values and purposes that might underpin such arguments. Furthermore, students were informed to invite and listen to each other and to show respect for the arguments of others, as prescribed in the National Agency of Education (Skolverket, 2000, 2011). Directly after the group discussions, the students wrote an individual reflection on whether they had come up with new arguments or perhaps changed their views after the discussion. Half a year after the discussions the students wrote individual arguments for or against the use of biotechnology.

2.4 Data collection

Data was collected from all the students' individually written arguments from before and after each of the peer discussions and from their arguments six months after the intervention. Furthermore all the groups' video recordings were collected and analysed.

2.5 Data analyses

In order to examine what the students focused in their collaborative argumentation for the different issues in the teaching intervention all the groups' video-recordings were inductively analysed. In a previous study (Berne, 2014) where students' talk was categorized according to Mercer's (1995) categories, it was found that students preferably using cumulative talk did not advance in as many of the categories of ethical reasoning than the groups using preferably disputational or explorative talk. Hence, the present study explores what structures of collaborative argumentation that could take place in such a group. Students' arguments that could not be directly traced to the information given in the texts for their preparation were transcribed and analysed.

In order to analyse the structure of the students' argumentation and to be able to capture a possible progression in the students' argumentation, the

model for learning progression described by Berland & McNeill (2010) was explored. In this model 'responses to claims' are seen as more complex if students, besides articulating and defending claims also question, evaluate and revise them. These structures were originally described for analysing an argumentative discourse. However, the intention for the present study was to explore the argumentation in a group preferably using cumulative talk and to find out whether these structures could also add substance to the feature 'elaborate' from Mercer's (1995) description of cumulative talk. Hence, the model was modified to describe an argumentation discourse instead of an argumentative discourse. However, the description of 'students participation' from the original model could still be used for a cumulative group. According to the model, an argumentation is viewed as more complex if the students spontaneously engage in the argumentation discourse rather than if students' participation is prompted by their teacher (Figure 1).

Dimension of the argumentation process in the characteristics	Simple ——————————————► Complex		
Responses to claims	Claims are articulated, defended questioned OR evaluated	Claims are articulated, defended, questioned, AND evaluated	Claims are articulated defended, questioned, evaluated and revised
Students' participation	Student participating in argumentation discourse is prompted by their teacher	Teacher and students share responsibility for prompting the argument	Students spontaneously engage in argumentation discourse.

Figure 1 The dimension of the argumentation process for learning progression adopted from Berland & McNeill (2010) and modified.

To illustrate progression students' claims are marked directly in the excerpts: questioned (Q), evaluated (E) and revised (R). The claims articulated and defended are not marked since these responses are included in all dimensions. The coding is in accordance with the analysis from two scholars in science didactics and shown directly in the data section below. The students spoke in Swedish and the translation into English is as close to the students' words as possible with only small corrections to enable comprehension. The students' names are pseudonyms.

3. Findings

In the analysis of all the students' peer discussions it was found that the groups focused on diverse aspects of ethical reasoning for the different issues. When they discussed a general DNA-register the groups focused on the virtue ethics/rights and duties aspects. When the groups discussed GMO they focused on longer term/societal consequences and when they discussed stem cell research they focused on scientific content knowledge. Even if the students were informed about different perspectives of the issues beforehand and got catalysts to start the discussions, they were not prompted during the discussions, but spontaneously engaged in the argumentation discourse. Hence all groups reached a complex level in the characteristics of 'students' participation' for all the recurrent discussions (Figure 1). However, the students' argumentation process in the characteristics 'responses to claims' (Figure 1), analysed for the cumulative group, became more complex during the recurrent discussions. The short excerpts that are presented below just capture very restricted parts of this activity for the cumulative group.

3.1 DNA-register

Whether we should have a general DNA-register or not was the first issue for the students in this intervention to discuss. The students had got time for preparation and immediately started to discuss who should be in a future DNA-register. They also questioned the fact that the already existing PKU-register was opened to solve the murder of Anna Lindh, our previous foreign minister, demonstrating a virtue ethics approach to this issue.

Elin: But I think it was while damn that they went into this register only when Anna Lindh was murdered. (Q)
Josefine: Mm, why just Anna Lindh? (Q)
William: Yes.
Josefine: It makes me angry.
Elin: She is worth the same as everyone else! (Q)
/…/
Kajsa: It is no equality at all! (Q)
Josefine: No, I agree.
/…/

William: Politicians! That she apparently is more important than most others! (Q)

Kajsa: It is not the way it should be! (Q)

Elin: No.

/…/

Josefine: Thus, this murder of Olof Palme, they have not found anyone yet,

Elin: No

Josefine: supposedly they were afraid that it should be the same with Anna Lindh. (E)

William: I can understand a bit of that.

Josefine: You could understand a bit of that.

Elin: Why?

Josefine: Because that was big! But come on, we are worth as much as she! (Q)

Elin: Yes, I think so too, we really are. She should not get any extra! (Q)

Kajsa: No, really not!

This excerpt demonstrates how the students, in their first discussion, mostly add short claims to each other's claims. The students are upset that the PKU- register was opened to solve the murder of Anna Lindh, and not for anyone else and they question the equality. Even if Josefine evaluates the reason behind opening the register this is the only evaluation made in this section. Thus the argumentation process could be described as quite simple in the characteristic 'response to claims' (Figure 1).

3.2 GMO

The second issue was whether we should grow GMO or not. Before the discussion the students had been informed about how to construct a good argument according to the Toulmin (1958) model and the students had been requested to connect data to their claims. To start the discussion the students discussed possible negative longer term consequences for the ecosystem from genetically modified tomatoes, as they do not get soft when they ripe. They also discussed possible positive consequences, as these tomatoes could be easier to transport. At this point Elin makes a suggesting for taking action and not eating tomatoes in winter.

Elin: What I mean is, you do not have to eat tomatoes in winter! (Q)

Josefine: Yes, you mean that you could conserve the tomatoes? (E)

Elin: You could grow them in Spain, because Spain is not that far away, you could use these tomatoes in summer.

Josefine: You mean that it is more environmentally friendly? (E)

William: Mm

Josefine: I agree.

William: However, it will also become more expensive because it costs money. (E)

Josefine: It becomes cheaper! (E)

Kajsa: No, Swedish tomatoes in summer and no tomatoes at all in winter!

Elin: Yes, this is the way I thought.

Josefine: What? What did you say now?

Elin: Thus?

Josefine: No, but what you mean is that if tomatoes could stand the cold much better, then you could grow them also in winter. (E) Is this what you mean?

Elin: Thus, you do not have to eat tomatoes and such things that come from tropical countries, or do they not…? (Q)

/…/

Elin: Thus, do you have to eat it in winter then? (Q)

Josefine: I do not really understand what you mean.

Kajsa: You should eat Swedish tomatoes in summer and no tomatoes at all in winter, not even Spanish!

/…/

Josefine: Yes, but there are people who want to make dishes with tomatoes even in winter. (E)

Elin: But, I do not give a shit, could they not skip that?

In this excerpt the students put more effort than in the first discussion explaining their claims. Elin, with the help from Kajsa evaluates if it is necessary to eat tomatoes in winter. She argues that if we need to modify tomatoes for being able to transport them, we could instead stop eating tomatoes. However, Josefine does not really understand Elins' claim and asks questions for clarification and she, as well as William, try to evaluate the claim posed by Elin. Hence

the argumentation process in this discussion could be described as in between simple and complex in the characteristic 'response to claims' (Figure 1).

3.3 Stem cells

The last discussion was whether we should grow stem cells or not. Now the students had also been informed to use *scientific* content knowledge as data and it is clear that the students put more focus than in the previous discussions on content knowledge relevant to the issue. However when the students realize that stem cells come from embryos, the discussion focus on societal or longer term effects of stem cell research on our society.

> *Kajsa:* To grow stem cells for making different organs, thus develop a liver from a stem cell, that I think would be very good, because then you do not need donors and you save surgeries. (E)
>
> *William:* Money! (E)
>
> *Kajsa:* Yes, money as well and you do not need to have two surgeries, you just need one. (E)
>
> /…/
>
> *William*: Stem cells are more natural. (E)
>
> *Josefine*: But, I got to think about, where to get stem cells from?
>
> /…/
>
> *Kajsa:* from, from a human being.
>
> *Josefine*: But, is it not just embryos, thus embryos that have got stem cells? Is it not hard to find stem cells in a grown up person? (Q)
>
> *William:* Yes
>
> *Josefine*: Are there not very, very few? (Q)
>
> *Kajsa:* Do you not take them when you are born?(Q)
>
> *William:* From new-borns.
>
> *Kajsa:* Thus, you can take someone and make stem-cells, however you could not do like that…(E)
>
> *William*: If you take stem cells from certain persons, then I think it could be rather offensive. However if you take stem cells from all people in Sweden, or from the whole world, then it is nothing special. (E)
>
> *Josefine:* However, if you for example take stem cells from an embryo and save those stem cells, that is similar to… those stem cells should become a human being! (E)

/…/

Elin: My point of view concerning stem cells is that you should be allowed to use them, because you cold repair parts of a injured body and that is good, however, at the same time I think that if we repair, we are so incredibly many people on earth now, so if you repair all injures that are all over the world then I think we will become too many at the end. (E) (R)

William: Yes

Elin: Thus I do not think we should always cure all illness, I think the swine flu for example, actually I think some are meant to die. (E) (R)

This excerpt shows that the students, at the end of this discussion, spend even more time than in the previous discussion explaining their claims. From the beginning the students are very positive to stem cell research. However, after the students have discussed where to find stem cells and realized that they preferably are found in embryos, they start questioning the need of stem cell research. Furthermore, Elin starts to revise the previous claims. Hence the students question, evaluate and revise claims and their argumentation process in the characteristic 'response to claims' could be described as complex (Figure 1).

4. Discussion

Although the benefits from SSI studies have been reported earlier, most studies are from students discussing one specific issue (e.g. Evagorou & Osborne, 2013). Hence the contribution of the present study is to show, from empirical data, how recurrent discussions might be critical for the students to fully discuss claims. The students in the present study did not question, evaluate and revise claims until the third discussion. Furthermore the students, in the aforementioned studies, did not prepare before discussing with peers, hence the contribution of the present study is also to show how critical the students' preparations appear to be, in order for the students to make more reflected arguments. The preparations seemed to help the students to get to know what data could be appropriate and what could be viewed as an acceptable motivation for or against the issue at hand (Sampson & Clark, 2008). It seemed as the students brought their initial arguments into the peer discussion. In the discussion their arguments were questioned, evaluated or revised, with the result that the students could intertwine content knowledge or ethical consid-

erations into their initial arguments or revise their arguments.

In the present study it also seemed important that the students were explicitly taught argumentation (Osborne et al., 2004) and scientific content knowledge (Berland & McNeill, 2010) prior to the peer discussions. When the students had been informed about the Toulmin (1958) model, they started to evaluate claims, when they had been informed to use their *scientific* content knowledge as data for claims they searched for content knowledge.

The students from the first peer discussion spontaneously engaged in the argumentation discourse. This finding is in line with the findings from Zohar & Nemet (2002) who argued that students' implicit argumentation skills often are ignored at school, and hence even a short exposure to argumentation may reinforce skills already present. Moreover it seemed that students benefit from discussing different SSIs, not just for the sake of interest and to be able to fully discuss claims, but also for the students to experience different approaches to SSIs.

Argumentation, especially collaborative argumentation, suggests that students should benefit from being critical to the arguments expressed by themselves or by their peers (Skolverket, 2000, 2011). However, for students using preferable cumulative talk, recurrent peer discussions seemed to be crucial for the students to fully engage in the discussion. These findings are similar to the findings from Zohar & Nemet (2002) describing an increase in the complexity of the students' discourse during their intervention. In their study they described an increase in the frequency of explicit conclusions and an increase in the number of ideas units per conversational turn. This increase of complexity is similar to what is found in the present study where students' discourse in argumentation can be described as students moving from questioning claims to also evaluating or revising claims. When claims are questioned and evaluated students gain the opportunity to connect data to their claims (Toulmin, 1958), advance in their ethical reasoning (Jones et al, 2007) and view the issues from different perspectives (Sadler et al., 2007). However, even if the students in my study revised their claims I question if revising claims is mandatory for viewing a process in argumentation as complex. This, I believe, needs further study.

Since the sample in the study is limited to only one class, no generalisations can be made. However, the study can highlight the role that recurrent peer-to-peer discussions play in enabling students to gain the skills that define

an informed and reflective scientifically literate citizenry. Furthermore, the study can show how a teacher researcher is not just an ideal position for gathering data (Roth, 2007) it is an ideal position for implementing research insights into the classroom.

References

Berkowitz, M. W., & Simmons, P. (2003) Integrating Science Education and Character Education. In D. Zeidler (Ed.), *The role of moral reasoning and discourse on socioscientific issues in science education,* Dordrecht: Kluwer Academic Publishers, 117-138.

Berland, L. K., & McNeill, K. L. (2010) A Learning Progression for Scientific Argumentation: Understanding Student Work and Designing Supportive Instructional Contexts. *Published online 4 May 2010 in Wiley Online Library (wileyinlinelibrary.com).*

Berne, B. (2014) Progression in Ethical Reasoning When Addressing Socioscientific Issues in Biotechnology. *International Journal of Science Education*, *36*(17-18), 2958-2977.

Brändén, H. (2002) *Genteknik, kloning och stamceller*. (Gene technology, cloning and stem cells) Stockholm: Ingenjörsvetenskapsakad. (IVA).

Dawson, V. (2010) Outcomes of Bioethics Education in Secondary School Science: Two Australien Case Studies. In Jones, A., McKim, A. & Reiss, M. (Eds) *Ethics in the Science and Technology Classroom. A New Approach to Teaching and Learning.* Rotterdam/Boston/Taipei: Sense Publishers, 68-86.

Dori, Y. J., Tal, R. T., & Tsaushu, M. (2003) Teaching biotechnology through case studies - can we improve higher order thinking skills of nonscience majors? *Science Education*, 87: 767-793.

Duncan, R. A., Rogat, A. et al., & Yarden, A. (2009) A learning progression for deepening students' understandings of genetics across the 5th-10th grades. *Journal of Research in Science Teaching, 46*(6), 655-674.

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

Evagorou, M., & Osborne, J. (2013) Exploring Young Students' Collaborative Argumentation Within a Socioscientific Issue. *Journal of Research in Science Teaching, 50* (2), 209-237.

Ford, M. (2008) Disciplinary authority and accountability in scientific practice and learning. *Science Education*, *92*(3), 404-423.

Jimenez-Aleixandre, M., P., & Erduran, S. (2008) Argumentation in Science Education: An Overview. In S. Erduran, & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in Science Education: Perspectives from Classroom-Based Research*. Dordrecht: Springer 3-25.

Jones, A., Mckim, A., Reiss, M., Ryan, B., Buntting, C., Saunders, K., et al. (2007) *Research and development of classroom-based resources for bioethics education in New Zealand.* Hamilton, NZ: Wilf Malcolm Institute of Educational Research, School of education, University of Waikato.

Jorde, D. (2007) Promoting science inquiry – new possibilities using ICT. *Conference-paper presented at Linnaeus Tercentenary 2007, Promoting Scientific Literacy: Science Education Research in Transaction- LSL Symposium Uppsala, May 28-29, 2007.*

Kolstø, S.D. (2001) Scientific literacy for citizenship: Tools for dealing with the science dimension of controversial socioscientific issues. *Science Education, 85*(3), 281-310.

McNeill, K. L., & Krajcik, J. (2009) Synergy between practices and curricular scaffolds to support

students in using domain specific and domain general knowledge in writing arguments to explain phenomena. *Journal of Learning Sciences*, *18*(3), 416-460

Mercer, N. (1995) *The Guided Construction of Knowledge. Talk amongst Teachers and Learners.* Philadelphia: Multilingual Matters LTD.

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.

Osborne, J., & Dillon, J (2008) Science Education in Europe: Critical Reflections. London. King´s College. *A report to the Nuffield Foundation.*

Ottander, C., & Ekborg, M. (2012) Students' experience of working with Socio Scientific Issues – a Quantitative Study in Secondary School. *Research in Science Education, 42*: 1147-1163.

Roth, W-M. (2007) *Doing teacher-research a handbook for perplexed practitioners*. Rotterdam: Sense Publishers.

Sadler, T. D. (2011) Situating Socio-scientific Issues in Classrooms as a Means of Achieving Goals of Science Education. In: T. D. Sadler (Ed.), *Socio-scientific Issues in the Classroom - Teaching, Learning and Research*. Dordrecht Heidelberg London New York: Springer, 1-10.

Sadler, T. D., Barab, S. A, & Scott, B. (2007) What Do Students Gain by Engaging in Socioscientific Inquiry? *Research in Science Education*, 37: 371-391.

Sadler, T. D., & Zeidler, D. L. (2005) The significance of content knowledge for informal reasoning regarding socioscientific issues: Applying genetics knowledge to genetics engineering issues. *Science Education*, 89, 71-93.

Sampson, V., & Clark, D. B. (2008) Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education,* 92(3), 447-472.

Simonneaux, L. (2013) Questions socialement vives and socioscientific issues : New trends of research to meet the training needs of post-modern society. In C.Bruguière, A.Tiberghien & P. Clément (eds). *9th ESERA Conference Selected Contributions. Topics and trends in current science education*. Dordrecht: Springer, 37-54.

Skolverket (The National Agency for Education) (2000) *Natural Science Programme. Programme Goal, Structure and Syllabuses.* Retrieved July 20, 2009, from: http://www3.skolverket.se/ki/eng/nv_eng.pdf

Skolverket (The National Agency for Education) (2011) *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011.* Stockholm: Skolverket

Toulmin, S. (1958) *The uses of argument*. Cambridge: Cambridge University Press.

Vygotsky, L., S. (1935/1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Zeidler, D. L., & Keefer, M. (2003) The Role of Moral Reasoning and the Status of Socioscientific Issues in Science Education. In D. Zeidler (Ed.), *The role of moral reasoning and discourse on socioscientific issues and discourse in science education* (pp. 7-38) Dordrecht, The Netherlands: Kluwer.

Zeidler, D.L., Walker, K.A., Ackett, W.A., & Simmons, M.L. (2002) Tangled up in views: Beliefs in the nature of science and responses to socioscientific dilemmas. *Science Education, 86*(3), 343-367.

Zohar, A., & Nemet, F. (2002) Fostering students´ knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching,* 39, 35-62.

Young children's reasoning when sorting pictures and objects

Marida Ergazaki[1], Renia Gasparatou[1] and Eftychia Valanidou[1]

ergazaki@upatras.gr

Abstract

This paper reports on a mixed-model case study that explores young children's ability to spontaneously categorize biological and non-biological entities presented to them as pictures or 3-dimensional objects. Our focus is set on the sorting criteria children activate depending on the stimulus medium and whether these criteria differ in terms of concerning (a) properties inherent to the entities *(intra*-entity criteria), or (b) associations emerging in the entities' interactions *(inter*-entity criteria). Conducting individual, semi-structured interviews with 120 preschoolers (age 5-5.5) at public kindergartens in the Patras area of Greece and analyzing their responses, we found that the frequency of intra- and inter-entity criteria did not differ much in the two sorting tasks. Children appeared to slightly prefer intra- rather than inter-entity criteria, regardless of whether they were sorting pictures or objects. Moreover, intra-entity criteria were slightly more frequent in the pictures-task rather than in the objects-task, while inter-entity criteria were almost equally activated in both. None of these differences was found to be statistically significant. Thus, both objects and pictures could be used within introductory learning environments for 5 year olds. Biology instruction could build, without medium-restrictions, on children's emerging potential to activate intra-entity relations, in order to help them enhance their ability to categorize taxonomically.

Keywords: biology education, categorization criteria, preschoolers, reasoning skills, stimulus-medium

---

[1] University of Patras, Patras, Greece

1. Introduction

Categorization is a basic thinking skill. People seem to categorize in many different ways, which are all useful for their lives (Murphy, 2002). The teaching of biology however, demands that children shift to taxonomic criteria (Ergazaki et al., 2016) and learn to form concepts like *animals* or *plants*. Research has shown that children of different ages hold alternative ideas regarding such concepts (Allen, 2015; Chen & Ku, 1998; Trowbridge & Mintzes, 1985; Trowbridge & Mintzes, 1988; Wright et al., 2015; Yen et al., 2004, 2007; Yorek et al., 2009); and these ideas seem to be rather resistant. For instance, Kattmann (2000; 2001) suggested that primary school students prefer to classify living organisms by appealing to their habitat and locomotion, even if taught about biological taxonomy. Thus, it seems reasonable to study what kinds of factors may facilitate or hinder taxonomic categorization.

Research suggests that children use different types of sorting criteria. According to Ware et al. (2013), children's criteria may be (a) *taxonomic*, (b) *shared-properties*, (c) *thematic*, or (d) *slot-fillers.* In other words, children may put entities in the same group, because they think that (a) they belong to the same kind (e.g. they are all animals), (b) they have common features (e.g. they are all of the same shape), (c) they interact in the real world in causal, functional, spatial, temporal or other ways (e.g. they are all 'school stuff'), and (d) they fill the same role in a familiar script or a schema, like for instance a bowl and some milk and cereals do in the breakfast script that children experience every day.

Having a closer look at the sorting criteria of Ware et al. (2013), we realized that the first two types (*taxonomic* & *shared properties*) may be characterized as *intra*-entity criteria, since they concern properties that are *inherent* to the entities. When appealing to the ontological status of the entities or their properties, children actually set their focus on each entity itself and compare it to the others. On the other hand, the last two types of Ware et al.'s criteria (*thematic* & *slot-fillers*) may be characterized as *inter*-entity ones, since they concern associations that emerge in the interactions *between* the entities. When appealing to the ways through which the entities can interact in the real world, possibly in the context of well-known rituals of everyday life, children do not set their focus on the entities themselves, but rather on what may happen between them.

Categorization performance may be influenced by several factors. The type of sorting task (e.g. free-sorting, match-to-sample, sorting-with-parent), the

instructions coming with it (e.g. the use of labels / names for the entities, the place where sorting is required to be performed, for instance a table or a transparent bag), and finally the way that the entities in question are represented to the participants (e.g. as pictures or objects) - the so-called stimulus medium - are some of them (Deák & Bauer, 1996; Gelman et al., 2005; Markman et al., 1981; Markman & Hutchinson, 1984; Ware et al., 2013; Waxman & Namy, 1997).

Regarding the stimulus medium in particular, it has been suggested that pictures may be more easily interpreted as representations of *kinds*, whereas objects may be more easily interpreted as representations of *individual* entities (Gelman et al., 2005; Gelman et al., 2008; Ware et al., 2013). According to Gelman et al. (2005), the child-mother discussions about objects in their study were more likely to focus on individuals (e.g. a specific 'Mr. Frog'), whereas the discussions about pictures were more likely to focus on kinds (e.g. 'frogs' in general). Moreover, according to Ware et al. (2013), the child-mother discussions about objects were more likely to draw upon thematic relations (e.g. 'a frog eating ice cream'), whereas the discussions about pictures were more likely to draw upon taxonomic relations (e.g. 'frogs are animals') and shared-properties (e.g. 'they have the same colour').

These findings were derived from studies with 2-5 year olds, but what is the situation with a slightly older preschoolers, who are actually more likely to take part in teaching interventions that aim at category building? The present study explores 5-5.5 year olds' ability to spontaneously categorize biological and non-biological entities presented to them as *pictures* or *3-dimensional objects*. Our focus is set on the criteria children activate depending on the stimulus medium and whether these differ particularly in terms of being *intra-* or *inter*-entity ones. Thus, the research questions we address here are the following:

- What types of criteria do preschoolers use to categorize spontaneously biological and non-biological entities presented to them as pictures?

- What types of criteria do preschoolers use to categorize spontaneously biological and non-biological entities presented to them as objects?, and

- Do picture-criteria & object-criteria differ in terms of being *intra-* or *inter*-entity?

2. Methods

2.1 Overview of the study

This is a mixed-model case study. The participants were 120 preschoolers (60 girls / 60 boys, age 5-5.5), attending four public kindergartens in the area of Patras, Greece. The schools were situated in semi-urban areas of medium / high socio-economic status and were selected due to the teachers' wish to facilitate our study. The children were already familiar with educational interactions, since they had been attending kindergarten for several months. Nevertheless, according to their teachers they were not engaged with *formal* categorization activities up to that point. They also had the opportunity to get familiar with the interviewer and give their own assent for taking part. The parents were informed about the study at the beginning, in order to confirm that they did not have any objections to their children's participation. Tracing young children's reasoning was performed through individual, semi-structured interviews; these were taken in quiet places at the schools and lasted 15-20 min each.

2.2 The interview protocol

Children were asked to complete two free-sorting tasks: the objects-task and the pictures-task. The tasks were modified versions of Gelman et al. (2005) and Ware et al. (2013). In both tasks, children had to (a) recognize what was represented by each picture or object, (b) create groups with them, and (c) provide justifications for the groups they created. The interview questions were along the lines of: *'Let's see what we have here. What do the pictures/objects depict? Can you make groups with them? Why did you make these groups? What do the pictures/objects in this group have in common?'* We used 24 items, in the form of 3-D toy-objects *and* in the form of pictures. Pictures were photographs of the objects in a sketch format, and they looked similar to the objects but without a 3-D effect. The items were divided into two sets, each with 4 animals, 4 plants and 4 artifacts of 4 different colours (Table 1).

	Colour	Animals	Plants	Artifacts
Set 1	Blue	Bird	Tree (with blue flowers)	Bag
	Brown	Rabbit	Tree (no leaves)	Box
	Yellow	Turtle	Sunflower plant	Chair
	Orange	Fox	Flowers	Shovel
Set 2	Green	Bird	Tree (green leaves)	Cup
	Beige	Dog	Tree (no leaves)	Box
	Pink	Flamingo	Flowers	Sunglasses
	Red	Ladybird	Tree (red leaves)	Stool

Each set had two versions: a pictures-version for the pictures-task, and an objects-version for the objects-task. When the pictures-task was performed using Set 1, the objects-task was performed using Set 2 (Figure 1). This was the case for half of the children.

Figure 1 Sets' combination (I).

The other half of the children did the reverse: they performed the objects-task using Set 1 and the pictures-task using Set 2 (Figure 2).

Figure 2 Sets' combination (II).

Moreover, the task-order was counterbalanced: half of the children were first given the pictures-task and the other half the objects-task.

2.3 The analytic procedure

The audio-recorded interviews were transcribed and prepared for coding in 'NVivo', a computer-supported environment for the analysis of qualitative data (Gibbs, 2005). Children's sorting criteria were identified in the interviews and coded into a series of main categories or types (Figure 3) by two authors (Cohen's kappa=0.90). The differences in the extent of these categories' use in the pictures-task and objects-task were statistically tested.
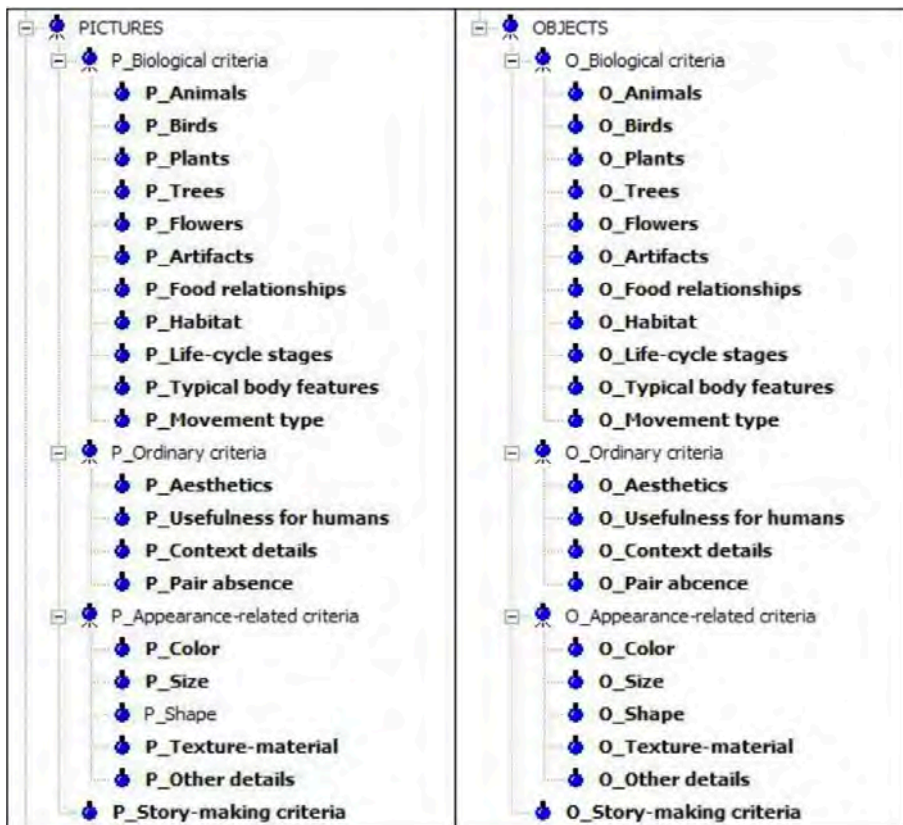
Figure 3 The coding scheme.

We then focused on whether the main categories of our coding scheme concerned *taxonomic*, *shared-properties* or *thematic* relations. In other words, we ascertained whether our main categories had to do with (a) properties that were inherent to the entities, or (b) associations emerging in their interactions. Based on this distinction, we finally characterized them as *intra-* or *inter*-entity, respectively. The differences in the extent of the use of *intra-* and *inter*-entity criteria in the pictures-task and objects-task were also statistically tested.

3. Findings

3.1 Children's criteria when sorting pictures or objects
Children formed groups of pictures or objects, appealing to four main types

of criteria: (a) 'biological', (b) 'ordinary', (c) 'appearance-related', and (d) 'story-making'.

The **biological criteria** involve ideas about the biological world. In both the pictures-task and objects-task, children put entities in the same group because they are all 'animals', 'plants', 'birds', trees', 'flowers' or 'artifacts'. In children's words:

- 'animals': *'Turtle, rabbit, bird and fox because they are all animals'* (pictures-task); *'Ladybird, flamingo and bird: this is the group of animals'* (objects-task)

- 'birds': *'Flamingo and bird, because they are both birds'* (pictures-task); *'I put flamingo and bird together because they are birds: they go together in the same group even though they have different colour'* (objects-task)

- 'plants': *'Sunflower, tree, tree with blue flowers and orange flowers: I put them in the same group because they are all plants … together'* (pictures-task); *'Pink flowers, tree without leaves, tree with red leaves and tree with green leaves: together because they are plants'* (objects-task)

- 'trees': *'Brown tree and tree with blue flowers: this is the group of trees'* (pictures-task); *'Tree with green leaves, tree with red leaves and tree without leaves: together because they are trees'* (objects-task)

- 'flowers': *'Yellow sunflower, orange flowers together: they're all flowers'* (pictures-task); *'Sunflower and orange flowers: this is the group of flowers'* (objects-task)

- 'artifacts': *'Chair, box, bag, shovel because they are things'* (pictures-task); *'Sunglasses and cup: together because they are both things'* (objects-task).

They also put entities together because they interact in the context of a 'food relationship' or a 'habitat', or they are part of a plant's life cycle. In children's words:

- 'food relationships': *'Sunflower and bird: together because the bird eats the leaves of this flower'* (pictures-task); *'They go together because the turtle eats the plants'* (objects-task)

- 'habitat': *'These go together because they live in the forest'* [turtle, bird, rabbit, fox] (pictures-task); *'Tree without leaves and bird: together because the bird makes its nest on the tree… the bird lives there'* (objects-task)

- 'life-cycle stages': *'The two trees cause the tree has leaves and then it drops them'* (pictures-task); *'Pink flowers, tree with green leaves, tree with red leaves and tree without leaves: I put these together because first it is a flower, then it has green leaves, then it has red leaves and in the end it is a naked tree'* (objects-task).
- Finally, they put entities together because they share some typical body features or a movement type. In children's words:
- 'typical body features': *'Pink flowers, tree with green leaves, tree without leaves & tree with red leaves: because they have roots'* (pictures-task); *'Flamingo, dog, ladybird, bird: together because they have legs'* (objects-task)
- 'movement type': *'I put these together because they can walk: the dog can walk on the grass, flamingo can walk in the water and the bird and ladybird can walk, too'* (pictures-task); *'I put these together because they can both fly: flamingo & bird'* (objects-task)

The **ordinary criteria** involve ideas encountered in everyday life. In both the pictures-task and objects-task, children put entities in the same group because they look good together ('aesthetics'), they are all useful to us ('usefulness for humans'), they can be found in the same spatial or temporal context ('context details'), and finally they do not match with anything else ('pair absence'). In children's own words:
- 'aesthetics': *'Rabbit and flowers: because they look nice together'* (pictures-task); *'Beige box, sunglasses, cup: because box looks nice with things in it'* (objects-task)
- 'usefulness for humans': *'Because the bag is like the box. We put things in the bag and we put things in the box'* (pictures-task); *'Because we use the sunglasses to protect ourselves from the sun, we use the cup for drinking water, we use the box to put things in it and we use the stool to sit'* (objects-task)
- 'context details': *'Chair and bag: when people go to the office they take their things in a bag and they sit on a chair to work'* (pictures-task)*; 'This tree which does not have leaves in the winter goes with the fox which gets out in the winter and with the shovel which we use to clean snow in the winter'* (pictures-task); *'I put the chair and the bag because this is the school bag and this is the school chair'* (objects-task); *'Box, cup, sunglasses and stool: I put these together because when we go for picnic we take all these with us'* (objects-task)
- 'pair absence': *'Bag, shovel, chair and box together, because they don't match with*

anything of the rest and they go together' (pictures-task); *'Stool, cup & sunglasses, because these are the last and they don't match with anything of the rest'* (objects-task).

The **appearance-related criteria** have to do with external, visible features of the entities. In both the pictures-task and objects-task, children put entities together because they are of the same 'colour', 'size', 'shape', 'texture / material' or they share 'other details'. In children's own words:

- 'colour': *'Sunglasses and flamingo: together because they're both pink'* (pictures-task); *'Bird and bag: together because they are both blue'* (objects-task)

- 'size': *'Bird and rabbit: because the two of them are small'* (pictures-task); *'Tree with red leaves and flowers: together because they are both big'* (objects-task)

- 'shape': *'Box and stool: because they are both the same shape, square'* (objects-task)

- 'texture / material': *'Cup and sunglasses look the same, they're glass'* (pictures-task); *'Box, stool and tree without leaves: they look like wood* (objects-task)

- 'other details': *'Stool, sunglasses and red with red leaves: together because they have some small lines here and there'* (pictures-task); *'Basket and bag: Because it is almost the same... i.e. the basket [brown box] has handles and the bag has handles'* (objects-task)

The **story-making criteria** involve ideas that allow entities to function as the building blocks of imaginary stories. In children's own words:

- *'Stool, ladybird, pink flowers, sunglasses, box and glass: because the little girl picks up the pink flowers and one of these flowers has a ladybird on it and then the girl goes to her house, gets in, seats on the stool to have some rest, drinks water with the cup, wears her sunglasses and gets out for another walk' walk'* (pictures-task)

- *'The three trees with the box: I have put these together because there was a hidden treasure in the box somewhere among the trees and some people went to look for it and they were digging and they found it in the forest and in the soil'* (pictures-task)

- *'This is a fox. She has found some flowers in the same color and she wants to give them to her husband'* [fox & flowers] (objects-task)

- *'The turtle goes on holiday and takes the bag with it'* [turtle & bag] (objects-task)

As shown in Table 2, the appearance of the four main types of criteria was very similar in both the pictures-task and the objects-task. The *biological criteria* were the most popular type in both. They were used (a) for the formation of 221 groups by 106/120 children in the pictures-task, and (b) for the formation of 226 groups by 101/120 children in the objects-task. The *ordinary* criteria were the second most popular type of criteria in both tasks. They were used (a) for the formation of 141 groups by 75/120 children in the pictures-task, and (b) for the formation of 123 groups by 74/120 children in the objects-task. The order of the last two types of criteria was reversed for the two tasks: in the pictures-task, the third type was the *appearance-related* criteria, used for the formation 72 groups by 31/120 children; whereas in the objects-task, it was the *story-making* criteria, used for the formation of 52 groups by 29/120 children. Thus, the criteria with the lower frequency were (a) in the pictures-task, the story-making criteria (used for the formation of 45 groups by 25/120 children), while (b) in the objects-task, the appearance-related criteria (used for the formation of 46 groups by 20/120 children).

Table 2 The frequency of use of the four main types of criteria in the two free-sorting tasks.

Types of criteria	Pictures-task	Objects-task
'Biological'	221 groups / 106 children	226 groups / 101 children
'Ordinary'	141 groups / 75 children	123 groups / 74 children
'Appearance-related'	72 groups / 31 children	46 groups / 20 children
'Story-making'	45 groups / 25 children	52 groups / 29 children

The differences of the four main types of criteria regarding the times they were used, or in other words the number of groups they gave rise to, were *not* found to be statistically significant ($\chi^2(3)=6.4$, $p\gg.05$).

3.2 Children's criteria when sorting pictures or objects: intra-entity or inter-entity ones?

In order to characterize the biological, ordinary, appearance-related and story-making criteria we identified children's responses as *intra*-entity or *inter*-entity, we had to take an intermediate step by defining which of them concerned *taxonomic, shared-properties* or *thematic* relations. This intermediate step along with the final one, are both summarized in Table 3 and are described below.

Starting with our *biological* criteria, we noted that some ('animals', 'plants',

'birds', trees', 'flowers', and 'objects') had to do with *taxonomic relations*, some ('movement type', 'typical body features') with *shared-properties*, and some others ('food relationships, 'habitat', and 'life-cycle stages') with *thematic relations* (see check marks in Table 3).

Table 3 Characterizing our criteria types as intra- and inter-entity.

Underlying relations	Taxonomic	Shared-properties	Thematic
Types of criteria			
'Biological'	✓	✓	✓
'Ordinary'	-	-	✓
'Appearance-related'	-	✓	-
'Story-making'	-	-	✓
	Intra-entity criteria		Inter-entity criteria

Moreover, all our ordinary ('aesthetics', 'usefulness for humans', 'context details', and 'pair absence') and story-making criteria had to do with *thematic relations*, whereas all the appearance-related criteria ('colour', 'size', 'shape', 'texture / material', and 'other details') had to do with *shared-properties* (see check marks in Table 3). In other words, when children appealed (a) to some biological criteria by saying for instance *'They go together because the turtle eats the plants', or* (b) to ordinary criteria by saying for instance *'I put the chair and the bag because this is the school bag and this is the school chair'*, or (c) to story-making criteria by saying for instance *'The turtle goes on holiday and takes the bag with it'*, their reasoning strands were actually based on different *thematic* relations. Similarly, when children appealed (a) to some other biological criteria by saying for instance *'I put these together because they can both fly: flamingo & bird'*, or (b) to appearance-related criteria by saying for instance '*Bird and bag: together because they are both blue'*, their reasoning strands were actually based on different *shared-properties*. Finally, when children appealed to the more advanced biological criteria by saying for instance *'Flamingo and bird, because they are both birds'*, or '*Sunflower, tree, tree with blue flowers and orange flowers: I put them in the same group because they are all plants … together'*, their reasoning strands were actually based on *taxonomic* relations.

As explained in the introduction, criteria that have to do with the entities' ontological status or properties, could be characterized as intra-entity criteria, while criteria that have to do with thematic relations, could be characterized

as inter-entity criteria. Thus, those biological criteria that drew upon taxonomic relations or shared-properties, along with the appearance-related criteria that all drew upon shared-properties were characterized as intra-entity criteria. Moreover, those biological criteria that drew upon thematic relations, along with the ordinary and the story-making criteria that all drew on thematic relations, as well, were characterized as inter-entity criteria (see check marks in Table 3).

As shown in Table 4, the frequency of intra-entity and inter-entity criteria did not differ much between the pictures-task and objects-task. Children appeared to slightly prefer intra-entity rather than inter-entity criteria, regardless whether they were sorting pictures or objects. More specifically, intra-entity criteria were used (a) for the formation of 263 groups by 104/120 children in the pictures-task, and (b) for the formation of 237 groups by 93/120 children in the objects-tasks; whereas inter-entity were used (a) for the formation of 216 groups by 82/120 children in the pictures-task, and (b) for the formation of 210 groups by 81/120 children in the objects-task. Intra-entity criteria were slightly more frequent at the pictures-task rather that the objects-task (263 groups by 104/120 children vs. 237 groups by 93/120 children), while inter-entity criteria were almost equally activated in both (216 *vs.* 206 groups by 82/120 and 81/120 children respectively).

Table 4 The frequency of inter-entity and intra-entity criteria in the two free-sorting tasks.

	Pictures-task	Objects-task
Intra-entity criteria	263 groups / 104 children	237 groups / 93 children
Inter-entity criteria	216 groups / 82 children	210 groups / 81 children

However, the differences between intra-entity and inter-entity criteria regarding the times they were used, or in other words the number of groups they gave rise to, were not found to be statistically significant ($\chi^2(1)= .33$, $p \gg .05$). The same was valid for the number of children who used (a) intra-entity criteria only (pictures-task: 38, objects-task: 39), (b) inter-entity criteria only (pictures-task: 16, objects-task: 27), and (c) both intra-entity and inter-entity criteria (pictures-task: 66, objects-task: 54) ($\chi^2(2)=4.03$, $p \gg .05$).

4. Discussion

The participants sorted pictures or objects by drawing upon a rich conceptual toolkit. They used 'appearance-related' criteria to justify some of their groups, but they did not seem to be perception-restricted, since they often used ideas that had to do with the biological world, as well as with the world of everyday life and the world of imagination. Children's toolkit was rather dominated by a series of 'biological' criteria that drew upon *taxonomic, shared-properties* and *thematic* relations. In fact, (a) animals or animal kinds like birds, (b) plants or plant kinds like trees or flowers, and (c) artifacts, were used very frequently for pictures or objects groups, while the distinction between animals, plants and artifacts was present as well. Moreover, children appeared ready to apply biological knowledge about 'food relationships', 'habitats' and 'life cycles' in order to meet the sorting requirements of both tasks. This was valid for their knowledge about the 'movement type' or some 'typical body features' of living organisms, as well.

On the other hand, children used a series of 'ordinary' criteria concerning thematic relations encountered in everyday life. This indicates their potential to combine everyday knowledge and apply it in problem-solving situations like the ones of our interview protocol. The most popular type of ordinary criteria, the 'context details', is probably the best example. Forming pictures or objects groups by appealing to the fact that we typically find the represented entities in the same location, or that we use them at the same periods or in the same way, highlights not only children's natural tendency to bring everyday knowledge to school, but also their creative ways of making sense of the world. Part of this creativity, probably a more primitive one, is expressed through the use of another type of sorting criteria - 'story-making'. Children integrated several entities into imaginary stories they invented in order to cope with the pictures-task and the objects-task.

Thus, the children applied both knowledge and creativity in order to form their groups. Focusing on the biological knowledge in particular, one could suggest that it is quite promising, especially since it is expressed without being explicitly prompted. Children appealed to the entities' ontological categories, to their shared morphological or functional properties, or to their thematic relations in the context of food chains, habitats and life-cycles, just in order to justify the groups they spontaneously formed for us. Nevertheless, we could

not miss the fact that even though children appealed to kinds of 'animals' or 'plants' to ground their groups, and sometimes they did make the higher-level distinction between 'animals', 'plants' and 'artifacts', the over-arching distinction of 'living/non-living' seemed to be totally absent from their reasoning. In fact, none of the children divided the represented entities in the group of 'living' and the group of 'non-living' in any task. This is obviously something we need to take into account when we design learning environments for young children. The 'living/non-living' distinction is very important for understanding the biological world (Inagaki & Hatano, 2006). Children need to build it in order to be able to leave 'personifying reasoning' (Inagaki & Hatano, 2002) behind them; in other words, to be able to predict properties and functions of living organisms which are not very familiar to them, not by drawing upon their external similarities or differences with humans, but upon their ontological identity as 'living beings'.

The four main types of sorting criteria we traced were used to a similar extent regardless of whether our informants were sorting pictures or objects. Notable differences were also absent when we translated these criteria types to intra-entity or inter-entity ones. So, unlike the findings of Gelman et al. (2005) or Ware et al. (2013), it seems that in our free-sorting tasks the use of intra-entity and inter-entity criteria did not depend on the stimulus medium. Intra-entity criteria were prioritized with both pictures and objects. Many factors might account for this. Our participants were older than those of the above studies. They also had specific instructions to group the items themselves while being with the interviewer; they didn't just reason about the items in free mother-child interactions. Moreover, they performed the tasks in a school setting (not in a laboratory one), where some of them have probably already learnt to prioritize certain types of intuitions. Finally, they were asked to reason about different types of items ('animals - plants - artifacts' instead of 'animals - food - artifacts').

Our findings show that both objects and pictures could be used within introductory learning environments for participants around the age of five. It seems that they can already spontaneously activate intra-entity conceptual relations concerning kind and shared-properties, which are the stepping stone for understanding the biological world. Biology instruction could build on children's emerging potential in order to help them enhance their ability to

categorize taxonomically, and finally build the overarching distinction between 'living' and 'non-living'.

Acknowledgments

References

Allen, M. (2015) Preschool children's taxonomic knowledge of animal species. *Journal of Research in Science Teaching, 52*(1), 107-134.

Chen, S. H. & Ku, C. H. (1998) Aboriginal children's alternative conceptions of animals and animal classification. *Proceedings of the National Science Council Republic of China,* Part D, *8*, 55-67.

Deak, G. O. & Bauer, P. J. (1996) The dynamics of preschoolers' categorization choices. *Child Development, 67*(3), 740-767.

Ergazaki, M., Gasparatou, R., Spai, A., Dimitrakopoulou ,A. & Kyriazidou, M. (2016) How do preschoolers spontaneously form categories?. In J. Lavonen, K. Juuti, J. Lampiselkä, A. Uitto & K. Hahl (Eds.), *Electronic Proceedings of the ESERA 2015 Conference. Science education research: Engaging learners for a sustainable future,* Part 15 (co-ed. E. Glauert & F. Stylianidou), (pp. 592-598). Helsinki, Finland: University of Helsinki.

Gelman, S. A., Chesnick, R. J. & Waxman, S. R. (2005) Mother–child conversations about pictures and objects: Referring to categories and individuals. *Child Development, 76*(6), 1129-1143.

Gelman, S. A., Waxman, S. R. & Kleinberg, F. (2008) The role of representational-status and item-complexity in parent–child conversations about pictures and objects. *Cognitive development, 23*(2), 313-323.

Gibbs, G. R. (2005) *Qualitative Data Analysis: Explorations with NVivo.* Maidenhead, England: Open University Press.

Inagaki, K. & Hatano, G. (2002) *Young children's naïve thinking about the biological world (Essays in Developmental Psychology).* New York: Psychology Press – Taylor and Francis Group.

Inagaki, K. & Hatano, G. (2006) Young Children's Conception of the Biological World. *Current Directions in Psychological Science, 15*(4), 177-181.

Kattmann, U. (2000) Do students have an implicit theory of animal kinship. In B. Andersson, U. Harms, G. Helldén, & M.-L. Sjöbeck. (Eds), *Research in Didaktik of Biology. Proceedings of the 2nd Conference of European Researchers in Didaktik of Biology* (pp. 61-83). Göteborg, Sweden.

Kattmann, U. (2001) Aquatics, Flyers, Creepers and Terrestrials–students' conceptions of animal classification. *Journal of Biological Education*, *35*(3), 141-147.

Markman, E.M. & Hutchinson, J.E. (1984) Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16*(1), 1-27.

Markman, E. M., Cox, B., & Machida, S. (1981) The standard object-sorting task as a measure of conceptual organization. *Developmental Psychology, 17,* 115–117.

Murphy, G.L. (2002) *The big book of concepts*. MIT Press.

Trowbridge, J.E. & Mintzes, J.J. (1985) Students' alternative conceptions of animals and animal classifi-

cation. *School Science and Mathematics*, *85*(4), 304-316.

Trowbridge, J.E. & Mintzes, J.J. (1988) Alternative conceptions in animal classification: A cross age study. *Journal of Research in Science Teaching*, *25*(7), 547-571.

Ware, E.A., Gelman, S.A. & Kleinberg, F. (2013) The medium is the message: Pictures and objects evoke distinct conceptual relations in parent-child conversations. *Merrill-Palmer Quarterly, 59*(1), 50-78.

Waxman, S.R. & Namy, L.L. (1997) Challenging the notion of a thematic preference in young children. *Developmental Psychology, 33*(3), 555-567.

Wright, K., Poulin-Dubois, D. & Kelley, E. (2015) The animate–inanimate distinction in preschool children. *British Journal of Developmental Psychology*, *33*(1), 73-91.

Yen, C.F., Yao, T.W. & Chiu, Y.C. (2004) Alternative conceptions in animal classification focusing on amphibians and reptiles: A cross-age study. *International Journal of Science and Mathematics Education*, *2*(2), 159-174.

Yen, C.F., Yao, T.W. & Mintzes, J.J. (2007) Taiwanese students' alternative conceptions of animal biodiversity. *International Journal of Science Education*, *29*(4), 535-553.

Yorek, N., Sahin, M. & Aydin, H. (2009) Are animals 'more alive' than plants? Animistic-anthropocentric construction of life concept. *Eurasia Journal of Mathematics, Science & Technology Education*, *5*(4), 369-378.

Promoting biological literacy skills of gymnasium students during the operation of a new competence-based biology curriculum

Anne Laius[1], Aveliis Post[1] and Miia Rannikmäe[1]

anne.laius@ut.ee

Abstract

Sustainable societies require a creative workforce able to engage in reasoning skills in order to cope with the rapidly changing world. There is a major need for developing creative and reasoned problem solving skills that lead to the competence of decision making. The goal of this research is to determine the influence of a new competence-based biology curriculum for gymnasium students' scientific creativity and socio-scientific reasoning skills in the context of lactose intolerance. The longitudinal study incorporated 44 representative secondary schools and 10th grade (N=1116) and 12th grade students (N=848) were tested against biological literacy components. The results reveal that three years of biology studies have not significantly affected the students' learning outcomes in enhancing the competences associated with scientific creativity and socio-scientific reasoning skills, and the new biology curriculum has not supported the students' ability to transfer biological and chemical knowledge into everyday situations. The data were gathered with an 8-item test of assessing students' biological literacy components, including scientific creativity and socio-scientific reasoning. Statistical analysis (Wilcoxon Signed Rank Test) of collected data reveal that there was a statistically significant increase in mean scores related to students' scientific creativity tasks, and these changes were meaningful as effect size (Cohen's d) was 0.34. Similar statistics for socio-scientific reasoning tasks did not show a statistically significant change. The results show that the implementation of a new competence-based biology curriculum has had a bigger effect on students' scientific creativity skills (mean increase 0.29) than to socio-scientific reasoning skills (mean increase 0.04).

---

[1] Science Education Centre, University of Tartu, Estonia

1. Introduction

During the past few decades there have been many international developments focusing on ways to change secondary school science, including biology. Two main arguments have been put forward to justify such change, the decrease in student interest in science (Ekborg et al., 2013) and need to broaden the curriculum with more emphasis on development of competences and transferable skills rather than solely academic knowledge and skills (Beier, 2014). The role of biology education is to promote biological literacy and offer support for coping with the problems of everyday life which involve biological knowledge.

In 2011, Estonia introduced a competence-based biology curriculum, intended to initiate a paradigm shift from memorization of knowledge to the development of competences and transferable skills. The goal for this research was to assess the possible change in students' biological literacy levels in terms of scientific creativity and socio-scientific reasoning skills during the operation of the competence-based biology curriculum. According to this goal the following research questions were proposed: (1) *What possible differences are there between male and female students' scientific creativity and socio-scientific reasoning skills both in 10th and 12th grade?* and (2) *Has the competence-based biology curriculum been effective for increasing the students' scientific creativity and socio-scientific skills during three years of studies from 10th to 12th grade?*

2. Theoretical background

Recently suggested ideas for increasing the interest towards learning science, including biology, are making use of socio-scientific issues (SSIs) (Lee et al., 2013). SSIs are shown to be effective in developing students' creative problem solving, decision making and socio-scientific reasoning skills (Eastwood et al., 2013). The use of SSIs has been effective for engaging students in learning biology (Lenz & Willcox, 2012), and also for development and measurement of 21st century skills (Bellotti et al., 2014). Scientific creativity (ScCr) and socio-scientific reasoning (SsR) are important components within the cognitive strand of 21st century skills, and are emphasized in many new science educa-

tion standards (Zhou et al., 2016).

With the emergence of increased attention to competences required by citizens for a knowledge-based society, schools and educational systems around the world have been called upon to make changes to their curricula (Voogt & Roblin, 2012). Solving problems, which have socio-scientific content in biology lessons help to develop students' to make reasonable, independent decisions, which are based on biological knowledge. Students' development of their reasoning for the purpose of making informed decisions depends on their understanding of the concepts used (Zeidler et al., 2013). These concepts will have contextualized meanings for each individual and will have an impact on students' understanding.

The need to emphasize decision-making as part of science education has long been noted by science educators (Millar & Osborne, 1998; Zeidler et al., 2005). Decision-making and reasoning are valued skills in society. Good decision-making and reasoning skills are highly valued competences that Estonian stakeholders (e.g. employers) expect from graduating students (Laius et al., 2016).

Creative thinking and problem solving skills are adjustable tools for successfully handling various kinds of unfamiliar problems, which enhance adaptive behaviours in these new settings (Kashani-Vahida, et al., 2017). Creative thinking techniques have been repeatedly shown as promoting creative solutions in many settings (Tsai, 2014). Scientific creativity requires awareness of scientific problems, and this is an important part of being a good scientist (Usta & Akkanat, 2015). Biology teachers must develop students' scientific (biological) literacy, creativity and system thinking (Estonian Curriculum, 2011), and creativity should be developed in everyday problem-solving (Basadur et al., 2014). The topic of lactose intolerance provides a good context for exploring these aspects.

Nowadays an ability to solve problems creatively is a key performance indicator (Trilling & Fadel, 2012) because new problems occur every day. Exploring socio-scientific issues provides a good way of increasing interest in learning science, including biology (Lee et al., 2013). SSIs are proving to be effective in developing students' creative problem solving, decision-making and socio-scientific reasoning skills (Eastwood et al., 2013; Sadler, 2005). In the last decade, science education researchers have made significant advances in using SSIs as contexts for transforming science learning opportunities; but

assessment of SSI-related learning outcomes has left behind other developments in SSI-based teaching and research (Zeidler et al., 2013).

3. Methodology

3.1 Study design

The longitudinal scientific literacy test was conducted with 10th grade students in the fall of one school year and again after three years at the end of 12th grade.

3.2 Sample and procedure

The longitudinal research was conducted in a representative sample of 44 different Estonian schools with 1116 (598 female and 518 male students) in the 10th grade (16–17 years old) and 848 (436 female and 412 male students) in the 12th grade (17–18 years old). The data was collected at the beginning of the implementation of the new curriculum and 3 years later after the implementation.

A pilot study was conducted among upper secondary school students. The reliability, calculated using Cronbach alpha for the overall instrument, was 0.63 (0.62 for 10th grade and 0.64 for 12th grade), and it was considered as acceptable for this under ten item test instrument (Loewenthal, 2004). The data gathering period was fall 2011 (10th grade) to spring 2013 (12th grade).

3.3 Instrument

The instrument was part of a bigger study that focused on socio-scientific contextual problem solving tasks. The instrument used in this current study was based on a socio-scientific contextual situation of lactose intolerance that consisted of 8 tasks focusing on students' abilities of transferring biology knowledge into new everyday situations. All 8 tasks involved socio-scientific problem solving at different levels of difficulty. The situation was based on a family (a mother and triplets – two boys and a girl), and their food provision problems arising because one of the boys suffered from lactose intolerance. Two of the tasks were designed to measure the following aspects of divergent thinking: (1) scientific creativity and (2) socio-scientific reasoning skills.

The task of scientific creativity asked for possible advantages of lactose tolerance in human beings during evolution and the different answers were coded

as follows: (1) no answer or biologically wrong – 0; (2) one possible adequate advantage – 1; (3) two possible adequate advantages – 2; and (4) three or more possible adequate advantages of lactose tolerance – 3.

The socio-scientific reasoning task asked the students to reason about their decisions made on creating or not creating special menus for lactose intolerant people in a private restaurant: (1) no answer – 0; (2) just describing the decision – 1; (3) reasoning about the decision from one aspect – 2 and (4) reasoning about the decision from two or more aspects – 3.

3.4 Data analysis

Data was analysed using IBM SPSS Statistics 20. This was used to describe students' response frequency distribution and to determine how responses to single items within the sub-sections varied between four exam sub-groups, location groups and grades. The Mann-Whitney U test was used to determine the significance of differences (Cohen et al., 2007). Cohen's d was used to calculate the effect size to eliminate sample size influence.

4. Results and analysis

The mean results of the study (Table 1) revealed that in the 10th grade, the students' scientific creativity (ScCr) skills were seen as considerably lower than the skills of socio-scientific reasoning (SsR), but as a result of gymnasium studies within the new curriculum, the results of 12th grade students for the skills of ScCr and SsR levelled off. The data showed that the results of ScCr tasks were statistically significant at the 0.01 significance level, but those of SsR tasks were not significant.

Table 1 The results of scientific creativity and socio-scientific reasoning tasks among 10th and 12th grade students.

Skills	Mean (SD)			Wilcoxon Signed Ranks test		Effect Size (Cohen's d)
	10th grade	12th grade	Difference	Z	Asymp. Sig. (2-tailed)	
Scientific creativity	**1.44** (0.63)	**1.73** (0.68)	0.29	-8.979	**0.000****	0.215*
Socio-scientific reasoning	**1.54** (0.61)	**1.59** (0.68)	0.04	-0.875	0.382	0.038

* Cohen's d>0.2 shows that the difference is meaningful

** Statistical significance at level 0.01

The current study gave an overview of the levels of c students' level of ScCr and SsR skills and illustrates that the majority of 10th grade students stayed at the lower levels (0 or 1) with both ScCr (N=1116) and SsR (N=848) tasks scores. Very small increases for 12th grade students revealed low effectiveness of the new competence-based biology curriculum so far.

Mean results of test in the beginning of the 10th grade showed that gymnasium students had a moderate level of ScCr and SsR skills, and they did not evolve during three years of studies. The focus of the new competence-based curriculum promoting scientific ScCr and SsR was shown not to be applied into teaching and learning in biology lessons.

The test results both of 10th grade and 12th grade students (Table 2) did not show any statistically significant differences between female and male students, although there was a tendency for girls to perform better than boys in all cases. As the effect sizes (Cohen's d) were also below 0.20, it can be concluded that both in scientific creativity and socio-scientific reasoning skills there was no significant gender difference.

Table 2 Comparison of the test results for male and female students.

	Scientific creativity Mean (SD)	Socio-scientific reasoning Mean (SD)
10th Grade		
Male students	**1.09** (0.74)	**1.43** (0.67)
Female students	**1.20** (0.87)	**1.50** (0.70)
Difference	-0.11	-0.07
Wilcoxon Signed Ranks test Z	-1.851	-1.406
Asymp. Sig. (2-tailed)	0.064	0.160
Effect Size (Cohen's d)	0.136	0.102
12th Grade		
Male students	**1.54** (0.67)	**1.47** (0.62)
Female students	**1.64** (0.70)	**1.59** (0.72)
Difference	-0.10	-0.12
Wilcoxon Signed Ranks test Z	-1.408	-1.913
Asymp. Sig. (2-tailed)	0.159	0.056
Effect Size (Cohen's d)	0.146	0.179

5. Discussion and conclusions

The results of the current research indicated that in both areas of scientific creativity and socio-scientific reasoning there are similarities with previous studies about the perception that ScCr and SsR skills of students are comparatively low (Bellotti et al., 2014; Zhou et al., 2016).

The results of the study revealed that in the 10th grade, the students' scientific creativity skills were considerably lower than the skills of socio-scientific reasoning, but as a result of gymnasium studies within the new curriculum, the results of 12th grade students within the assessed skills levelled off.

The comparative statistical analysis show that the mean results of 12th grade students were significantly higher only within scientific creativity skills and not in socio-scientific reasoning skills. Even this increase of results was not meaningful in a pedagogical sense as the effect size was below 0.20 because of the large sample size.

The results of tests of both 10th grade and 12th grade students, showed that the scientific creativity and socio-scientific reasoning skills were not gender specific.

In conclusion, the findings indicated that the students' ability to transfer their biological and chemical knowledge into new situations, in relation to scientific creativity and socio-scientific reasoning skills, was quite low at the beginning of gymnasium studies, and it did not increase significantly during the three years of studying biology at gymnasium level. Although the biology curriculum had changed towards competence-based learning outcomes at the beginning of the current study, it is apparent that this period of time has been too short for meaningful changes in students' results.

This research confirmed the evidence from literature that the use of SSIs is equally effective both in motivating students' engagement in biology studies and in assessing the biological literacy skills of students (Eastwood et al., 2013).

In conclusion, it can be said that the instrument developed was suitable for assessment of ScCr and SsR skills, but the role of these skills in school biology classes is still undervalued as the teachers are not prepared to foster these skills among students by purposefully assessing them, although the curriculum has been updated for this purpose. To enhance the implementation of the new competence-based curriculum into pedagogical practice, the pre-service

training of future biology teachers needs to be more focused on fostering biological literacy and its cognitive skills, including scientific creativity and socio-scientific reasoning skills. There is also a need for large-scale in-service courses for biology teachers to obtain the outcomes of biology education that meet the needs both of the new curriculum and society.

Acknowledgements

References

Basadur, M., Gelade, G. & Basadur, T. (2014) Creative problem-solving process styles, cognitive work demands, and organizational adaptability. *The Journal of Applied Behavioral Science, 50* (1), 80–115.

Beier, Y. (2014) The collaborative advantage. The rewards of a collaborative culture are significant, but so is the effort to get there. *Communication World*, 22–25.

Bellotti, F., Bottino, R.M., Fernández-Manjón, B. & Nadolski, R. (2014) Guest Editorial: Game Based Learning for 21st Century Transferable Skills: Challenges and Opportunities. *Journal of Educational Technology & Society*, 17 (1), 1–3.

Cohen, L., Manion, L. & Morrison, K. (2007) *Research methods in education* (6th Edition). London: Routledge.

Eastwood, J.L., Sadler, T.D., Sherwood, R.D. & Schlegel, D.M. (2013) Students' Participation in an Interdisciplinary, Socio-scientific Issues Based Undergraduate Human Biology Major and Their Understanding of Scientific Inquiry. *Research in Science Education*, 43 (3), 1051–1078.

Ekborg, M., Ottander, C., Silfver, E. & Simon, S. (2013) Teachers' Experience of Working with Socio-scientific Issues: A Large Scale and in Depth Study. *Research in Science Education*, 43, 599–617.

Estonian Curriculum (2011) N*ational Curriculum for basic schools and upper secondary schools*. Regulation of the Government of the Republic of Estonia, RT I, 14.01.2011, 2.

Kashani-Vahida, L., Afrooz, A. G., Shokoohi-Yekta, M., Kharrazi, K., Ghobari, B. (2017) Can a creative interpersonal problem solving program improve creative thinking in gifted elementary students? *Thinking Skills and Creativity*, *23*, 175–187.

Laius, A., Post, A., & Rannikmäe, M. (2016) Determining support of Estonian stake-holders for a new competence-based science education curriculum. *International Journal of Education and Information Technologies, 10*, 14–24.

Lee, H., Yoo, J., Choi, K., Kim, S.-W., Krajcik, J., Herman, B. C. & Zeidler, D.L. (2013) Socio-scientific Issues as a Vehicle for Promoting Character and Values for Global Citizens. *International Journal of Science Education*, 35 (12), 2079–2113.

Lenz, L. & Willcox, M.K. (2012) Issue-Oriented Science: Using Socio-scientific Issues to Engage Biology Students. *American Biology Teacher (University of California Press),* 74 (8), 551–557.

Loewenthal, K. M. (2004) *An introduction to psychological tests and scales* (2 ed.). Hove, UK: Psychology Press.

Millar, R., & Osborne, J. (1998) *Beyond 2000: Science education for the future: A report with ten recommenda-*

tions. King's College London, School of Education.

Sadler, T.D. (2005) Evolutionary theory as a guide to socioscientific decision-making. *Journal of Biological Education, 39* (2), 68–72.

Trilling, B. & Fadel, C. (2012) *21ˢᵗ Century Skills: Learning for Life in Our Times*. Jossey-bass, A Wiley Inprint.

Tsai, K. C. (2014) A Review of the Effectiveness of Creative Training on Adult Learners. *Journal of Social Science Studies*, *1* (1), 17-30.

Usta, E. & Akkanat, C. (2015) Investigating Scientific Creativity Level of Seventh Grade Students. *Social and Behavioral Sciences*, *192*, 1408–1415.

Voogt, J. & Roblin, N.P. (2012) A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies. *Journal of Curriculum Studies*, 44(3), 299–321.

Zeidler, D. L., Sadler, T. D., Simmons, M. L. & Howes, E. V. (2005) Beyond STS: A research-based framework for socioscientific issues education. *Science Education*, *89* (3), 357–377.

Zeidler, D.L., Herman, B.C., Ruzek, M., Linder, A., & Lin, S.S. (2013) Cross-Cultural Epistemological Orientations to Socioscientific Issues. Journal of Research in Science Teaching *50* (3), 251–283.

Zhou, S.; Han, J.; Koenig, K.; Raplinger, A.; Pi, Y.; Li, D.; Xiao, H.; Fu, Z. & Bao, L. (2016) Assessment of scientific reasoning: The effects of task context, data, and design on student reasoning in control of variables. *Thinking Skills and Creativity*, 19, 175–187.

Biology: the ultimate science for teaching an understanding of scientific evidence

Ros Roberts[1]

rosalyn.roberts@durham.ac.uk

Abstract

Recent school science curriculum developments in many countries emphasise that scientists derive evidence for their claims through different approaches; that such practices are intimately bound up with disciplinary knowledge; and that the quality of data should be appreciated. This position paper focuses on the role of Biology to understand evidence, an essential component of 'scientific practice'. Biology is an empirical science, using evidence to support claims. Yet biological practice is diverse – including, *inter alia*, observations, lab-based experimentation, field trials, ecological surveys, randomised controlled trials – so how can we teach, within the time-constraints of the curriculum, to help pupils really understand about evidence in biology? In this paper biology is shown to be the ultimate context for teaching about evidence and 'scientific practice'. The paper draws on a body of research that presents an understanding of the validity of data as a set of conceptual relationships, shown on a concept map. Using examples from biological practice, the paper shows how teachers can illustrate the application of the network of all these ideas and their inter-relationships within the biology curriculum, to help pupils develop the necessary 'thinking behind the doing'. The paper explores ways in which this understanding is inherently related to underpinning disciplinary ideas of biology.

Keywords: biological practice, concept map, concepts of evidence, investigations, practical work

---

[1] School of Education, Durham University, UK

1. Introduction

The science curricula in many countries now include not only 'the products' of science - substantive facts, theories and laws, sometimes referred to as the content knowledge (e.g., OECD, 2013) - but also "the processes and characteristics of the scientific enterprise" (Roberts, 2011, p. 12). A common feature is to "understand the methods by which science derives the evidence for the claims made by scientists, [and] to appreciate the strengths and limits of scientific evidence" (Millar & Osborne, 1998, p. 2004). Large-scale international science assessments such as PISA (OECD, 2013) and TIMSS (Jones et al., 2013) also reflect this curriculum emphasis of understanding 'scientific practice'.

Roberts & Johnson (2015) report that many international curricula now, in addition to the vital substantive knowledge-base of science, include an understanding of the diversity of empirical practice; the important relationship between substantive knowledge and this 'doing' aspect; and the importance of pupils being able to use their understanding to evaluate empirical work and reason with evidence as well as being able to carry out practical work. Roberts & Johnson (2015) put forward a detailed case for how curricula might be framed and how teaching may be structured to support an understanding of evidence which underpins 'scientific practice'. This position paper will build on their work and will argue that biology as a subject, due to its nature, provides an ideal opportunity for teaching about scientific evidence.

2. Scientific practice

Scientific practice is not the same as 'school science practice'. In school science we aim to teach about scientific practice, but what we enact in the school curriculum - often but not exclusively through our use of practical work – is, understandably, not necessarily what scientists do in their practice. This distinction is important. Scientific practices vary across science disciplines and also within them, and not only because of the specific subject knowledge required. How research is designed and conducted to solve different problems varies according to what is being investigated (for instance, using different approaches such as tightly controlled lab-based research; or the observation of phenomena; in field surveys which suggest links between factors; or in Randomised Controlled Trials (RCTs)) each with concomitant effects on the validity of the data and the strengths of the claims made. Research may also

differ in, *inter alia*, the equipment employed, the manual skills required and the specific techniques selected. This presents a challenge to curriculum developers and teachers. How can this diversity of scientific practice be taught and evidence understood within the limitations of a school curriculum?

2.1 Chains and nets

Research into real-world practice can give us some insights that can help to frame the problem. Kinchin and several colleagues over the years have extensively researched expertise and their analysis is represented in Figure 1. They have shown that expert practice, despite appearing to consist of linear chains of practice, is underpinned by complex nets of ideas. The understanding represented by the nets makes the difference between competent enactment of a particular practice and the ability to make evaluative decisions about the practice. Experts therefore draw on a complex net of underpinning ideas to make decisions as they practice which may make it appear deceptively 'straightforward'; and yet, crucial to this is the networked understanding that enables them to make informed and evaluative decisions.
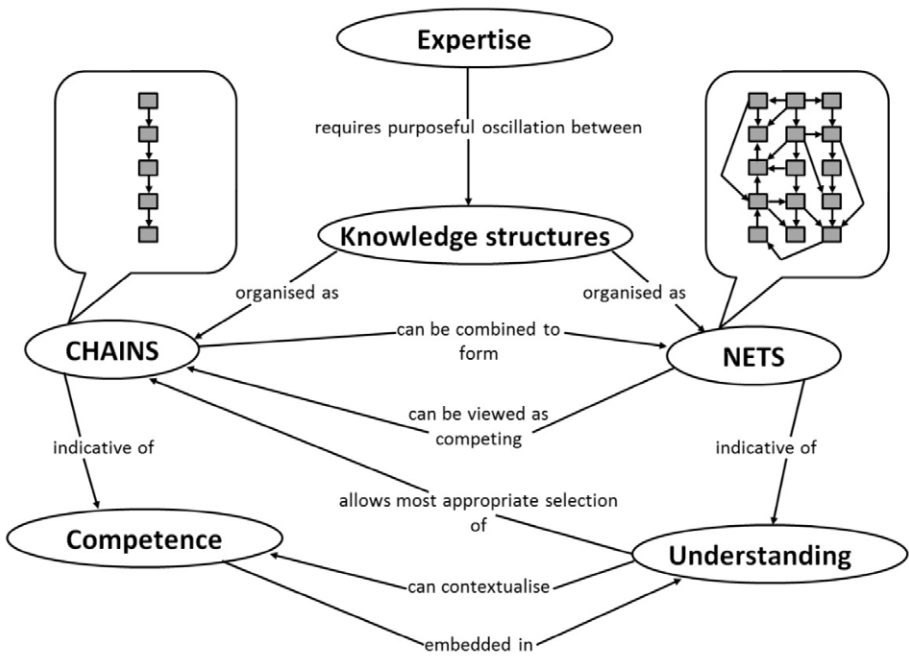
Figure 1 A dual-processing knowledge structures perspective on the nature of expertise (from Kinchin & Cabot 2010:161).

In empirical sciences there are, of course, myriad practices which endeavour to collect valid data – and biology, in particular, encompasses a very diverse range – which could each be represented by chains, but the importance of Figure 1 is that it draws attention to the understanding, represented by the nets, that underpins expert practice. Roberts (2016) argues that in scientific practice, it is this understanding of evidence, the 'thinking behind the doing', that enables evaluative decisions to be made to optimise the quality of the data collected.

While school curricula may not aim to develop expertise in school pupils, we do expect pupils to be able to "develop *understanding* of the nature, processes and methods of science, through different types of scientific enquiry that help them to answer scientific questions about the world around them … (and) develop their ability to evaluate claims based on science through critical analysis of the methodology, evidence and conclusions" (as stated in the National Curriculum for England (DfE, 2014, pp.3-4, emphasis added), which is not atypical of other countries' curricula. Figure 1 therefore has implications for what and how we teach pupils to meet these specific curriculum aims.

Practical work in science classrooms is often represented as chain-like practice – as relatively linear processes to be followed, recipe-like, with few decisions to make (Abrahams & Reiss, 2012). Of course, if the purpose of the practical is to illustrate some substantive idea or phenomenon (such as osmosis or photosynthesis) then following a protocol that someone else has already developed specifically to reduce any uncertainty so that the substantive ideas are clear may be useful (most particularly if the substantive idea has already been introduced and the practical is used to help reinforce the learning (Hodson, 1992; Millar, 1998)). If the purpose is to illustrate, by exemplification of different approaches, the diverse range of practices and protocols in science, or to develop manual skills, it too can arguably be useful. However, there is evidence that just *doing* a practical does not always result in meaningful learning (Abrahams & Reiss, 2012) and research shows that, for the most part, it does not provide a route to understanding evidence (Glaesser et al., 2009) which is the focus of this paper.

Many science curricula are written as 'descriptions of practice' to guide staff and pupils in doing investigations, breaking down the complexity of investigating into sequential processes such as 'planning', 'collecting data',

'analysing' and 'evaluating' but many pupils still struggle to design their own investigations (which Glaesser et al, 2009, show requires that they understand evidence) other than by close mimicking of 'unwritten recipes' of chain-like procedures with which they are already familiar (Roberts & Gott, 2006). Pupils seem to resort to similar chains of practice since they have not developed an understanding of evidence which can be applied to solve problems.

In short, chain-like practicals (the left of Figure 1) have their place in science teaching but are not an efficient or systematic ways to develop the understanding of evidence (the right of Figure 1) that underpins scientific practice as specified in the curricular aims. Lunetta et al. (2007, p.433) (using the term 'laboratory' synonymously with 'practical work') have suggested that "Much more must be done to assist teachers in engaging their students in school science laboratory experiences in ways that optimise the potential of laboratory activities as a unique and crucial medium that promotes the learning of science concepts and procedures, the nature of science, and other important goals in science education". Practical work in the curriculum is necessarily constrained by time and opportunity. Perhaps, instead of exposing pupils to a limited selection of chains of practice, and hoping that that they will understand scientific practice almost 'by osmosis', if a curriculum was framed in terms of understanding the ideas about the quality of data that underpinned such practice, and practical work was selected explicitly to support this understanding, there could be an efficient way to meet the curriculum aims.

2.2 Concept map

With this in mind, Roberts and Johnson (2015) have attempted to show the 'thinking behind the doing' of science encompassing most if not all of 'scientific practice' in schools as a concept map (Figure 2) which focuses on the quality, or validity of the data. Understanding can be represented on concept maps, where both the ideas and their inter-relationships can be shown (Novak & Cañas, 2007). This map centralises the question of the quality of the data since the degree of confidence in the validity gives it weight as empirical evidence for a claim. Understanding the quality of data is an essential component of scientific practice. Figure 2 is their attempt to articulate the net of ideas underpinning scientific practice, equivalent to that on the right of Figure 1. The concept map is based on the concepts of evidence (Gott et al., n.d.) and illustrates how scientific practice is underpinned by a conceptual under-

standing: an understanding about the quality of data. The map focuses on the validity of data generated in carrying out scientific investigations. The quality of empirical data is a central component where the ideas provide the foundation for understanding other aspects of scientific practice. Space does not allow elaboration upon these contingencies here but the reader is referred to Roberts & Johnson (2015) who discuss the implications of the map to teaching the Nature of Science and argumentation.

Viewing scientific practice as having a conceptual knowledge-base to be understood, rather than a series of procedures to be practiced, represents an ontological shift in its characterisation, which has implications for curriculum developers, teaching and assessment. As a knowledge-base, the ideas and understanding of evidence can be specified in a curriculum and explicitly taught and assessed, just as the substantive ideas of science are.

Broadly speaking, Figure 2 has two interrelated sides. On the left is thinking about variables and on the right thinking about measurement. The relationships between these ideas are the basis for decision-making ('the thinking behind the doing'). Since the arrows represent the conceptual links between the ideas, there is no implied sequence in the map and it is not a 'flow diagram'. The map shows the understanding behind the whole of 'scientific practice' in schools. It is important to note that some practices focus on just some areas of the map.

Roberts and Johnson (2015) have exemplified how the ideas in the map are used to make decisions when collecting data in both lab-based and ecological fieldwork; significantly the map does not privilege any particular approach to scientific practice (and this will be expanded on later) but shows the understanding regardless of approach. In terms of validity, there is no distinction between approaches (such as an 'experimental approach' or an 'observational' or 'historical approach') to finding patterns in data (Cleland, 2002). The key issue is what is appropriate depending on the circumstances.
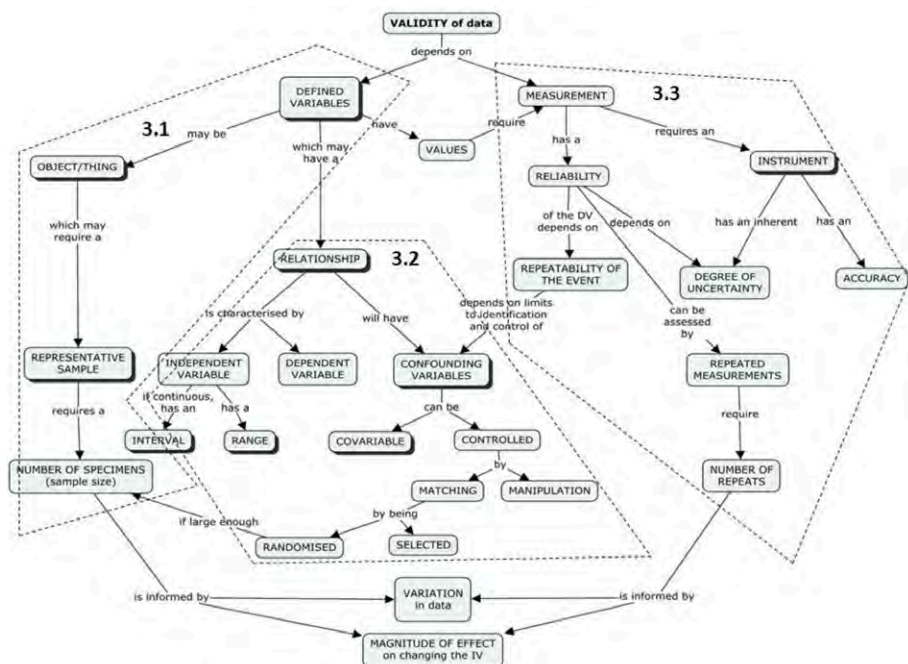
Figure 2 A concept map of the 'thinking behind the doing' of scientific practice (based on Roberts & Johnson, 2015:348). Concepts directly informed by substantive knowledge are highlighted with a shadow on the box. Concepts addressed in sub-sections 3.1, 3.2 and 3.3 of this paper are identified.

From the curriculum perspective (to understand evidence, vital to 'scientific practice') with the aims of both being able to investigate and evaluate others' work, it is worth noting that the same understanding is employed when evaluating scientific practice as it is when collecting valid data. Roberts & Johnson (2015) illustrate how an investigator uses these ideas and their inter-relationships to make decisions while conducting both lab-based and ecological investigations; while a case study (Tytler et al., 2001) - about sampling emissions from a cement-works when it changed the fuel burnt in its furnace - illustrates how the same ideas are important when evaluating scientific practice, in this case in an outdoor context. Essentially, when investigating - in the lab or in the field - decisions have to be made with an eye on the **validity of the data**[2] to be collected and, when evaluating, the same understanding is used to interrogate the data.

---

[2] Terms shown in bold in the account are ideas of evidence represented on the concept map (Figure 2).

In investigations into potential relationships between variables (just one type of scientific practice, but one which employs all the ideas shown on the map and hence used to illustrate the map), **variation in the data** (relative to the **magnitude of the effect** of the change in the independent variable) is key and this is affected by many factors which are shown (from left to right) on the concept map: factors to do with variation in the variables being investigated; in the ability to manipulate and control variables; and in the uncertainty of the measurement. Roberts & Johnson (2015, p. 359) illustrate how "decisions when investigating are based on nuanced application of these ideas, involving mental juggling as juxtapositions and contingencies are considered according to context". Understanding and applying these ideas is a far cry from the routine 'recipes' of many chain-like school practicals and employs higher-order thinking to meet the aims of the school science curriculum.

So how can this understanding be developed and why is Biology the ultimate science for addressing this? This understanding of the quality of evidence is inextricably linked with substantive understanding (the 'subject matter' of science – its facts, models, laws and theories) and those concepts directly informed by substantive knowledge are highlighted with a shadow on the box in Figure 2. It is the substantive 'subject matter' of biology that provides opportunities for a sophisticated understanding of the map to be readily taught (discussed later). The map emphasises the intimate integration of substantive knowledge with scientific practice. Neither stands alone, each is only as good as the other. The soundness of substantive knowledge depends on the quality of the originating data as evidence. They are inextricably bound (as discussed further in Roberts & Johnson (2015) and Johnson & Roberts (2016)).

3. The nature of biology and scientific practice

The variables of science are the creation of the substantive knowledge of the discipline. As Lederman et al. (2014, p. 68) state, investigators "need to have specific knowledge that has been melded into some curious pattern or question". Any limitations in understanding of pertinent substantive ideas affects what can be observed (Haigh et al., 2012). Substantive knowledge is fundamental to scientific practice. The many different approaches to research and the resultant validity of the data and strength for a claim can be viewed as the consequence of differences in the nature of the variables involved and

their measurement in any investigation. The degree to which variables can be isolated and their values manipulated, and the amount of variation in a defined variable influence how relationships are sought and the strength of any resultant claims. This is true of all scientific investigations, yet in the majority of school physics and chemistry investigations (typified as involving variables with very little variation, that can be easily manipulated and isolated and with well-established instruments with relatively little uncertainty) the opportunity to develop the understanding represented by the map may be missed. However, the nature of biology's 'subject matter' provides the ultimate context to develop an understanding of evidence, as represented by the map.

Biological practices vary and include *inter alia* tightly controlled lab-based research; the observation of phenomena; field surveys which suggest links between factors; and Randomised Controlled Trials (RCTs). These diverse approaches have concomitant effects on the validity of the data and the strengths of the claims made; all employ the underpinning understanding of evidence represented on the concept map. The map centralises the question of the validity of data since the confidence in the validity in any research practice gives it weight as evidence for a claim – it is this that all biologists are striving for, regardless of their approach, and is at the forefront of expert researchers' thinking whether they are researching in a lab or the field, doing 'classical experiments' or 'observational study' (Gray, 2014). These diverse practices can be related to the nature of biology's variables, and it is the nature of biology's variables and the resultant diversity of practice that allows *all* the ideas on the concept map to be explored in a biology curriculum. Working from left to right across the concept map (Figure 2) these will be exemplified.

3.1 Variables

The **defined variables** of biology include those that are the focus of physics and chemistry – length, mass, time, energy, substance, rate of chemical change etc. – but also include many with larger inherent variation, for example 'a species'; 'a community'; varying 'environmental factors'. To be able to investigate variables like these, whatever role they have in an investigation, a **sample size** will be required and the validity of the data acquired and claims made will depend on the **representativeness** of the sample. This is not to suggest that the inherent variation in many of the variables in biology is a problem! These variables are the subject matter of biology and are often the focus of

biological practice. The map points to how, if there is variation, this must be taken into account when understanding evidence. If there is little variation in the variable, such as in work with isolated genes or when clones are being used, the issue of variation has been reduced so much that a small sample size (sometimes reduced to an individual) is going to provide valid data, as it does in many chemistry and physics practices where, for instance, 'quality control' processes eliminate any potential cause of variation, e.g. in the purity of a substance, and therefore the opportunity to teach these important ideas about evidence – shown on the left of the map – may be overlooked. But in many instances in biology, such as work involving variables such as 'holly bushes', 'flow rate of a stream' or 'boys', the inherent variation will affect the validity of the data unless a sufficient sample of readings 'captures' the variation in each. In school practice the collection of large class datasets in such circumstances in biology provides an opportunity to teach explicitly about these ideas.

3.2 Relationships

In an investigation where a **relationship** is sought between variables (identified as the **independent** and **dependent variables** on the map; but, as Roberts & Johnson (2015) state when explaining the map, their terminology does not imply a causal link), the ability to isolate other potentially **confounding variables** affects the validity of the data. The identification of confounding variables absolutely draws directly on substantive knowledge and is limited by that knowledge. There has to be a reason for deciding upon a particular variable, if only because it might be relevant. The potential effects of confounding variables must be **controlled** in some way to isolate any relationship between the variables being investigated. But the nature of many of biology's complex variables can make this difficult. For instance, in research involving individual organisms, each organism is composed of a complexity of other, inseparable, **co-variables**. Even at a simpler level, variables such as substrate size and oxygenation of water in a stream are co-variables associated with the velocity of the water. It is often only in 'experimental' conditions that variables might be isolated and **manipulated**. The manipulation of variables is a means of control so that the **values** can be fixed at a constant value so as to isolate the relationship being investigated. In field trials, control can be achieved even if the values are changing over, say, 24 hours, by manipulating the experimental conditions so

that all specimens are subject to the same conditions (Roberts, 2001). The inability often to isolate variables in biology means that other approaches to deal with the effects of confounding variables must be used, so that the relationship can be explored but with consequent effects on the **variation in the data** collected. In ecological surveys, for instance, the values of the variable 'aspect of slope' can only be **matched** by **selection** of sites with similar values; and in the formation of comparisons groups samples may be matched on several key characteristics e.g. patients selected with similar conditions such as being of a similar weight, age and medical history. In randomised controlled trials (RCTs), subjects are assigned to treatment groups by a **random** process. With a large enough sample size it can be assumed the multifarious confounding variables will 'even out' so the only difference overall is the treatment applied to one group and not the other. Roberts & Johnson (2015) and Johnson & Roberts (2016) provide further exemplification.

3.3. Measuring variables

The concept map shows that another source of variation in the data is due to **measurement** issues. Measurements in biology can be both quantitative and qualitative. The concept of measurement (although often referred to as 'observation'; see, for instance Gray, 2014) can also be applied to categoric variables, where qualitative descriptions are the values. Here, the measurement entails the recognition of the defining features of the variable (for instance, a species), with the substantively-informed discernment of the observer acting like an instrument; and with an element of uncertainty introduced if identification is wrong. The quality of the data and the strength of any claim made from it are affected by **variation in the data.** There is a **degree of uncertainty** in all measurements but in some situations in biology this can be quite large (in comparison with many situations in the physical sciences) and this provides an opportunity to teach about the effect it has on the variation in the data explicitly. In situations where the variation caused by the nature of the variables is relatively low, the uncertainty associated with measurement can contribute significantly to the variation in the data. Such is the case in many physics and chemistry-focused investigations and then much attention is paid to improving the measurements to reduce the variation. In some situations in biology where the variation in the data, due to the nature of the variables, is large in relation to the effect of changes in the independent variable, the focus of the

efforts when collecting data may be relatively less on improving the measurement. For example, ACFOR scales based on estimates of relative abundance of a species in ecology have acknowledged uncertainty but pragmatic reasons may influence their adoption in situations where the other causes of variation in data are so large. Examples like this, in biology teaching, provide opportunities to explicitly address important concepts of evidence associated with measurement that affect the quality or **validity of data.**

The nature of many of biology's variables are complex, often involving inherent variation i.e. within a species, requiring sampling. Many variables cannot be easily isolated one from another and cannot always be manipulated by the investigator. All of these characteristics, as well as the quality of any measurements, result in variation in the data, which in turn affects the strength of the resultant claim and our explanation of any potential relationship found: causation may not necessarily be claimed. Concluding that data do show a relationship is one thing, explaining that relationship is another matter. The relationship may be causal, may be an association due to a common cause or may simply arise by chance. Substantive ideas will be used here to consider the question of causality and the relative merits of competing theories.

As biologists, we know that our biological practices must take account of the nature of the variables we are working with and this has resulted in a diverse range of approaches in biology to obtaining data that is as valid as possible. Biology provides the ideal opportunity to teach pupils about all the ideas required to understand evidence, as represented in Figure 2.

4. Teaching about evidence in biology

The significance of the concept map (Figure 2) is that it shows that there is a knowledge-base (the concepts of evidence) to understanding the quality of data that underpins scientific practice (the right of Figure 1). Roberts & Johnson (2015) point to the concept map (Figure 2) having implications for research in science education, policy, curricula and school practice. It is the last of these that will now be considered briefly.

Research with undergraduates (see for instance Taylor & Meyer, 2010; Wilson et al., 2010) shows that aspects of scientific research expertise indicative of network thinking are poorly developed as a result of school curricula and practice, wherein pupils often carry out chain-like practicals and scientific

practice is framed in terms of descriptions of practice rather than as a set of ideas. The concept map points to an alternative way to meet our curriculum aims of understanding scientific practice: it frames scientific practice as having a conceptual basis – a set of ideas that can be explicitly taught and assessed, using a range of engaging pedagogical activities, just as we teach the more familiar substantive concepts such as photosynthesis, succession or adaptation. As with other learning, both practical and non-practical teaching activities can be used, with the selection made based on the best route to the learning outcomes (of understanding the ideas and the relationships in the map).

Support for activities suitable for school pupils that focus on these often-ignored ideas is available (see Gott et al, 1997; 1998; 1999; Gott et al, n.d.; Roberts & Gott, 2002; Roberts & Reading, 2015) but a teacher needs to be alert to the possibilities.

In biology, many opportunities to teach this understanding explicitly present themselves due to the nature of biology. All the concepts discussed in this paper have been taught explicitly by the author and colleagues; and key points to be considered when teaching to develop this understanding from this experience are summarised in Roberts & Reading (2015). They suggest that pupils should be introduced to, and develop a secure understanding of, elements of the map through activities where the understanding of evidence ideas is the focus of the teaching activity and the pupils' learning. Trying to learn about evidence when the focus of the teaching is on substantive ideas seems not to work – explicit teaching that focuses on evidence means that pupils aren't distracted from this by the more familiar subject matter. Once pupils have developed that understanding then applying it in more sophisticated biological contexts is possible. As Roberts & Reading (2015, p.38) state: "The challenge is therefore:

1. to plan a curriculum that enables a progression in pupils' understanding of the ideas of evidence;

2. to 'map' this across the progression planned for the substantive content and the school's teaching sequence;

3. to include practical activities within this sequence which have as their focus the illustration or application of the ideas of evidence."

In terms of sequence, we focus on different sections of Figure 2 in our teaching, starting with variables with little inherent variation in contexts where values can be manipulated (the familiar 'fair testing' wherein control variables' values are 'kept the same') and then considering how a valid design might be established in situations where they can't, such as field trials (where many values – such as temperature - are not held at a constant value but are allowed to change across all treatments) to ecological selection of sites matched for key variables. Issues to do with the quality of measurement are then addressed so that students get an understanding of variation in the data before other sources of variation, such as in the sample, are introduced. We have worked with teachers who have decided to introduce pupils to ideas about evidence as a separate 'module', often at the start of the year, so that the ideas can be repeatedly addressed within practical work in the other more familiar substantive modules. Others have decided to introduce pupils to ideas explicitly as they arise through the substantive work they have planned – short 'nodules' of teaching about evidence as the opportunity arises.

Since this approach represents a conceptual basis for understanding evidence, decisions can be made about the sort of activities best used to teach the ideas. As with the substantive ideas of science, these can involve using both practical and 'non-practical' opportunities. Practical work is important in developing this understanding – during open-ended investigations pupils can make these decisions for themselves and see the effect of their decisions on the quality of the data and recognize this affects the strength of their claim. In genuinely 'open-ended' contexts where pupils are not focusing on getting 'the right answer' they can focus on the 'trade-offs' between sections of the map and their practice is characterized by trials and iterative working. We have also found that students learn lots from discussion involving their own and others' data. Having generated their own data, pupils seem better able to understand the sources of variation in it. The importance and meaning of simple statistical tests and graphical forms of data presentation can then be appreciated more in our experience. Having grappled with 'messy data' and discussed how best to get meaning out of it students appear to be in a better position to understand the conventions employed in handling data and presentation.

5. Conclusion

Since the concepts of evidence are validated across all scientific disciplines (Gott et al., 1999; Roberts & Gott, 1999) they can be taught in all science subjects. Chemistry's and physics' focus on variables that have such little variation means that, in school science at least, the ideas of representative sampling can usually be ignored. School chemistry and physics are usually studied in situations where variables can usually be isolated and manipulated and their investigations use well-developed and precise measuring instruments which means that learning opportunities have to be specifically developed to draw learners' attention to all the ideas of evidence on the map. However, the nature of biology and the topics addressed in school biology curricula enable *all* the key elements of the concept map to be readily addressed. The nature of biology's variables and the diverse practices employed and taught about in school biology curricula suggest that biology is the ultimate science to readily develop this understanding of all the ideas on the map.

References

Abrahams, I. & Reiss, M. (2012) Practical work: its effectiveness in primary and secondary schools in England. *Journal of Research in Science Teaching,* 49(8), 1035-1055.

Cleland, C.E. (2002) Methodological and epistemic differences between historical science and experimental science. *Philosophy of Science,* 69(3), 447-451.

Department for Education (2014) *Science Programmes of Study: Key Stage 4.* Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/381380/Science_KS4_PoS_7_November_2014.pdf

Glaesser, J., Gott, R., Roberts, R. & Cooper, B. (2009) Underlying success in open-ended investigations in science: using qualitative comparative analysis to identify necessary and sufficient conditions. *Research in Science and Technological Education*, 27(1), 5-30.

Gott, R., Duggan, S., & Johnson, P. (1999) What do practising applied scientists do and what are the implications for science education? *Research in Science and Technological Education,* 17(1), 97-107.

Gott, R., Duggan, S., Roberts, R., & Hussain, A. (n.d.) Research into understanding scientific evidence. Retrieved from http://community.dur.ac.uk/rosalyn.roberts/Evidence/cofev.html.

Gott, R., Foulds, K., Johnson, P., Jones, M. and Roberts, R. (1997*) Science Investigations 1.* Collins Educational: London

Gott, R., Foulds, K., Jones, M., Johnson, P. and Roberts, R. (1998) *Science Investigations 2.* Collins Educational: London.

Gott, R., Foulds, K., Roberts, R., Jones, M. and Johnson, P. (1999) *Science Investigations 3.* Collins Educational: London.

Gray, R. (2014) The distinction between experimental and historical sciences as a framework for improving classroom inquiry. *Science Education,* 98(2), 327-341.

Haigh, M., France, B., & Gounder, R. (2012) Compounding confusion? When illustrative practical work falls short of its purpose - A case study. *Research in Science Education, 42*(5), 967–984.

Hodson, D. (1992) Redefining and reorientating practical work in school science. *School Science Review*, 73(264), 65–78.

Johnson, P. & Roberts, R. (2016) A concept map for understanding 'working scientifically'. *School Science Review,* 97(360), 21-28.

Jones, L. R., Wheeler, G., & Centurino, V. A. S. (2013) 'TIMSS 2015 Science Framework.' In I.V.S. Mullis, & M.O. Martin, (Eds.). *TIMSS 2015 Assessment Frameworks* (Chp 2). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://timssand-pirls.bc.edu/timss2015/frameworks.html

Kinchin, I. M., & Cabot, L. B. (2010) Reconsidering the dimensions of expertise: From linear stages towards dual processing. *London Review of Education*, 8(2), 153–166.

Lederman, J. S., Lederman, N. G., Bartos, S. A., Bartels, S. L., Meyer, A. A., & Schwartz, R. S. (2014) Meaningful assessment of learners' understandings about Scientific Inquiry: The Views About Scientific Inquiry (VASI) questionnaire. *Journal of Research in Science Teaching*, 51(1), 65-83.

Lunetta, V. N., Hofstein, A., & Clough, M. P. (2007) 'Learning and teaching in the school science laboratory: An analysis of research, theory, and practice.' In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 393–441). Mahwah, NJ: Lawrence Erlbaum.

Millar, R. (1998) 'Rhetoric and reality: What practical work in science education is really for'. In J. Wellington (Ed.), *Practical work in school science: Which way now?* London: Routledge.

Millar, R., & Osborne, J. (1998) *Beyond 2000: Science education for the future: A report with ten recommendations.* London: King's College London.

Novak, J. D., & Cañas, A. J. (2007) Theoretical origins of concept maps, how to construct them and use them in education. *Reflecting Education*, 3(1), 29-42.

Organisation for Economic Co-operation and Development. [OECD] (2013) *PISA 2015: Draft science framework.* Paris: Author.

Roberts, D. A. (2011) 'Competing visions of scientific literacy: The influence of science curriculum policy image.' In C. Linder, L. Ostman, D. A. Roberts, P -O. Wickman, G. Erickson, & A. MacKinnon (Eds.), *Exploring the landscapes of scientific literacy* (pp. 11-27). Abingdon: Routledge.

Roberts, R. (2001) Procedural Understanding in Biology: the 'thinking behind the doing'. *Journal of Biological Education,* 35(3), 113-117

Roberts, R. (2016) Understanding the validity of data: a knowledge-based network underlying research expertise in scientific disciplines. *Higher Education*, 72(5), 651-668.

Roberts, R., & Gott, R. (1999) Procedural understanding: Its place in the biology curriculum. *School Science Review,* 81(294), 19-25.

Roberts, R. and Gott, R. (2002) 'Collecting and Using Evidence'. In Sang, D. and Wood-Robinson, V. (Eds.) *Teaching Secondary Scientific Enquiry*, ASE/John Murray: London.

Roberts, R. & Gott, R. (2006) Assessment of performance in practical science and pupil attributes. *Assessment in Education,* 13(1), 45-67.

Roberts, R., Gott, R. & Glaesser, J. (2010) Students' approaches to open-ended science investigation: the importance of substantive and procedural understanding. *Research Papers in Education*, 25(4), 377-407.

Roberts, R. & Johnson, P. (2015) Understanding the quality of data: a concept map for 'the thinking behind the doing' in scientific practice, *The Curriculum Journal*, 26(3), 345-369.

Roberts, R & Reading, C. (2015) The practical work challenge: incorporating the explicit teaching of evidence in subject content, *School Science Review,* June 2015, 96(357), 31- 39.

Taylor, C. E., & Meyer, J.H.F. (2010) 'The testable hypothesis as a threshold concept for biology students.' In J. H. F. Meyer, R. Land, & C. Baillie (Eds.). *Threshold Concepts and transformational learning* (pp. 179-192). Rotterdam: Sense Publishers.

Tytler, R., Duggan, S. & Gott, R. (2001) Dimensions of evidence, the public understanding of science and science education. *International Journal of Science Education,* 23, 815–32.

Wilson, A., Åkerlind, G., Francis, P., Kirkup, L., McKenzie, J., Pearce, D., & Sharma, M. D. (2010) Measurement uncertainty as a threshold concept in physics. *Proceedings of The Australian Conference on Science and Mathematics Education (formerly UniServe Science Conference).*16, 98-103.

Views of students in agriculture about the issue of 'pest animals': wolves in France

Laurence Simonneaux[1] and Jean Simonneaux[1]

laurence.simonneaux@educagri.fr

Abstract

This research is part of the CASSIS (Communication About Socio-Scientific Issues) project which is a collaborative research project between four universities in France, England (2) and New Zealand. This project is focused on a particular Socially Acute Question (SAQ) – the problem of controlling 'pest animals' when they influence farming practices. In France, we analysed the views of students from different kinds of agricultural training on the issue of the wolf. In particular, we analysed the registers they mobilize (cognitive or emotive or both), their engagement in the issue and their ethical position. Different viewpoints within and between classrooms (depending on location and type of training) were expressed. There were always potential conflicts in the classrooms.

Keywords: ethical position, pest animal, socially acute question, wolf

---

[1] UMR EFTS, ENSFEA, University of Toulouse, France

1. Introduction

This research is part of the CASSIS (Communication About Socio-Scientific Issues) project which includes four universities in France, England (2) and New Zealand. The overall aim of this project is to help people with opposing views to develop a collaborative reflection and to improve their reasoning by integrating wider perspectives. The context is the potential conflict that can arise when people's views about agriculture, sustainable practices, animal welfare and conservation intersect. This project is focused on a particular issue – the problem of controlling 'pest animals' when they influence farming practices.

In France, the problem with the wolf is not the result of re-introduction, but instead of their migration into agricultural areas from Italy. After previously disappearing, they are now present in more and more regions of France. They are seen as predators of sheep but they also produce ecosystem diversity. Although wolves are protected species, there is a growing demand for them to be excluded from mountain pastures and natural parks or even eradicated.

2. Research questions

The issue of control of wild animals is of interest in biology education in the topic of biodiversity conservation. In previous research (Simonneaux & Simonneaux, 2009), we saw that students' reasoning about 'pest animals' varies because of their emotional proximity with Socially Acute Questions (SAQs), and their socio-cultural origin. We observed that the greater the proximity between the question considered and the students (a local issue in which they are implicated because of their socio-cultural origins) the lower the high-level thinking, decision-making, critical thinking and reasoning are.

If the situation presented to the students contradicts their system of values, the effect can hinder critical reasoning (effectively 'blind them') and build resistance; if, however, it allows them to defend socio-cultural positions, it stimulates critical analysis. We consider that students' positions could influence their reasoning and their adherence to collaborative reflections.

Our main research question is: *Does the socio-professional background of agricultural students dominate over their position with regard to the wolf?*

Knowing students' positions should help to conceive didactic strategies to improve their collaborative reflection with people of diverging positions and

to improve their reasoning, integrating wider perspectives of the issue. The sub-questions are: *What are the views of students from different kinds of agricultural training on the issue of the wolf? In particular, what register do they mobilize (cognitive or emotive or both)? Who are the actants they recognise? How are students engaged in the issue? What is their ethical position?*

3. The wolf Socially Acute Question

We have previously identified the issue as a Socially Acute Question (SAQ) (Legardez & Simonneaux, 2006). Rather than referring to this issue as a socio-scientific issue, we believe that the term SAQ more accurately reflects the complexity of this issue at the public representation phase where risk, patterns of political and economic government and the notion of taking action are central (Simonneaux et al., 2013), as this approach better accommodates a discussion of how socially robust knowledge could be identified and/or critiqued.

At the end of the 18th century, there were 3,000 to 7,000 wolves in France. They were everywhere from the sea to the high mountains. After an organized eradication, the species had disappeared in the early 1930s. The first wolves naturally returned to France in 1992 from Italy. They have been a protected species since 1993 by the international Berne Convention. The population continues to grow. Now the estimated wolf population is around 300 animals. The wolves eat wild game but also attack flocks of sheep. In 2015, they killed nearly 9,000 sheep in France for food. If sheep are killed by predation, farmers are compensated by the State. The sheep population in France is about 6 million, which means that France is 51% self-sufficient in sheep meat. In neighboring countries (Italy and Spain), there are more wolves. In Italy there are 7 million sheep (67% self-sufficiency) and 600-900 wolves; in Spain, 14 million sheep (116% self-sufficiency) and 2,000 wolves. Sheep attacks by wolves are psychologically very hard to bear for breeders. Sheep farming is an economic issue but also an environmental and tourism issue through landscape maintenance in mountain areas. The migration and development of the wolf population in these areas has occurred because of several factors:

- Legal protection of the species;
- The reintroduction of wild ungulates (ibex) by hunters;

- Increasing populations of deer and wild boar;
- The reduction of farms. The area under plantations has expanded.

Since the return of the wolf, traditional husbandry methods have been questioned. It is recommended to have "patous" dogs, that is, big white dogs trained to be emotionally attached to the sheep, and living with sheep so they can scare the wolves. These dogs can ward off single wolves, but a pack of wolves can defy them. The reduction of 40% of sheep meat consumption in France during the last 20 years mean that mountain farms can be maintained only because of subsidies which account on average for 2/3 of the income of farmers. These economic constraints have resulted in an increase in herd sizes and reduced guarding by shepherds, and reduced deterrent for wolves. In areas where the wolf is present, the State subsidizes various protective devices and the use of shepherds, but not all measures prove satisfactory. The wolf appears to adapt to new practices that are implemented. In 2015, the State compensated the farmers to the tune of approximately € 2.5 million for this damage.

Although the species is protected, to limit the damage on livestock the State set a maximum quota of 36 wolves that could be shot between July 2015 and the end of June 2016. Some researchers argue that the shooting of wolves only shifts the problem or accelerates the spread of wolves by causing a breakdown of the packs. Different solutions are being considered by pro or anti-wolf groups:

- The population control by annual culling of a fixed number of wolves;
- The protection of the wolf;
- The eradication of the wolf.

4. Analytical framework to analyse answers to a 'photo questionnaire'

The photo questionnaire was composed of 4 pages with photos and an open question which provided an opportunity for students to reflect on their thinking. These questions were organized as follows: page 1 - an agricultural landscape; page 2 - a "natural" landscape, i.e. a forest; page 3 - a wolf; and page 4 - two photos of public demonstrations where participants are for and against wolves. Each page was presented separately so participants were unaware until

the third photo that the questionnaire was about 'the wolf reappearance controversy', see Appendix.

The students were from different areas of France (where wolves are present or not) and were trained in different types of agricultural training (Table 1 and Figure 1). They were between 16 and 18 years old (71 female and 101 male) and they had not received education on SAQ. This communication considers specifically the answers to page 3 (the wolf photo, see Appendix I) and the following question. Do students make connections to nature or to the controversy, or both? We assume that if participants write about the controversy it demonstrates that they are concerned.
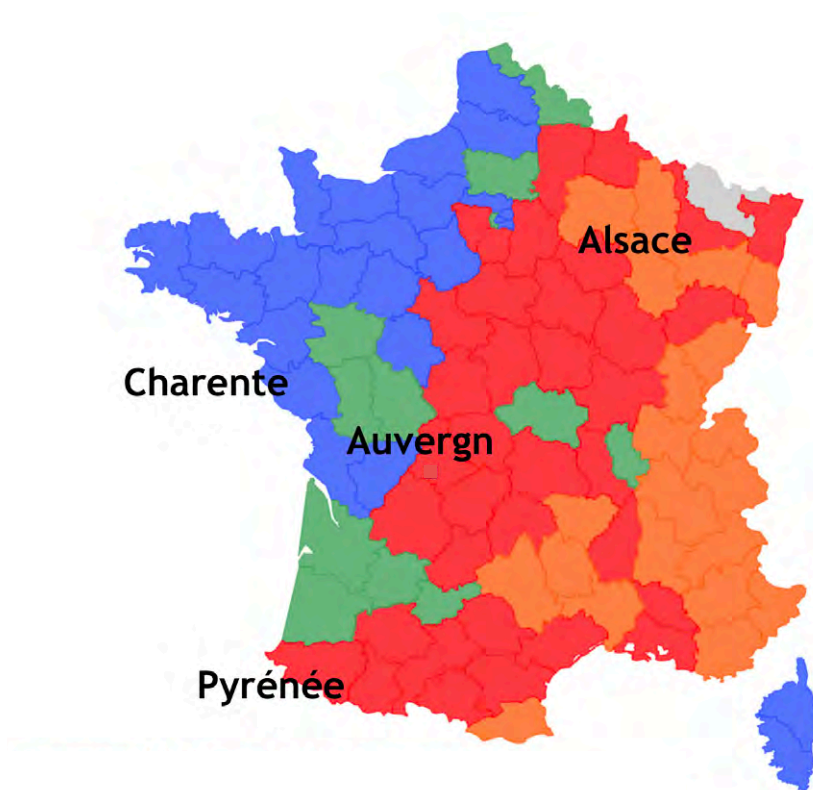
Figure 1 Repartition of wolves in France (Observatoire du loup - 08/2016) and locality of the school

Table 1 Division of the students from different areas.

Area	Auvergne (presence of wolves)		Pyrénées (presence of bears)	Charentes (no wolf in the area)			Alsace (presence of wolves)	
Type of training	Dog breeding training	Dog breeding training	Cattle production training	Plant production training	Cattle production training	Cattle production Training	Landscaping training	General and technology training
Number of students	N = 22	N = 23	N = 26	N = 10	N = 23	N = 9	N = 27	N = 32

.

Analytical step A: For the analysis of the answers to question 3, we used the nine attitude scales of attitudes towards animals developed by Kellert and Berry (1980). Their original scales consisted of the following attitudes: naturalistic (interest and affection for wildlife and the outdoors), ecologistic (concern for the environment as a system), humanistic (interest and strong affection for individual animals particularly pets), moralistic (concern for the right and wrong treatment of animals), scientistic (interest in the physical attributes and biological functioning of animals), esthetic (interest in the artistic and symbolic characteristics of animals), utilitarian (concern for the practical and material value of animals), dominionistic (interest in the mastery and control of animals typically in sporting situations), and negativistic (avoidance of animals due to dislike, indifference, or fear).

Analytical step B: In the cognitive register, we analysed the kind of reference they mobilized (professional, ecology, economy, farming), and we analysed how students who expressed an ethical position are situated between anthropocentrism, biocentrism and ecocentrism. Biocentrism is a strand in environmental ethics in which all living beings should be regarded as ends in themselves. Ecocentrism is an expansion of moral consideration to the non-living elements of nature, that is to say a biospheric egalitarianism in which species, communities, ecosystems have intrinsic value. Humans are part of the biotic community. Anthropocentrism separates humans from nature; humans are the undisputed master or guardian of nature; in this context the farmer is given this prevalent role.

Analytical step C: We used the Actor-Network Theory (ANT) (Callon, 1990; Latour, 1999; 2007). These French sociologists have developed and

grounded ANT in science and technology studies, so this theory is able to accommodate not just the technical and scientific aspects of an issue, but also allows space to explore the effect of the political, legal, moral, ethical and other issues that need to be considered (Latour, 2007, p.251). What is distinctive about ANT is that, as well as human beings, it treats non-human beings and objects as part of social networks. Consequently the participants - called actants - can be human or non-human, and all can be given an identity and a 'voice'. This theory emphasises that there is no hierarchy of actants and assumes that all entities in the network can be described in the same terms – referred to as generalised symmetry (Latour, 2007). He postulates that the shape of the network is determined by actants and their interactions. It is argued that when one is considering the performative actions of the non-human actants, their 'voice' is 'heard' through the actions/statements of the human actants. For example when considering the wolf issue, it could be argued that this non-human actant is central to the issue, and its presence determines not only the presence/actions of other non-human actants (e.g. ibex, sheep, dog) but also the purpose and focus of the human actants groups' performance – that is their opposition or acceptance of the presence of the wolf.

Analytical step D: We also used the work of the linguist Plantin (2011) to analyse emotions linked to argumentation. Participants may show their emotions in different ways. Plantin distinguishes two structural axes along which emotions are constructed in discourse: the axis of approval (pleasure-displeasure) and the axis of the intensity (strong emotion - weak emotion). In terms of approval, everything related to life is generally viewed positively, and things related to the death negatively. Here it can be the life or the death of wolves or sheep.

Co-authors separately analysed the answers and were able to come to agreement on all of the categories that emerged from the data.

5. Results

We could distinguish different student positions: protection of wolves, control of wolf population, eradication of wolves, stance against anti-wolf positions, stance against poachers.

Students involved with dog training mainly associated the photo of the wolf with Nature, whereas the others connected it also with breeding and predation, or farmers' economic problems.

Examples of association with Nature by dog training students:

This photo reminds me of the wilderness.

This picture evokes a natural environment in the wild where animals live peacefully.

Examples of association with wolf by dog training students:

It's a wolf. We must protect the wolves, there are fewer and fewer, either because of a lack of food or because of hunting and poachers who kill the wolves! The wolf is not a dangerous animal. It is afraid of humans, we must remove this opinion in the minds of idiots! Maybe even break them because they make me mad!

We see here that these remarks can be very emotionally charged from the start.

This image shows a beautiful wolf; its hierarchy is perfect and just. He is the ancestor of the dog by human intervention. It's my favorite animal. It is noble and royal and he was unjustly persecuted by humans though he was afraid of them.

Examples of association with wolf and predation, by other students:

There were also very emotionally charged points of view. They may include misconceptions: rapid breeding of wolves, attack of cows or people.

It's a protected species that is useless. It should be hunted as it kills the sheep, which can create a sudden loss. We don't produce animals to be eaten by the wolf! Rapid reproduction which creates migration to northern France. They hunt in packs and therefore can kill cows and can happen to kill people. ERADICATE THEM!

The **actants** that were considered by the dog training students were wolves, dogs and poachers, whereas other groups identified mainly farmers and sheep. When analysing the photos on pages 3 and 4, some participants expressed strong **emotions.**

"The bloody greens are useless, they want to protect this thing! Bullshit! One can see that they are not farmers and do not have children They live in cities and don't see what is happening around. They don't see beyond the tip of their nose. I will put wolves in the city, they will be scared. ERADICATION of WOLF"

"Should not KILL the wolves! We should rather kill these bastards that kill them. Shoot them in the ass then they will understand the pain! Is there something prestigious about killing a wolf at 200 meters?"

Concerning students' **attitudes** toward wolves (Table 2), students in two areas where the wolf is present, Auvergne and Alsace, involved in dog training and general technology training, expressed more naturalistic attitudes than the other students. The symbolic characteristics of wolves in the aesthetic attitude were linked to the virtues associated with wolves and different qualifiers were stated. In all areas, the students stated aesthetic attitudes towards wolves and attributed them virtues (e.g. wolves are strong, brave, powerful, fantastic, awesome, proud, intelligent, elegant, brave, fascinating, etc.). Considering the moralistic attitude, we distinguished whether students were concerned about the wrong treatment of farm animals (predation by wolves) or wrong treatment of wolves (culling, eradication). Moralistic attitudes (for the wolf protection or against the sheep predation) were voiced in every group. Ecologistic, humanistic, utilitarian and scientific attitudes were not articulated very much. Six students in dog training wrote about scientific considerations (dogs come from wolves through genetic selection), and two in general technology training wrote about pollution, species disappearance and climate change. No student had a dominionistic attitude towards the wolf, and only two communicated their negativistic attitude (fear or dislike).

Students in dog training strongly supported the protection of wolves. Students from the Pyrénées who were trained in cattle breeding had different positions in the same classroom (protection vs eradication). The students at Charentes and Alsace were mainly for the protection of wolves (though nine were for its control, two for the right to hunt them, and seven were strongly for their eradication). Some expressed acute opposition and accused ecologists or people from cities. Some believed that the wolf was reintroduced by humans, some of them advocated stopping the reintroduction or to reintroduce them in remote locations from the farms, for example in parks.

Considering **ethical** positions, dog training students and students in general technology training, and even cattle breeding students of Charentes were mostly biocentrist. Students from the Pyrénées were mostly anthropocentrist. In the other classrooms the two positions were equally represented. We did not find a gender difference.

Table 2 Students' attitudes towards wolves.

Attitudes	Auvergne		Pyrénées	Charentes			Alsace	
	Dog training	Dog training	Breeding 1ere	Plant production	Cattle production	Cattle production	Landscaping	General & technology
N	32	23	26	10	23	9	27	32
Naturalistic	17	9	5	4	7	1	5	19
Ecologistic	2	0	0	0	0	0	2	4
Humanistic (affection)	2	2	0	0	0	0	2	2
Moralistic Wolf protection	7	5	2	1	4	5	9	10
Moralistic Sheep predation	5	4	11	6	10	6	16	6
Scientistic	8 *(dog genetic selection)*	2 *(dog genetic selection)*	0	0	0	0	0	2 *(pollution, climate change, disappearance of species)*
Esthetic Virtues	13	15	9	2	15	2	15	22
Utilitarian	1	0	0	0	0	0	0	0
Dominionistic (on the contrary, it cannot be dominated)	0	0	0	0	0	0	0	0
Negativistic	0	0	0	1 (dislike)	1 (fear)	0	0	0

6. Conclusion

This study indicated different viewpoints within and between classrooms (depending on location and type of training). There were always potential conflicts in the classrooms. This analysis shows that we may have "to cool down or to heat up the didactical strategies" when setting situations for communi-

cation. We predict that when students are very concerned and express strong emotions, it is necessary to "cool down" the issue to help them to interact within the dialogue at least, and understand that there are different perspectives. If students do not know much about the issue, it is necessary to "heat up" the issue by emphasizing the controversies to help students to understand its complexity (Simonneaux, 2013). Since there are potential conflicts, to improve students' reasoning and reflection, we have decided to develop their distance and critical thinking by exchanging with students of different socio-cultural backgrounds. Thus we have chosen in the CASSIS project to cool down the issue of 'pest animals' (badgers accused of transmitting bovine tuberculosis, possums accused of destroying the environment and wolves accused of killing sheep), by proposing a strategy in three main steps: (i) construction of maps of controversies, (ii) interaction on blogs of English, New Zealander and French students, and (iii) review of maps of controversies.

References

Callon, M. (1990) Techno-economic networks and irreversibility. *The Sociological Review*, 38(S1), 132-161.

Kellert, S.R. & Berry, J.K. (1980) *Phase III: Knowledge, Affection and Basic Attitudes toward Animals in American Society*. Washington, DC: United States Government Printing Office.

Latour, B. (1999) *Pandora's hope. Essays on the reality of science studies*. Cambridge, MA: Harvard University Press.

Latour, B. (2007) *Reassembling the social. An introduction to Actor-Network-Theory*. New York, NY: Oxford University Press.

Legardez, A. & Simonneaux, L. (2006) *L'école à l'épreuve de l'actualité - Enseigner les questions socialement vives*. Issey-les-Moulineaux: ESF.

Plantin, C. (2011) *Les bonnes raisons des émotions : principes et méthode pour l'étude du discours* « émotionné ». Bern: Peter Lang.

Simonneaux, L. (2013) Questions socialement vives and socioscientific issues: New trends of research to meet the training needs of post-modern society. In C.Bruguière, A.Tiberghien & P. Clément (eds). *9th ESERA Conference Selected Contributions. Topics and trends in current science education*. Dordrecht: Springer, 37-54.

Simonneaux, L., Panissal, N. & Brossais, E. (2013) Students' perception of risk about nanotechnology after an SAQ teaching strategy. *International Journal of Science Education*, 35(14), 2376-2406.

Simonneaux, L. & Simonneaux, J. (2009) Students' socio-scientific reasoning on controversies from the viewpoint of Education for Sustainable Development. *Cultural Studies of Science Education*. 4, 3, 657-689.

Appendix – Photo pages from the photo questionnaire

3 – Pouvez-vous indiquer en quelques lignes ce qu'évoque pour vous cette photo ?

Photo page 3: Can you indicate in a few lines what this photo evokes for you?

4 – Pouvez-vous indiquer en quelques lignes ce qu'évoquent pour vous ces photos ?

Photo page 4: Can you indicate in a few lines what these photos evoke for you?

Chapter 4:
Teaching strategies and teaching enviroments

Acquiring diagnostic skills in the field of experimentation in the Wadden Sea's Teaching & Learning Laboratory

Lea Brauer[1] and Corinna Hößle[1]

lea.brauer@uni-oldenburg.de

Abstract

Diagnostic skills are gaining increasing importance in discussion about the professionalism of teachers in science education research. This paper follows the definition of Hößle (2014, p.1) that "diagnostic competence is […] the ability of teachers to make judgmental statements of the standing and processes of learning as well as characteristics of pupils. These statements are based on a specific and well-grounded question in a focused, theoretical, methodology controlled and reflected process." This study is focused on the diagnostic skills of 12 future teachers who took part in a 5-part diagnosis-oriented intervention: i) Participating in a theoretical seminar; ii) Designing teaching material; iii) Discussing their designs; iv) Practicing them in the "Wadden Sea´s Teaching & Learning Laboratory" (Teaching, Diagnosing teaching skills, and Diagnosing pupils´ experimentation skills); and v) Participation in a reflection seminar. The diagnostic skills of the future teachers were improved with written and filmed vignette tests in a pre-post-design. Following the post-test, the future teachers were interviewed about the procedure of the module and their diagnostic self-assessment, which offers a comparison to the real diagnostic skills of the future teachers. The results of the study show that the 5-part intervention had a positive effect on the diagnostic skills of the future teacher.

Keywords: diagnostic skills, pre-post-design, Wadden Sea´s Teaching & Learning Laboratory, written and filmed vignette tests

---

[1] Oldenburg University, Biology Didactics Division, Oldenburg, Germany

1. Introduction

Early PISA (Programme for International Student Assessment) results revealed that German pupils were ranking below the average in an international comparison, which caused the so-called "PISA shock". The issues "teaching professionalism" and "good teaching" have received pressing relevance since, and have been publicly discussed in subject-related didactics. On the basis of these results, professional capacity for action and the related professional knowledge of teachers are topical subjects in the discourse on quality development of our education system (Artelt & Gräsel, 2009; Baumert & Kunter, 2006; Hesse & Latzko, 2011). For an accurate diagnosis of pupils' learning processes, special skills are required that are called diagnostic skills, and serve to identify different learning difficulties and learning processes of pupils (Baumert & Kunter, 2006). These diagnostic skills have been declared to be a teacher's key competence (Baumert & Kunter, 2006; Dübbelde, 2013; Frey & Jung, 2011; Helmke, 2009; Hesse & Latzko, 2011; Schrader, 2008).

In 2004, the Conference of Ministers of Education and Cultural Affairs passed standards for teacher training setting definite targets for diagnostic training, thus taking up a clear position regarding the debate about teachers' professionalism (KMK, 2004). In 2008 this conference declared the states' common standards for specific science and didactics in teacher training valid, and thus confirmed the importance of diagnostic competences again (KMK, 2008). Hence, diagnostic competences are considered to be an important "adjusting screw" in teacher training (Hußmann & Selter, 2013), and attempts are made at enhancing diagnostic skills early, in order to impart far-sighted diagnostic skills to future teachers with regard to learning processes (Hößle, 2014).

2. Theoretical background

According to Hesse & Latzko (2011, p.25), diagnoses manifest themselves as "explicit statements about conditions, processes or traits of persons achieved via a reflected and methodically controlled process". Moreover, a diagnosis is characterized by the fact that an assessment is feasible with the help of existing categories, terms or concepts (Helmke, 2009). The present study supports this definition.

Different forms of diagnosing are distinguished. The formal (explicit, scientific) diagnosis "is achieved professionally, i.e., object-oriented, theo-

ry-controlled and systematically with scientifically verified methods" (Hesse & Latzko, 2011, p.25) and implies a communicative assessment level (Ingenkamp & Lissmann, 2008). The informal (implicit, everyday) diagnosis represents another form of diagnosis. It consists of subjective assessments and intuitive judgements that are made in everyday lessons mostly unconsciously (Hesse & Latzko, 2011, Helmke, 2009). Informal diagnoses usually lead to direct decisions (Schrader, 2011), because teachers continuously observe, judge and assess their pupils. According to Hesse (2014), implicit diagnoses are not at all wrong in everyday lessons, as long as the teacher is aware of the deficient accuracy of such diagnoses. In order to accurately assess pupils, diagnostic skills are required besides theoretical diagnostic knowledge. Furthermore, knowledge of fundamentals, criteria, and instruments of pedagogical diagnosis (Helmke, 2009) as well as knowledge of degree and development of competences should be available. According to Helmke (2009), a well-aimed and scientifically oriented planning, organization, and reflection of teaching and learning processes, as well as their individual assessment and systematic evaluation, are crucial tasks of the teaching profession. According to Bromme (2008), teachers should acquire sound knowledge and skills quite early in their training which they then can apply and improve during their work. Therefore, the standards for teacher training require skills to be imparted that enable teachers to properly identify developmental levels, learning potentials, learning impairments, and learning progress, as well as to develop suitable supportive measures (KMK, 2014).

The term "diagnostic competence" is often found in specialized publications. However, the construct has not yet been defined unequivocally (Steffen, 2015). According to Helmke (2009, p. 121), diagnostic competence is the ability to "appropriately assess persons or groups of persons (e.g. classes) and to form exact diagnostic judgements". This definition is supported by Hesse & Latzko (2011) and Schrader (2011). Most didactics experts agree that besides the performance, the learning behaviour must also be diagnosed (Hößle, 2014; Ingenkamp & Lissmann, 2008; Schrader, 2011). The presented definitions of the term "diagnostic competence" show that it means competences in their entirety, which is why the term "diagnostic skills" is also often used (Hesse & Latzko, 2011; Steffen, 2015; Weinert, 2002). Although the term "diagnostic competence" is frequently found in the literature, it is not as yet based on a scientific model describing its successive development (Steffen, 2015). Therefore, only the term "diagnostic skills" is used in the present study.

3. Research design and method

The objective of this research project is to understand how teaching in the "Wadden Sea's Teaching & Learning Laboratory" influences the diagnostic skills of 12 future teachers. Therefore, the research question of the present study is:

> How do diagnostic skills of prospective teachers improve by their active participation in the "Wadden Sea's Teaching & Learning Laboratory" with regard to diagnosing their pupils' experimentation skills and difficulties?

3.1 Participants in the study

Twelve future teachers of biology (seven female and five male students) participated in the study. They were future teachers of a Master's programme and attended the seminar "Teaching and Learning in the Wadden Sea's Laboratory".

3.2 Design of the study

The present study is a purely qualitatively oriented study with the focus to access the diagnostic skills and self-assessments of the future teachers, and thus to get clues to the development of diagnostic skills and to any difficulties and their causes. This way, the difficulties occurring even following the seminar and their possible causes had to be recorded just like the influences improving the diagnostic skills. Because of the qualitative orientation of the present study, the filmed and written vignette tests and the semi-structured interviews were combined as methods for data acquisition. The instruments of data acquisition for the present study were applied in a pre-post procedure during the winter term 2015/2016, and are described in more detail in Figure 1.
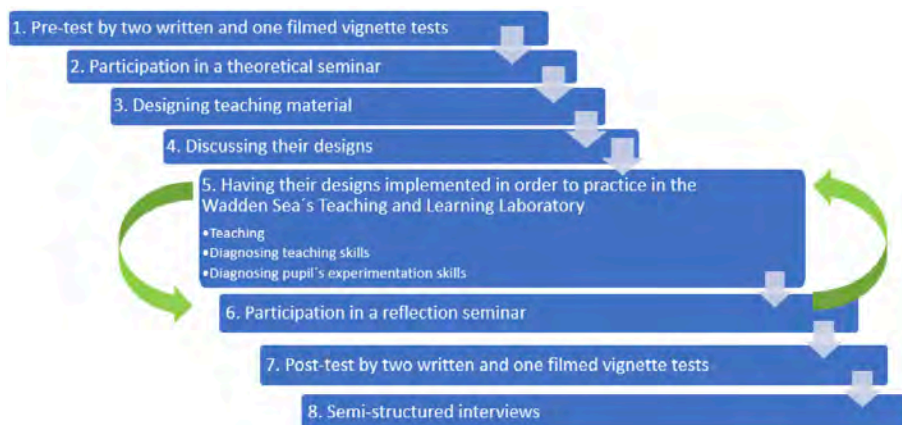
Figure 1 Structure of the study.

Pre-test

The pre-test served to determine the initial diagnostic skills of the participating future teachers with regard to pupils' experimentation skills. For this purpose, the following instruments were used: i) a **video-based vignette test** presenting the pupils during their experimental work; and ii) two **written vignette tests** presenting pupils' answers concerning experimental tasks that had to be diagnosed. The students were asked to diagnose the learning activities represented in the vignette tests in writing within the frame of an open assessment format following Baer & Buholzer (2005).

Participation in a theoretical seminar

The intervention phase of the study comprised a diagnostic seminar at Oldenburg University, including theoretical backgrounds such as definition of diagnosis, diagnosis processes, training diagnosis with examples of experimentation situations of pupils, experimentation processes and experimentation skills of pupils. In the theoretical part of the seminar, the fundamentals of hypothesis-generating as well as hypothesis-testing experimentation were imparted and fundamentals of diagnosis were dealt with.

Designing teaching material / Discussing their material

Following this theoretical introduction, the future teachers were asked to design 90-minute teaching material concerning the topic "Wadden Sea" in small groups of 2-3. Their task was to choose suitable experiments from an available pool of experiments, to reduce them didactically and to adapt them to the

requirements of different learning levels. The learning concepts designed by the future teachers were then discussed in plenary with regard to their methodological approaches.

Having their designs implemented in order to practice in the Wadden
Sea's Teaching and Learning Laboratory

Subsequently, the future teachers had the opportunity to test their learning concepts three times in the "Wadden Sea's Teaching & Learning Laboratory" according to the series presented below. During the practical phase, three different tasks were distributed:

a) one future teacher each accompanied the pupils during the lessons;

b) another future teacher diagnosed the teaching performance of the fellow-future teacher using an observation sheet;

c) a third future teacher diagnosed the pupils' learning performance by means of an individual diagnosis sheet.

The future teachers passed through this cycle three times, which enabled them to continuously exchange their roles and to test and adapt the learning sequence several times. Following each practical phase, the supervising university lecturers offered the future teachers individual discussions on the observed teaching skills.

Participation in a reflection seminar

A seminar session covering additional criterion-based reflections on the individual teaching activities in the small groups, as well as discussions on the diagnostic observations, always followed the teaching phase in the laboratory. To this end, the future teachers received a list of questions in advance, from which they had to select three questions per session. Priority was given to reflections on the individual teaching behaviour (e.g. *which learning objectives did I reach and how was this documented? How did I succeed in involving the pupils?*), and to a detailed treatment and analysis of the diagnosis sheet, which the future teachers had used to diagnose their pupils' learning processes. Then, the future teachers were asked to adapt their learning concepts to the diagnosed learning conditions of their pupils on the basis of the reflections.

Post-test

Following the intervention phase, a post-test was carried out corresponding to the pre-test in order to determine the diagnostic skills of the future teachers.

Semi-structured interviews

These interviews were performed in order to determine the future teachers` concepts and self-assessments, attitudes as well as difficulties and their causes with regard to diagnosing. The interviews also recorded the diagnostic self-concepts of the future teachers, based on the fact that according to Frey (2006, p.34) "*the respective person can provide the best information on the skills to be assessed.*"

3.3 Pre- and post-test instruments

The instruments of the study are described in more detail below:

Video-based vignette test

The video-based vignette test shows an experimental situation of four Year 9 pupils of an Intermediate Secondary School planning and performing an experiment concerning the question: *"Have snails got a sense of smell?"*

Written vignette tests

In addition, two written vignette tests were designed covering the topics "*Pitcher plants*" and "*Isolation*". These vignettes present the tasks completed by pupils in the eighth year with clearly diagnosable experimentation skills. As an example, the written vignette "*Pitcher plant*" is presented below (Figure 2). It displays the result produced by a pupil dealing with the topic "*Growth of a pitcher plant*". Among other tasks, the Year 8 pupil was asked to plan an experiment in order to investigate the optimum growth conditions for a pitcher plant. For this purpose, brief information text on the pitcher plant was given to the pupil and a material pool was made available as a stimulus for organizing the experiment. The formulated tasks covered mainly the partial competences "Make assumption / Propose hypothesis", "Identify / describe relations", and "Plan experiment" according to Nawrath et al. (2013).

Help to the pitcher plants

Linda bought two pitcher plants two weeks ago and put them on the window-sill in her room. Her plants are not doing well. The pitchers have been getting brownish and have been withering. Linda wonders why her pitcher plants do not grow well.

Pitcher plants belong to the carnivorous plants. The pitchers at the leaf tips are traps for flies. The insects are digested by means of a liquid and serve the plants as additional source of nutrients. Pitcher plants are indigenous to tropical mountainous regions and therefore prefer the following habitat conditions: high day-time temperatures of approximately 25-35 °C all year round, high air humidity (ca. 70 %), and humid soils without stagnant moisture.

Anna, 8th year of an Intermediate Secondary School, deals with the tasks as follows:

Tasks:
1. State the reasons why the pitcher plants cannot grow well.
Too cold, different climate, air humidity too low, not enough insects to feed, too much stagnant water
2. Choose one of the reasons stated above, which you would like to investigate.
Air humidity
3.1 Describe the relation between the withering of the plant and the reason you want to examine.
The plants' prosperity with regard to air humidity
3.2 Formulate an if ... then ... statement.
If the air humidity is too low for the plant, it cannot grow.

4. Now, please plan an experiment in order to verify your assumption. For this, you may take the necessary things from the material desk. Choose the things required for your experiment from the available materials.
First, draw a sketch of your experiment:

5. Describe your experimental set-up.

The pitcher plant is put into a glass case and I am spraying water into it at regular intervals. The hygrometer measures the air humidity and the heating lamp is used to warm up the air.

6. Give reasons for your choice of materials:

The glass case serves the purpose of insulation and of accumulating heat and air humidity. And the water is necessary for the air humidity. The hygrometer is used for measurement and verification. The heating lamp provides heat.

Available materials:

Pitcher plants   thermometer   various glass vessels   spray bottle

hygrometer   heating lamp   glass case   fertilizer

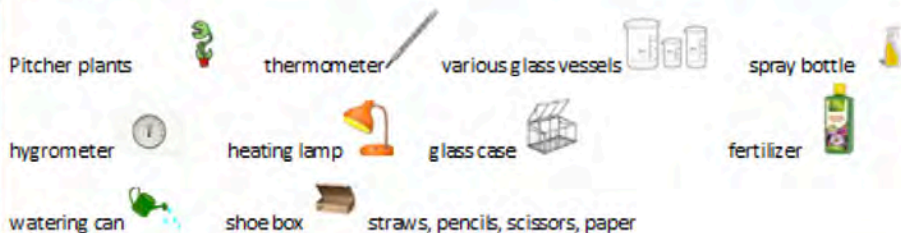watering can   shoe box   straws, pencils, scissors, paper

Figure 2 Written vignette: Pitcher plant.

3.4 Analysis procedure

A coding manual for experimentation skills was developed in accordance with the model established by Nawrath et al. (2013) in order to analyse the vignette tests. The coding manual consisted of categories obtained deductively and was supplemented inductively. The coding rules with standard examples enabled the treated vignettes to be analysed consistently, thus increasing the quality of the coding manual (Hamann & Jördens, 2014). See Table 1.

Table 1 Excerpt from the coding manual "Pitcher plant ".

Vignette		Partial competences		Coding rule with standard examples
Task	**Transcription**	**Partial competence**	**Traits to be diagnosed**	
1. State reasons why the pitchers' growth might be impaired	Too cold, different climate, air humidity too low, not enough insects to feed, too much stagnant water	Make an assumption / Propose hypothesis	- Identifies various potential causes from the pool of causes - Does not describe/explain the assumption	It must be noticed that causes are listed. Likewise, it must be stated that an explanation is lacking. Example of complete diagnosis: pupil states different causes, but does not consider them in more detail and does not notice the opposite cause, respectively.

The semi-structured interviews were analysed in accordance with the qualitative contextual analysis by Mayring (2015) using deductively and inductively obtained categories. For the final analysis, only those interviews were considered that yielded interesting insights in addition to the vignette tests.

In order to consider the criteria of qualitative research, the content validity was determined on the basis of expert assessments and verified by means of piloting. Piloting enabled the quality of vignettes and of the related diagnosis tasks to be verified. Unintelligible items were rephrased or removed.

Gradual levels were developed for each vignette for a definitive classification of the future teachers' diagnostic skills. These levels are based on the difficulties to be diagnosed and on features determined in the coding manual. Altogether, the gradual levels yield a good survey of the starting situation of the individual future teacher's diagnostic skills and additionally show their changes. As an example, the levels for the written vignette "Pitcher plant" are presented in Table 2.

Table 2 Gradual levels for the written vignette "Pitcher plant".

Partial skills in experimentation	Level 1	Level 2	Level 3
Make assumptions / Propose hypotheses	Does not identify assumptions and hypotheses made by pupil	Identifies the assumptions and hypotheses proposed by pupil. Does not consider the lacking arguments.	Identifies all assumptions and hypotheses proposed by pupil and considers the lacking arguments.
Identify and describe relations	Does not notice that pupil identifies a relation between the plant's prosperity and air humidity. Does not notice that pupil can formulate a respective cause-and-effect dependency.	Notices that pupil identifies a relation between two variables (plant's prosperity and air humidity).	Notices that pupil identifies a relation between two variables (plant's prosperity and air humidity). Notices that pupil can formulate a respective cause-and-effect dependency
Plan experiment	Does not notice that pupil has only minor experimentation skills in planning an experiment and has the following difficulties: -No control experiment -No comprehensible/ complete experimental plan -No complete inscription of sketch/incomplete sketch -Simultaneous testing of variables Does not notice that pupil can describe the experimental set-up in spite of incomplete sketch, chooses materials correctly and can explain her choice.	Notices that pupil has only minor experimentation skills in planning an experiment and has the following difficulties: -No control experiment -No comprehensible/ complete experimental plan -No complete inscription of sketch/incomplete sketch -Simultaneous testing of variables	Notices that pupil has only minor experimentation skills in planning an experiment and has the following difficulties: -No control experiment -No comprehensible/ complete experimental plan -No complete inscription of sketch/incomplete sketch -Simultaneous testing of variables Notices that pupil can describe the experimental set-up in spite of incomplete sketch, chooses materials correctly and can explain her choice.

The three-grade levels are based on the corresponding coding manual and on the theoretical fundamentals concerning the pupils' difficulties in experimentation (Hammann et al., 2006; Schmiemann & Mayer, 2013). On the basis of these fundamentals as well as by means of piloting, three levels have gained acceptance for all vignettes. In principle, level 1 means that the aspects to be diagnosed are not identified. Level 2 means that the aspects to be diagnosed are noticed by the test subject, however, mostly without detailed comment and diagnosis. The third level corresponds to a complete diagnosis of the vignettes. Since these levels cannot be considered to be static, further intermediate stages between the levels can be described. Thus, it is quite probable that students are found ranking between two levels. The results obtained for the

individual test subjects are presented in the form of a dartboard (Figure 3) in order to illustrate the changes occurring in single categories
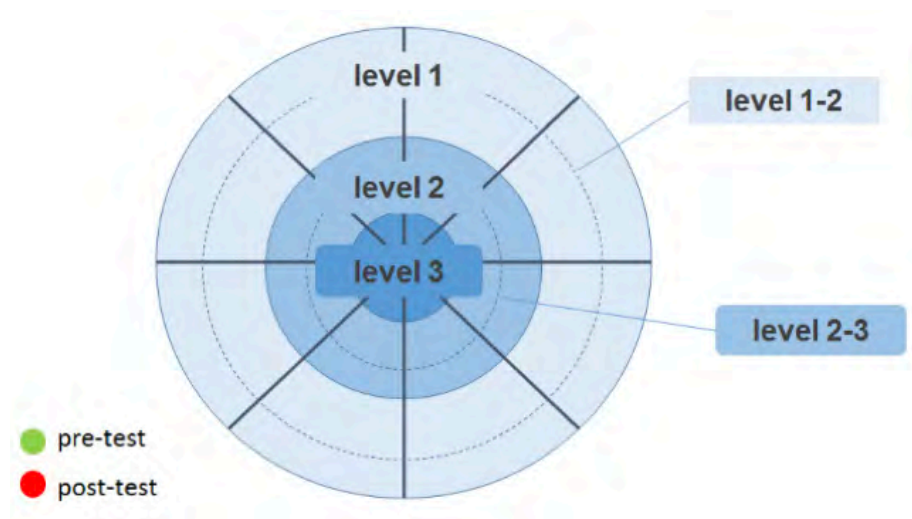
Figure 3 Description of dartboard.

4. Findings

4.1 Pre- and post-test

The pre-test clearly shows that future teachers have big difficulties in identifying pupils' typical learning difficulties during experimentation. For example, unsystematic approaches, omitted control tests, typical achievement of effects by manipulating the experiment, and interference in a running experiment by changing the variables (Hammann et al., 2006; Schmiemann & Mayer, 2013) are rarely noticed. Future teachers are seldom able to reconstruct the causes, pointing to their pupils' difficulties in experimentation. It is evident that the future teachers have difficulties in correctly describing their pupils' experimentation skills, which indicates that the future teachers are lacking skills in using the didactic vocabulary. Moreover, the group of future teachers was quite heterogeneous with different diagnostic skills due to greatly varying experiences previously gained during the course of their studies. The results of the post-test show an improvement of diagnostic skills allowing us to assume that the module with the 5-step phasing has a positive effect on the development

of diagnostic skills.

Some examples of the diagnosis of the written vignette "pitcher plant" in the category "plan experiment" are shown below:

The future teacher ANHA diagnosed in the pre-test: "*how is the humidity? And how much water will be splashed into the glass case?*" This diagnosis is positioned into the level 1, because she did not diagnose that the control experiment is missing, did not diagnose that the plan and the sketch of the experiment was not complete, and did not diagnose the simultaneous testing of variables. Furthermore, she did not notice that the pupil could describe the experimental set-up in spite of the incomplete sketch, chose materials correctly and could explain her choice. In the post-test she diagnosed: "*it is not distinguishable if the glass case is open or closed, how low should the humidity be? The reference to the optimum is not shown, when the pupil will measure? The lettering is not complete and the control of the experiment is missing.*" This diagnosis is placed into the level 2-3, because she was able to diagnose a lot of aspects of level 3. But as well in the post-test she did not notice that the pupil could describe the experimental set-up in spite of the incomplete sketch, chose materials correctly and could explain her choice.

This future teacher should learn to describe the diagnosis in more detail and think about reasons of the difficulties of the pupil.

The written vignette "pitcher plant" revealed that all future teachers improved in the category "Plan experiment" to be diagnosed. Furthermore, it is evident that a large number of future teachers improved from level 1-2 to the third level, seven future teachers reaching level 3 in the post-test, two future teachers improving to level 2-3, and one future teacher to level 2. In the post-test, none of the future teachers thus ranked below level 2 in the category "Plan experiment" to be diagnosed.

So, by the end of the semester, the future teachers had fewer difficulties in noticing that pupils often omitted the control test while planning their experiments. Likewise, the future teachers were considerably more frequently able to diagnose permanent interferences during experimentation as a learning difficulty in the post-test. Moreover, incomplete organization of experiments and unsystematic approaches during the experiments were noticed more often than in the pre-test.

4.2 Interviews

Some examples of the interviews concerning the question regarding the difficulties of diagnosing are shown below:

The future teacher KATH said: "*we tried [...] to find the mistakes, which the pupils made. But just in this lesson everything went right. And because of that, we found nothing*" (line 445-457). "*While pupils are planning an experiment, they have some ideas about why they are doing it that way and then we have to ask them again and again to explain their experimental work. And as a teacher you have always to ask them why they are doing this and that. [...] And the problem is, that the pupils don't write their thoughts down. It is really difficult to find out about their ideas.*"

Another future teacher ASAX said to this aspect: "*If you want to analyse pupil's conceptions, it is difficult to find out at this moment what conception it could be. And if you found out the pupil's conception how can you treat with it? [...] Should I comment that? Or ask some more questions or should I ignore that? It isn't easy to know how important that conception is at this moment and for me it is really difficult to find it out.*"

Regarding the diagnostic skills, the future teacher KATH said: "*And of course at this time I know more than during the time in the seminar, sure. But perhaps in an implicit way.*"

The female future teacher ANHA described her problems as follows: "*Well, sometimes, it already starts being difficult, when you are thinking you are now going to pay attention to pupils' concepts, to really notice them at that moment and to identify them. How can I then comment myself to them at the same instant?*" This test subject emphasized that she found it particularly hard to flexibly respond to diagnosed features during the lessons. This is also due to the fact that the future teachers are not really able to focus on their pupils while teaching them, because they are still highly focused on the course of the lesson.

For example, a female future teacher SIPE described her diagnostic skills as follows: "*Well, I surely have potential for improvement. The foundation has been laid, though. In my opinion, the practical part is tremendously important.*"

ASAX described her diagnostic skills as "*developable*". She also said: "*I don't feel uncertain, but I am not competent in diagnosing. So, I think I am in the middle.*"

The future teacher GIJO felt "*uncertain*" concerning diagnosing pupils' skills. She said: "*I always write down what seems to be important, but I don't know if it is relevant for the diagnosis.*"

In summary, the future teachers know about their difficulties in diagnosing pupils' experimentation skills and they can describe these difficulties. Regarding the self-concept of their diagnostic skills, most of them feel uncertain, but not bad. Some of them have the awareness that their diagnoses are implicit, not goal-oriented and not professional. But they feel more certain in diagnosing pupils' experimentation skills after finishing the module.

5. Discussion and conclusions

It is highly probable that participation in the seminar, which promoted the future teachers' teaching experiences as well as introducing them to the theoretical fundamentals of correct experimentation and diagnosis, had a positive influence on the development of their diagnostic skills. Nevertheless, the future teachers continued to have difficulties in diagnosing properties of pupils, for example, achievement of effects, and amalgamation of the single experimentation-steps, as well as lacking reasons for assumptions proposed. In addition, it is obvious that future teachers identified experimentation errors of pupils rather than existing experimentation skills, such as, e.g., independently organizing an experiment, correctly describing an experimental sketch or noticing own mistakes during experimentation. It is also evident that many future teachers merely made suggestions for improving the experimental work, thus not making diagnostic statements. The future teachers need pronounced experimentation knowledge and skills to establish better diagnoses, and they need more opportunities to train in their diagnostic skills.

Besides their difficulties in diagnosing their pupils' experimentation skills, future teachers also had difficulties in identifying certain concepts among their pupils in the vignette tests. Although pupils' concepts were treated as a crucial point in the seminar, the future teachers were rarely able to identify them in the video-based vignette during the post-test. This result is supported by the interviews, in which the future teachers emphasized that they continued to be unsure how to accurately diagnose learning processes and to identify pupils' concepts, in part even after attending the module. Nevertheless, the future teachers considered their skills in diagnosing learning processes to be acceptable but in need of improvement. However, one crucial point is that the future teachers became aware of the importance of considering the diagnosis.

These results reveal that the presented seminar actually promoted the

diagnostic skills of future teachers. However, additional courses and advanced training are required to qualify students to diagnose learning processes reliably.

References

Artelt, C. & Gräsel, C. (2009) Diagnostische Kompetenz von Lehrkräften – Gasteditorial. *Zeitschrift für pädagogische Psychologie*, *23* (3-4), 157-160.

Baer, M. & Buholzer, A. (2005) Analyse der Wirksamkeit der berufsorientierten Ausbildung für den Erwerb von Unterrichts- und Diagnosekompetenzen. *Beiträge zur Lehrerbildung 2*, 243-248.

Baumert, J. & Kunter, M. (2006) Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9* (4), 469-520.

Bortz, J. & Döring, N. (2006) Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Springer, Heidelberg.

Bromme, R. (2008) Lehrerexpertise. In: Schneider, W. & Hasselhorn, M. [Hrsg.]: *Handbuch der pädagogischen Psychologie.* 159-167. Hogrefe, Göttingen.

Dübbelde, G. (2013) Diagnostische Kompetenzen angehender Biologie-Lehrkräfte im Bereich der naturwissenschaftlichen Erkenntnisgewinnung. Available on: https://kobra.bibliothek.uni-kassel.de/bitstream/urn:nbn:de:hebis:34-2013122044701/3/DissertationGabiDuebbelde.pdf

Frey, A. (2006) Methoden und Instrumente zur Diagnose beruflicher Kompetenzen von Lehrkräften – eine erste Standortbestimmung zu bereits publizierten Instrumenten. *Zeitschrift für Pädagogik. 51.* Beiheft. S.30-46.

Frey, A. & Jung, C. (2011) Kompetenzmodelle und Standards in der Lehrerbildung. In: E. Terhart, H. Bennewitz & M. Rothland [Hrsg.]: *Handbuch der Forschung zum Lehrerberuf* , 540-572. Waxmann, Berlin.

Friebertshäuser, B. & Langer, A. (2013) Interviewformen und Interviewpraxis. In: B. Friebertshäuser, A. Langer & A. Prengel [Hrsg.]: *Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft.* Beltz Juventa, Weinheim.

Grube, Chr. & Mayer, J. (2010) Wissenschaftsmethodische Kompetenzen in der Sekundarstufe I: eine Untersuchung zur Entwicklung des wissenschaftlichen Denkens. In: U. Harms & I. Mackensen-Friedrichs [Hrsg.]: *Lehr- und Lernforschung in der Biologiedidaktik (Band 4), Heterogenität erfassen – individuell fördern im Biologieunterricht*, 155-168. Studien Verlag, Innsbruck.

Hammann, M. (2004) Kompetenzentwicklungsmodelle. *Der mathematische und naturwissenschaftliche Unterricht, 57* (2004), 4, 196-203.

Hammann, M. & Jördens, J. (2014) Offene Aufgaben codieren. In: D. Krüger , I. Parchmann & H. Schecker [Hrsg.]: *Methoden in der naturwissenschaftsdidaktischen Forschung.* Springer Spektrum, Heidelberg.

Hammann, M; Thi, T.H.P. & Ehmer, M. (2006) Fehlerfrei Experimentieren. *Der mathematische und naturwissenschaftliche Unterricht, 59* (2006), 5, 292-299.

Helmke, A. (2009) Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. Kallmeyer / Klett, Seelze-Velber.

Hesse, I. (2014) Pädagogisch-psychologische Diagnostik für Lehrkräfte – Herausforderung, Aufgaben, Probleme. In: A. Fischer, C. Hoessle, S. Jahnke-Klein, H. Kiper, M. Komorek, J. Michaelis, ... J. Suits [Hrsg.]: *Diagnostik für lernwirksamen Unterricht*, 15-39. Schneider, Hohengehren.

Hesse, I. & Latzko, B. (2011) Diagnostik für Lehrkräfte. Barbara Budrich, Opladen.

Hößle, C. (2014): Lernprozesse im Lehr-Lern-Labor Wattenmeer diagnostizieren und fördern. In: A. Fischer, C. Hößle, S. Jahnke-Klein, H. Kiper, M. Komorek, J. Michaelis, … J. Suits [Hrsg.]: *Diagnostik für lernwirksamen Unterricht,* 144-156. Schneider, Hohengehren.

Hußmann, S. & Selter, C. (2013) Das Projekt dortMINT. In: S. Hußmann & C. Selter [Hrsg.]: *Diagnose und individuelle Förderung in der MINT-Lehrerbildung. Das Projekt dortMINT*, 15-27. Waxmann, Münster.

Ingenkamp K.-H. & Lissmann, U. (2008) Lehrbuch der Pädagogischen Diagnostik. Beltz, Basel.

KMK (2004) Standards für die Lehrerbildung – Bildungswissenschaften. Aktualisiert 12.06.2014. Available on: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf

KMK (2008) Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung. Aktualisiert 12.02.2015. Verfügbar unter: http://www.kmk.org/bildung-schule/allgemeine-bildung/lehrer/lehrerbildung.html

Lamnek, S. (2010): *Qualitative Sozialforschung.* Beltz, Basel.

Mayer, J. (2007) Erkenntnisgewinnung als wissenschaftliches Problemlösen. In: D. Krüger & H. Vogt [Hrsg.]: *Handbuch der Theorien in der biologiedidaktischen Forschung – Ein Handbuch für Lehramtsstudenten und Doktoranden,* 178–186. Springer, Berlin.

Mayring, P. (2015) Qualitative Inhaltsanalyse. Grundlagen und Techniken. Beltz, Basel.

Nawrath, D., Maiseyenka, V. & Schecker, H. (2013): Experimentierfähigkeit. In: H. Schecker, D. Nawrath, H. Elvers, J. Borgstädt, S. Einfeldt & V. Maiseyenka [Hrsg.]: *Modelle und Lernarrangements für die Förderung naturwissenschaftlicher Kompetenzen,* 8-18. Behörde für Arbeit, Soziales Familie und Integration, Hamburg.

Pawek, C. (2009) Schülerlabore als interessenfördernde außerschulische Lernumgebungen für Schülerinnen und Schüler. Dissertation Universität Kiel. Available on: http://www.dlr.de/schoollab/Portaldata/24/Resources/dokumente/Diss_Pawek.pdf [Abrufdatum 19.12.2014].

Rehm, M. & Bölsterli, K. (2014) Entwicklung von Unterrichtsvignetten. In: D. Krüger, I. Parchmann & H. Schecker [Hrsg.]: *Methoden in der naturwissenschaftsdidaktischen Forschung*, 213-225. Springer Spektrum, Heidelberg.

Schmiemann, P. & Mayer, J. (2013) Experimentieren Sie! Biologieunterricht mit Aha-Effekt. Selbstständiges, kompetenzorientiertes Erarbeiten von Lehrplaninhalten. Cornelsen, Berlin.

Schrader, F.-W. (2008) Diagnoseleistungen und diagnostische Kompetenzen von Lehrkräften. In: M. Schneider [Hrsg.]: *Handbuch der Pädagogischen Psychologie 10,* 168-177. Hogrefe, Göttingen.

Schrader, F.-W. (2011) Lehrer als Diagnostiker. In: E. Terhart, H. Bennewitz & M. Rothland [Hrsg.]: *Handbuch der Forschung zum Lehrerberuf,* 683-698. Waxmann, Berlin.

Steffen, B. (2015) Negiertes Bewältigen. Eine Grounded-Theory-Studie zur Diagnose von Bewertungskompetenz durch Biologielehrkräfte. Logos, Berlin.

Weinert, F.E. (2002) Leistungsmessungen in Schulen. Beltz, Weinheim.

Weinert, F.E. & Schrader, F.-W. (1986) Diagnose des Lehrers als Diagnostiker. In: H. Petillon, J.W.L. Wagner & B. Wolf [Hrsg.]: *Schülergerechte Diagnose*, 11-29. Beltz, Weinheim.

Wellnitz, N. & Meyer, J. (2012) Beobachten, Vergleichen und Experimentieren: Wege der Erkenntnisgewinnung. In: U. Harms & F.X. Bogner [Hrsg.]: *Lehr- und Lernforschung in der Biologiedidaktik*, Band 5, „Didaktik der Biologie – Standortbestimmung und Perspektiven". Internationale Tagung der Fachsektion Didaktik der Biologie im VBIO, Bayreuth 2011. Studienverlag, Innsbruck.

Zehren, W. & Hempelmann, R. (2014) Kognitive und motivationale Effekte durch regelmäßiges Forschendes Experimentieren im Schu lerlabor. *LeLa Magazin, 9*, 8-10.

Retrieval-based learning in the context of inquiry-based learning

Anne Cohonner[1] and Jürgen Mayer[1]

anne.cohonner@uni-kassel.de

Abstract

In the field of science education, inquiry-based learning is advocated to be a promising method for teaching science concepts. Its benefits with regard to retaining content knowledge seem to be particularly unclear, and suitable means for ensuring retention have not yet been found. Research from cognitive psychology points to the benefit of retrieval-based learning for the retention of scientific content. With the purpose of improving science education, the present study aims to analyse the efficacy of this instructional technique for the retention of scientific concepts. Therefore, a study involving 6th and 7th graders ($N = 93$) was conducted. An inquiry activity which includes retrieval-based learning was compared with inquiry tasks that incorporate rereading or no extra guidance (control) to build up content knowledge about the biological concept of adaptation. Overall, retrieval-practice and rereading outperformed the control condition, indicating the need for embedding learning techniques that increase access to domain information during inquiry. However, retrieval practice and rereading did not differ significantly, although different types of learners benefited from the learning techniques on a long-term basis. The results are discussed concerning their theoretical relevance and practical implications.

Keywords: inquiry learning, learning strategy, memory, retrieval practice

---

[1] University of Kassel, Kassel, Germany

1. Introduction

To achieve long-term goals, learning environments in science education actively involve students in knowledge construction by implementing self-organized learning methods. Inquiry-based learning can been counted as such an approach by engaging students in activities used by scientists. Nevertheless, the efficacy of inquiry-based learning has been continuously questioned (Kirschner et al., 2006). Despite the extended body of research that has focused on enhancing processing during knowledge construction, suitable means to encourage learning of content knowledge effectively remain unclear. Referring to research from cognitive psychology, retrieving information during learning seems to be necessity to gain long-term learning outcomes (Roediger, & Karpicke, 2006). Practicing active retrieval by reconstructing knowledge from memory can be a mnemonic enhancer relative to rereading. In accordance with Karpicke (2012), we use the term *retrieval-based learning* to refer to both the underlying cognitive process and to learning activities (assessment strategies or tests; Schroeder et al., 2007) that encourage learners to engage in retrieving information. As current studies have found large learning gains when retrieval-based learning is used to teach scientific concepts, it can be assumed that retrieval processes during inquiry-based learning could be a helpful technique to enhance retention.

2. Theoretical framework

Altering long-term memory is the ultimate goal of all learning processes as well as of science education in particular. In this context, it seems to be beyond question that engaging learners actively in knowledge construction leads to more meaningful understanding. One approach in science education that coincides with this constructivist paradigm is inquiry-based learning. Learning science through inquiry is supposed to encourage learners to develop deeper understanding by actively engaging them in a scientific discovery process. In order to gain understanding during scientific discovery, students have to master two related problem spaces: a hypothesis space and an experiment space (Klahr, 2000). While the hypothesis space can be defined as a search space consisting of possible explanations for a phenomenon and is guided by prior knowledge of scientific concepts, the experiment space consists of experiments and requires inquiry skills for mastery (Klahr, 2000). In this way, inquiry-based

learning has a twofold significance for science teaching, as it can improve both inquiry skills and understanding of scientific concepts.

Even though inquiry-based learning is considered an important learning approach in science teaching, inconsistent findings about its effectiveness prevail in the research (Zhang, 2016). Evidence from educational research reveals that the acquisition of inquiry skill is to the detriment of content knowledge (e.g. Hof, 2011). This finding is supported by several meta-analyses pointing to higher effects for inquiry skills ($d = .52$, $d = .40$, $d = .78$) than for content knowledge ($d = .16$, $d = .20$, $d = .37$) (Bredderman, 1983; Shymansky et al., 1990; Lazonder & Harmsen, 2016). It is argued that students' cognitive capacities cannot be sufficiently used to focus on inquiry skills and content knowledge alike (Kirschner, et al., 2006). Novice learners in particular, with little prior knowledge, turn their whole attention towards the process of inquiry which leads to poor acquisition and retention of scientific concepts. As researchers aiming to help learners overcome these hurdles they show that optimal guidance seems to be the key. Although much effort has been made in the field of scientific reasoning, the acquisition of scientific content knowledge still remains deficient and suitable means for achieving long-term goals remain unclear (e.g. Arnold, 2015). These limitations of inquiry-based learning point to the need to focus on learning techniques that enhance retention of scientific content.

In accordance with Bjork & Bjork (2011), we argue that instead of exclusively focusing on scaffolds that encourage encoding during instruction, educational research has to concentrate on learning techniques that require retrieving knowledge if long-term learning outcomes are to be achieved. From this perspective, the act of retrieving information from memory (called retrieval-based learning) seems to be a more powerful learning activity than instructions that help encoding (Kapricke & Aue, 2015). Such retrieval processes are made manifest in assessment strategies like testing (Schroeder, et al., 2007). From an educational standpoint, however, assessment is seen as a neutral event, and is frequently applied in the form of high-stakes tests to evaluate learning outcomes. However, cognitive science research challenges this fundamental idea by showing that tests can be effective learning tools compared to rereading (for a review see Roediger, & Karpicke, 2006). Further, initial findings reveal large learning gains for scientific concepts (McDaniel et al., 2011).

Several possible explanations for why retrieval-based learning is such a robust mnemonic enhancer have been proposed, with the semantic elaboration account (Carpenter, 2009) and the episodic context account (Karpicke & Aue, 2015). Both accounts define the mechanism that exerts the mnemonic effect as an elaboration on existing memory and creation of additional retrieval routes. However, this positive effect greatly depends on students' ability to successfully recall correct content (Kornell et al., 2011). According to Kornell, et al.'s (2011) framework, only information that is retrieved successfully is strengthened, as non-retrieved content could not be remembered. In contrast to rereading, which strengthens all content to the same degree, retrieving correct content is crucial for the benefits of retrieval-based learning.

In contrast to these robust findings, research that explores the effects of retrieval-based learning as a function of individual differences (e.g. need for cognition) is rare (Dunlosky et al., 2013). Furthermore, much less is known about the promises of retrieval-based learning to promote learning in collaborative instruction. The extent to which retrieval-based learning has a positive effect for learners with different learning abilities, for complex learning materials in authentic educational settings, still remains an open question (Erichsen, & Mayer, 2015).

3. Research questions

To investigate the effects of retrieval-based learning and rereading, in the context of inquiry-based learning for complex learning material, we examined three questions:

- How does the consolidation of content knowledge during inquiry influence students' achievement and learning process (search for hypotheses during inquiry)?
- Which learning technique (retrieval-practice versus rereading) is more beneficial for the acquisition and retention of knowledge?
- Are the learning techniques (retrieval-practice versus rereading) for different kinds of learners concerning the tendency to engage in and enjoy cognitive endeavours (Need for Cognition; NFC) more beneficial?

4. Method

4.1 Sample and research design

A total of 120 6th and 7th graders from five different secondary schools in Germany were recruited. Two children were not given the permission by their parents and 25 were not able to participate during one of two learning occasions of the study. This resulted in a sample of 93 participants with a mean age of 12 years ($SD = .64$), and 58% were boys. The topic was new to all students, and the teachers confirmed that the students had not visited the Experimental-Lab Biology FLOX before.

The study followed a 3 (learning condition: retrieval-practice, rereading and no consolidation (control)) x 2 (retention interval: 10 minutes and 1 week) mixed factorial design. In addition, the study included 3 learning conditions with random assignment of the students: a condition which included retrieval-based learning during inquiry ($n = 30$), a condition which had reading guidance to consolidate scientific concepts ($n = 30$), and a control condition in which no consolidation strategy was implemented ($n = 33$). To ensure that the groups were comparable, the cognitive abilities (operationalized through the grades in Biology and German; due to the German grading system the grades range from 6 (unsatisfactory) to 1 (excellent)) and age were compared. The results from a one-way ANOVA showed that there were no significant differences in the cognitive abilities, $F_{Biology}(2,87) = .52$, $p = .597$, $F_{German}(2,89) = .35$, $p = .706$, or age, $F(2,90) = .80$, $p = .452$ between the groups. These results suggest successful random group assignment.

All students were tested 10 minutes after the inquiry task (T1) and one week later (T2) with the same content knowledge questionnaire. There were several reasons behind the choice of the time intervals (T1 and T2): From a theoretical standpoint, storage of information in long-term memory can already occur after 10 minutes; a central time marker is the next day, as sleep is essential for knowledge consolidation (Tetzlaff et al., 2012). These findings coincide with evidence from cognitive psychology showing the benefits of retrieval-based learning after short delays (Roediger & Karpicke, 2006). Furthermore, the delay of one week (T2) is due to the fact that educators often want to build on previous knowledge acquired during the last lesson, which is usually one week prior.

4.2 Learning environment

Students in all three conditions engaged in an inquiry task that covered the biological concept of animal adaptation using a specific organism, water fleas (*Daphnia magna*), as an example. The concrete task invited pupils to solve "The Mystery of Water Fleas' Migration" in a guided inquiry task in an extra-curricular learning environment, the Experimental Biology Lab FLOX. The phenomenon of water fleas' daily migration and the research question "How do water fleas react to light?" formed the starting point for an open-ended investigation. Students were confronted with the mystery that water fleas descend at dawn to the depths of the body of water, even though their principal food is located in the upper layers. At dusk, they return to the upper levels. Scientific findings generally explain this behaviour through avoidance of ultraviolet radiation as well as avoidance of visually orientated predators. At night, water fleas migrate to the illuminated water layers in order to find food. The relations between various aspects of the phenomenon as well as hypotheses stated by the learners while exploring the hypothesis space are outlined in Figure 1.
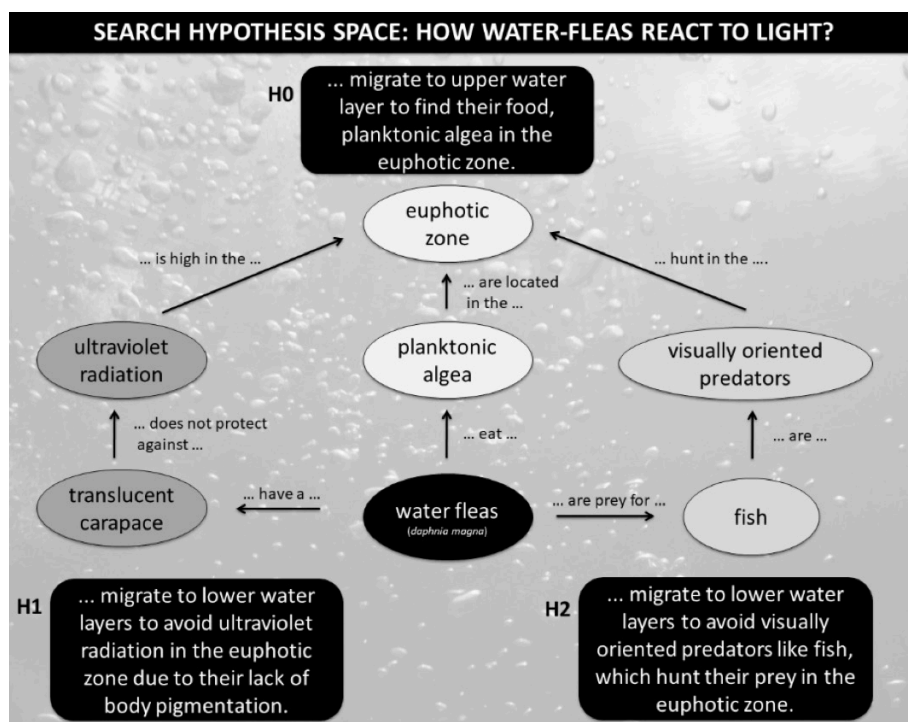

Figure 1. Overview of scientific concept that is necessary for searching for hypotheses.

After being presented with the phenomenon, students engaged in a scientific inquiry process by following the idealized inquiry process. All students completed the inquiry process in the same way, the only difference was in the learning task students had to solve before generating hypotheses or evaluating data. In the retrieval-based learning condition students were provided with a task in which they had to solve questions in order to call learned information actively to mind. The content which they had to remember was based on the information outlined in Figure 1. In the rereading condition, students were given set text to read to help them remember content (for an example see Figure 2). Students in the condition without a consolidation activity did not recall any information, but received a different task, so that the total amount of learning time remained the same. During the consolidation activities, as well as throughout the inquiry process, supervisors were instructed not to answer any questions concerning content knowledge.

4.3 Procedure and instruments

The intervention consisted of three consecutive learning sessions, each spaced one week apart. The first learning session was a lesson at school which was standardized for all students by using a self-constructed computer-based science program. The lesson was computer-based to minimize teaching style-dependent effects, and to verify that all students received the same information it was constructed by the researchers. After a short introduction by trained supervisors, the students worked at their own pace. It provided students with basic content knowledge. Short videos and reading material covered all relevant content knowledge (e.g. anatomy of water fleas, aquatic ecosystems, animal adaptations) for the inquiry unit, without anticipating the content to be learned. The information covered in the computer program was developed using information obtained from Meier & Wulff (2012). To ensure that the materials were read carefully and learning time was utilized effectively, the program contained control mechanisms (e.g. controlling the reading time by specifying the minimum duration of a trial). All students were given 30 minutes to work on the program. Afterwards, a questionnaire about the need for cognition (Preckel, 2014) was administered, and socio-demographic information was gathered.

One week later, the second learning session took place in an outreach lab, the Experimental-Lab Biology FLOX at the University of Kassel. Here,

students had three hours to solve "The Mystery of Water Fleas' Migration" in a guided inquiry task. At the beginning of the inquiry task, supervisors, which had a passive role during inquiry, introduced learners to the phenomenon and research question. Then, students had three hours to complete the inquiry task in small groups of up to five. Students also completed either the retrieval-based learning task, the rereading task, or the control condition task, depending on which condition they were assigned during inquiry at the point at which domain specific knowledge was required (before searching for hypotheses and evaluating data). The tasks comprising short texts for the rereading condition, questions (multiple-choice format) for the retrieval-based learning condition (see Figure 2), and a distractor task for the condition without consolidation, were computer-based and introduced by supervisors.
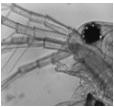
Rereading condition	Retrieval-based learning condition
When viewed under a microscope, the compound eyes of water fleas are one of their most prominent features. With them, water fleas aren't able to see shapes, but they can distinguish between light and dark. This targeted response to light shows that the crustaceans are well adapted to their environment. The targeted reaction is important in order to find their food.	With their compound eyes, water fleas are able to ... ☐ perceive colours, e. g. to find other members of their species. ☐ perceive shapes, e.g. to flee from predators. ☑ distinguish between light and dark, e.g. to find their food. ☐ distinguish distances for orientation in the pond.

Figure 2 Sample passage and question used in the rereading and the retrieval-based learning conditions (translated from German by the authors).

Therefore, students interrupted group work during inquiry and spend five minutes in total working at their own pace on the different computer-based activities. After completing the activity, students continued with the inquiry task and were asked to generate a possible hypothesis and evaluate their data on their own in order to get authentic information about students' abilities. At the end of the inquiry task a paper-and-pencil questionnaire to evaluate students' content knowledge was applied (T1). It consisted of seventeen multiple choice items relating to the scientific knowledge that is outlined in Figure 1. To go into detail, for each connection that is displayed in the figure an item was developed in order to reflect the knowledge sufficiently. All items had four possible answer options, only one of which was correct. Accordingly, the maximum knowledge score was seventeen. Item difficulty, internal consisten-

cy and discrimination parameters for the questionnaire were calculated using classical test theory. The item difficulty was appropriate (*p* .58) and the test was found to be reliable ($\alpha = .79$). Further, the discrimination parameters were all above r_{it} .30.

The last learning session took place one week later at school. Students were given the same paper-and-pencil test that was used at the first test occasion (T1). After students completed the test, the intervention ended with a discussion of students' results during the inquiry process and a short science lesson about animal and water flea adaptation.

4.4 Data analysis

Statistical analyses were performed using mixed-model ANOVA and one-way ANOVA to identify significant differences between groups and learning abilities. All results were significant at the .05 level unless otherwise stated. Pairwise comparisons were Bonferroni-corrected to the .05 level. To estimate effect sizes, partial eta squared (η_p^2) was used. To calculate performance during the inquiry process, students' answers during inquiry were analysed. The assigned scores were entered into SPSS for frequencies analysis.

5. Findings

5.1 Student achievement

The mean proportions of correct responses at each retention interval are shown in Figure 3. Test performance was entered into a 3 (learning condition) x 2 (retention interval) mixed-model ANOVA to test whether learning conditions differed significantly over time. The ANOVA revealed that the retention of content knowledge differed depending on learning condition, $F(2,90) = 17.83$, $p < .0001$, $\eta_p^2 = .11$. As evident from Figure 3, both consolidation strategies outperformed the condition without consolidation. Furthermore, there was a main effect of retention interval, $F(1,90) = 11.76$, $p < .001$, $\eta_p^2 = .11$, with retention decreasing over time.
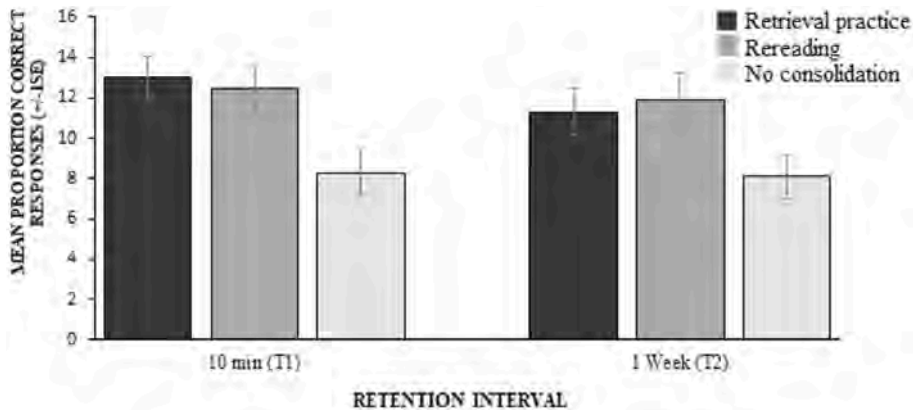
Figure 3 Mean proportion of correct responses for the three learning conditions at each retention interval.

The interaction between learning condition and retention interval was also significant, $F(2,90) = 3.96$, $p = .022$, $\eta_p^2 = .08$. This effect was caused by the decrease in retention in the retrieval-based learning condition (13% retrieval-based learning versus 5% rereading). In order to further examine whether conditions with retrieval-based learning outperform other consolidation strategies, two separate one-way ANOVAs were conducted. The conditions' performance differed reliably at the immediate test after 10 minutes (T1), $F(2,90) = 21.81$, $p < .0001$, $\eta_p^2 = .33$, as well as at the delayed test after one week (T2), $F(2,90) = 11.47$, $p < .0001$, $\eta_p^2 = .20$. Post hoc comparisons showed that the retrieval-based learning condition significantly outperformed the condition without a consolidation strategy (T1: $p < .0001$; T2: $p = .001$). In contrast, however, the retrieval-based learning condition and the rereading condition did not differ significantly at either retention interval.

5.2 Analysis of the learning process

To examine how the learning conditions influenced the learning process, the students' answers were analysed. Two independent raters, who were knowledgeable about the content, scored the formulated hypotheses and the evaluation of the data. Only the results for generating hypotheses are presented. Accuracy of the justification of the hypotheses were scored by awarding three point for fully correct, two or one point for partially correct justifications and zero points for scientifically incorrect justifications (derived from links in Figure 1). Inter-rater correlation was calculated and ranged from $\kappa = .81$-$.86$

depending on the justification for the different hypotheses. From frequency distribution analyses it emerged that students in the retrieval practice (26%) or rereading condition (40%) were more likely to formulate more specific hypotheses in contrast to students in the control condition (10%). Separate analysis for the possible hypotheses (see Figure 1) revealed that students in the rereading (56%) and retrieval-practice (33%) conditions most often mentioned H1. Hypotheses that considered the preference of light H3 were mentioned less frequently (27% and 17%). When it comes to the justification of hypothesis, however, students showed shortcomings. In fact, fully correct justifications are mentioned only by 3% of the retrieval-practice condition and 4% of the rereading condition. Partially correct results were stated by 9% and 14% of the retrieval practice condition and by 14% and 22% of the rereading condition. In contrast, the performance of the control condition remained at an unsatisfactory level as students were not able to formulate fully correct justifications whereas 6% and 4% were able to formulate partially correct justifications. Overall, this finding indicates that embedding information during inquiry seems to be necessity to promote searching for hypotheses.

5.3 Explaining students' learning patterns as a function of learner characteristics

To evaluate the hypothesis that learning outcomes relate to differences in learning abilities concerning the tendency to engage in and enjoy cognitive endeavours, the students were divided into groups with high Need for Cognition (NFC) ($>$ 3.32) and low NFC ($<$ 3.32) using a median split of individual NFC scores. Test performances were entered into two separate one-way ANOVAs, one for each test occasion. The ANOVA for the first test occasion after 1 week revealed a significant main effect for NFC, $F(1,59) = 5.84$, $p = .019$, $\eta_p^2 = .10$.

To further analyse this finding, students falling in the upper and lower quartile (called 'high performers' and 'low performers' respectively) of the distribution of NFC scores were selected. Figure 4 shows that high performing students benefited more from the rereading condition, whereas low performing students from retrieval-practice, $F(1,27) = 4.26$, $p = .049$, $\eta_p^2 = .14$. As can be seen from Figure 4 the learning outcome in the retrieval practice condition is strikingly similar (recalling about 70%), whereas in the reading condition large differences were observed (83% versus 55%). Low performers, however, operated at a similar level like students in the control condition without support.
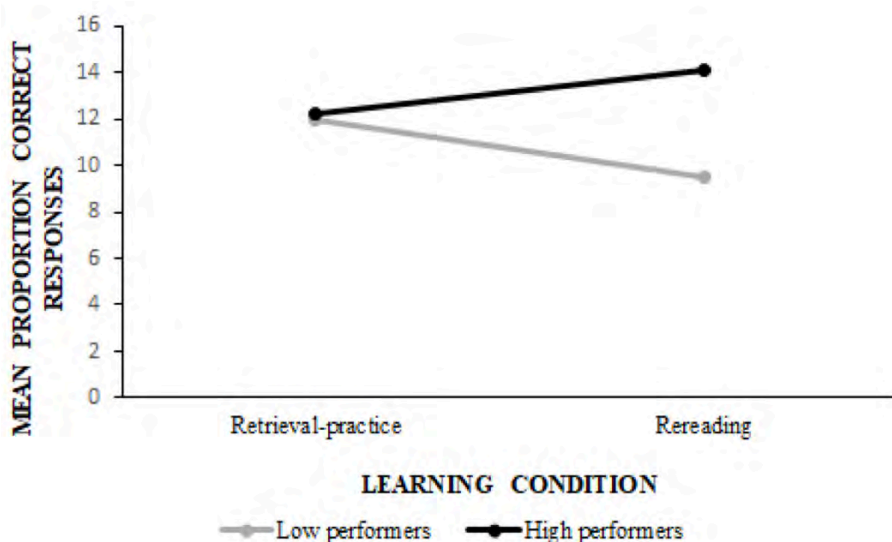
Figure 4 Mean proportion of correct responses for high and low performers in the retrieval-practice and rereading condition after one week.

6. Discussion

In order to address the deficiencies and demands of current discourse on inquiry-based learning we analysed the benefit of retrieval-based learning compared to rereading or no consolidation strategy, for the long-term learning outcome of scientific content knowledge during inquiry-based learning.

Consistent with previous findings the present study showed that engaging students in an inquiry process is not a guarantee to learn scientific content on a long-term basis. Even though students were provided with domain knowledge before the inquiry task, acquisition and retention of scientific concepts remained unsatisfactory. This points to the need for guiding the inquiry process. We focused on scaffolds that help to access previously studied information in order to use it for inquiry and to help retention of content. Accordingly, the present study showed that accessing knowledge from long-term memory could be a powerful mnemonic enhancer. From previous research, we expected that retrieval-based learning would be more beneficial for retention than rereading the learning material (Roediger & Karpicke, 2006). However, contrary to our prediction, learning outcomes in the retrieval-based learning condition and the rereading condition did not differ significantly. This result

might be explained in two ways:

Firstly, the complexity of the material could have had an influence on learning outcomes in the retrieval-based learning condition (van Gog, & Sweller, 2015). As scientific content can be seen as rather complex, the positive effect of retrieval-based learning might have been reduced by the complexity of the content, and learning outcomes could have been primarily due to students' ability to recall complex information successfully (Kornell, et al., 2011). In fact, misconceptions produced during retrieval might have remained or been repeated by students at the test one week later (T2). Here, the rereading condition could have had an advantage, because only correct information was strengthened.

Secondly, further use of the information during the inquiry process (to generate hypotheses) could influence acquisition and retention of content knowledge. The present study showed the advantage of rereading as the students reactivated only correct information. In comparison with students in the retrieval-practice condition, learners were able to generate more specific hypotheses by more likely engaging in focused processing of information.

This finding is important when considering inquiry-based learning as appropriate for learning scientific content because it reveals the need for embedding domain information during inquiry in order to build up sustainable knowledge. The importance of a scientific introduction is well stated in the literature (e.g. De Jong, 2006), but the importance of access to this knowledge during inquiry is not. This is somehow surprising because in the course of knowledge construction students have to rely on their prior knowledge to process new information in working memory, in order to actively integrate them into existing memory traces, which is of particular importance during the hypothesis space. Nevertheless, to overcome the shortcomings during generating hypotheses, students need further support.

Furthermore, the study showed that NFC is a personality characteristic that effects learning. A high NFC had a beneficial influence on learning and was a predictor for higher learning outcomes. We found that retrieval practice is more beneficial for low performers, who are less motivated to engage in demanding learning processes, than for high performers. In contrast, rereading had a greater effect for high performers. Interestingly, low performers in the rereading condition showed a similar achievement to students in the control

condition. This finding is important when considering optimal guidance during inquiry-based learning. In fact, it points to the need for embedding a different kind of support for individuals with specific learner characteristics. As inquiry-based learning is stated to be an ineffective method for achieving suitable learning outcomes for low achieving students (Klahr & Nigam, 2004), this finding seems to be important and points to important implications for optimizing inquiry-based learning. Retrieval-based learning represents a learning technique that could overcome the stated problems by providing those learners with domain specific information during the learning process.

Overall, the practical implication, although tentative, points to the importance of engaging students generally in the retrieval process during instruction instead of being added on at the end of learning processes. It is therefore prudent for educators to not only focus on adequate guidance to help the encoding of information during inquiry, but also on sufficient methods to help consolidation. Nevertheless, the results would be of limited value if students could not retain the learned scientific knowledge for a longer period of time or apply it. Therefore, the next step will be to investigate the potential benefit of retrieval-based learning over a longer retention interval (up to six weeks), and the application (transfer) of the acquired knowledge.

Acknowledgments

References

Arnold, J. (2015) *Die Wirksamkeit von Lernunterstützungen beim Forschenden Lernen: eine Interventionsstudie zur Förderung des Wissenschaftlichen Denkens in der gymnasialen Oberstufe*. Logos Verlag: Berlin.

Bjork, E.L. & Bjork R.A. (2011) Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*. New York: Worth Publishers.

Bredderman, T. (1983) Effects of activity-based elementary science on student outcomes: A quantitative synthesis. *Review of Educational Research*, 53(4), 499–518.

Carpenter, S.K. (2009) Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 35(6), 1563-1569.

De Jong, T. (2006) Scaffolds for scientific discovery learning. In J. Elen & R. E. Clark (Eds.), Handling complexity in learning environments: *Theory and research* (pp. 107-128). London: Elsevier.

Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J, & Willingham, D.T. (2013) Improving students'

learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, 14(1), 4-58.

Erichsen, A. & Mayer, J. (2015) Feedbacktests beim Forschenden Lernen. In D. Krüger, P. Schmiermann, A. Möller, A. Dittmer & L. Kotzebue (Eds.), *Erkenntnisweg Biologiedidaktik 14* (pp. 125-139). Online available [http://www.bcp.fu-berlin.de/biologie/arbeitsgruppen/didaktik/Erkenntnisweg/2015/Eriksen.pdf].

Hof, S. (2011) *Wissenschaftsmethodischer Kompetenzerwerb durch forschendes Lernen: Entwicklung und Evaluation einer Interventionsstudie*. Kassel Univ. Press: Kassel.

Karpicke, J.D. (2012) Retrival-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science*, 21(3), 157-163.

Karpicke, J.D. & Aue, W.R. (2015) The testing effect is alive and well with complex materials. *Educational Psychology Review*, 27(2), 317-326.

Kirschner, P.A., Sweller, J. & Clark, R.E. (2006) Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75-86.

Klahr, D. (2000) *Exploring Science: The cognition and development of discovery processes*. Cambridge: MIT Press.

Klahr, D. & Nigam, M. (2004) The equivalence of learning paths in early science instruction: effects of direct instruction and discovery learning. *Psychological Science*, 15(10), 661-667.

Kornell, N., Bjork, R.A. & Garcia, M.A. (2011) Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language*, 65(2), 85-97.

Lazonder, A.W. & Harmsen, R. (2016) Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, 86(3), 1-38.

McDaniel, M.A., Agarwal, P.K., Huelser, B.J., McDermott, K.B. & Roediger, H.L. (2011) Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology*, 103(2), 399–414.

Meier, M. & Wulff, C. (2012) *Klein, aber oho: Bau und Lebensweise des Wasserflohs. Unterrichtsreihe für das Sammelwerk RAAbits Biologie Realschule, 6*. Heidelberg: RAABE Verlag.

Preckel, F. (2014) Assessing Need for Cognition in Early Adolescence: Validation of a German Adaptation of the Cacioppo/Petty Scale. *European Journal of Psychological Assessment*, 30(1), 65-72.

Roediger, H.L. & Karpicke, J.D. (2006) The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, 1(3), 181-210.

Schroeder, C.M., Scott, T.P., Tolson, H., Huang, T.-Y. & Lee, Y.-H. (2007) A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching*, 44(10), 1436-1460.

Shymansky, J.A., Hedges, L.V. & Woodworth, G. (1990) A reassessment of the effects of inquiry-based science curricula of the 60's on student performance. *Journal of Research in Science Teaching*, 27(2), 127–144.

Tetzlaff, C., Kolodziejski, C., Markelic, I. & Wörgötter, F. (2012) Time scales of memory, learning, and plasticity. *Biological Cybernetics*, 106, 715-726.

Van Gog, T. & Sweller, J. (2015) Not new, but nearly forgotten: the testing-effect decreases or even disappears as the complexity of learning materials increases. *Educational Psychology Review*, 27(2), 247-264.

Zhang, L. (2016) Is inquiry-based science teaching worth the effort? Some thoughts worth considering. *Science and Education*, 25(7), 897-915.

"Why do parrots talk?" co-investigation as a model for promoting family learning through conversation in a natural history gallery

Emily Harris[1] and Mark Winterbottom[1]

mw244@cam.ac.uk

Abstract

Research into how and what families learn in science museums and other informal science learning settings suggests that parent-child interactions play an important role in shaping children's learning experiences. Our exploratory case study set out to discover and analyse learning happening within family groups during a visit to a traditional museum natural history gallery. Research methods were influenced by a growing body of literature that looks for learning in family visitor talk. Conversations of 18 families were recorded as they explored a gallery after being introduced to six learning games which fostered a 'climate of inquiry' and which were designed to spark family dialogue. Our findings indicate that families adopt a range of interactional approaches for building meaning together in a museum gallery. These approaches fell along a spectrum that varied according to the level of co-investigation and co-operation between group members. We suggest that family learning could be supported in informal learning contexts through simple, low-cost learning strategies that encourage dialogue and co-investigatory behaviours.

Keywords: family learning, learning talk, natural history museum, objectbased learning

---

[1] Faculty of Education, University of Cambridge, Cambridge, UK

The full version of this paper is published in the Journal of Biological Education:

Harris, E. & Winterbottom, M. (2018) "Why do parrots talk?" co-investigation as a model for promoting family learning through conversation in a natural history gallery, *Journal of Biological Education*, 52(1), 89-100.

Web-based training in school to increase learning success during dissection in biology classes

Katharina Luther[1], Jonathan Wrede[1] and Marc Gerhard[1]

gerhard@bio.uni-frankfurt.de

Abstract

Dissecting organs at school offers special learning opportunities. Due to ethical considerations, dissections in classrooms should be used for maximum educational benefit. Therefore we analyzed the effect of a web-based training (WBT) which prepared students for the dissection of a pig heart. A part of the WBT is a short video demonstrating the steps of the dissection and reflecting the exact situation in the classroom (the same dissection tray, preparation utensils and object to be dissected). In this study 8 classes were prepared for performing a dissection with the help of a WBT. A control group comprised 8 other classes with students not being prepared for the dissection day. Due to the German school system, half of the group belonged to a gymnasium and the other half to an integrated comprehensive school, both of them representing two different types of schools. For the study 401 students were given questionnaires at three points in time (pre-test, post-test and delayed post-test), and 77 out of 122 teams were monitored through teacher observation. The study showed that WBT accompanying an actual dissection increased students' learning success, helping the students conduct a more competent dissection, since they made fewer wrong cuts and they worked faster. By presenting the real situation with the video as a part of the WBT students knew what to expect. This decreased feelings of disgust and thus failures. The delayed post-test indicated that content knowledge afterwards was better if students were prepared with the help of the WBT. Based on our results of the investigation, the use of WBT at schools makes sense. Slower learners and visual learners in particular benefited from the clearer structure and the use of realistic video

---

[1] Department of Bioscience Education, Goethe University, Frankfurt, Germany

clips for preparation. The intensive preparation at home resulted in better conceptual understanding and better memorization of terms after the dissection day.

Keywords: e-learning, flipped classroom, learning success, peer feedback, pig heart dissection, web-based training

1. Introduction

Biology teachers try to use real objects during lessons since they presume that it might be conducive to learning (Petto & Russell, 2003). The advantages of using real animals in learning have long been recognized (Mayer and Hinton, 1990; Keiser & Hamm, 1991; Offner, 1993; Hepner, 1994). However, there has been considerable discussion of the pros and cons of animal dissections (Orlans, 1988). Ethical concerns often make students reluctant to participate in the dissection of animals, and some students refuse (Balcombe, 1997). Virtual dissections have been discussed and used as an alternative many times. Several studies have indicated that factual knowledge about respiration, cardiovascular anatomy and physiology can be improved when substituting the dissection (Leonard, 1992; Erickson & Clegg; 1993, Lilienfield & Broering, 1994; Samsel et al., 1994; Lunsford & Herzog, 1997; Lombardi et al., 2014). Nevertheless, current German teaching methodology in Biology demands the study of real objects as a fundamental form of studying in Biology lessons, such as the preparation of plants and animals (Köhler, 2012; Gropengießer et al., 2013; Otteni, 2014; Graf, 2016; Killermann et al, 2016). One ethical implication of the debate is that if dissection is used in schools, it should be used for maximum educational benefit. However, unprepared students do not achieve the same learning success as primed students. In several studies, e-learning resources prior to laboratory sessions reduced cognitive load and improved students' self-efficacy and understanding of practical activities (Jones & Edwards, 2010; Peteroy-Kelly, 2010; Al-Khalili & Coppoc, 2014; Whittle & Bickerdike, 2015). The current study investigates whether the preparation for a dissection improves students' results in problem-based learning. Due to the large differences between the real object and two-dimensional abstractions, such as worksheets used to prepare for a pig heart dissection (Spörhase-Eichmann, 2005), a film sequence was used for the preparation. The delivery of videoclips through

e-learning opportunities is much easier than giving an USB flash drive to each student or sending an e-mail with a huge attachment.

WBT was therefore designed to optimize the preparation for the dissection day. We hoped for positive effects similar to the ones in Steed's study (1999:31) "Studies have shown that interactive versions of training programs increase the learner's understanding of the course material by as much as 56% over the classroom version."

The film sequence used in our WBT shows only the beginning of a dissection. We ensured that the video showed the optimum procedure for the dissection without revealing any results. To this end the arterial openings were indicated as the ideal site to place a cut with the scissors and the cutting technique with the knife in the direction of the ventricle walls shown. However, the video did not show the opened heart with a view of the inner ventricles or the structure of the heart valves. Therefore problem-based learning is still possible, which is more nurturing and enjoyable compared with conventional instructional approaches (Albanese & Mitchel, 1993). In our problem-based learning the students needed to find out the anatomy and function of the heart valves. The colour and moving display of the object in the video of the WBT should have decreased the distance of the students from the task. To achieve maximal recognition effect from the video, the preparation instruments, preparation trays and utensils were the same as for the dissection in school. This should have reduced the degree of abstraction, since the students knew where, when and with what each cut was to be made (Hommel & Stränger 1994). The overall research question in this study was:

Do students who perform a dissection after having a preparatory WBT show i) better dissection attitudes before, during and after the actual dissection, ii) better dissection skills during the dissection, and iii) better content knowledge afterwards, compared to students who perform the dissection without the prior WBT?

We pursued the following six working hypotheses:
1. Students show a more positive attitude towards the real object, since they are familiar with it and have been shown an open-minded handling of the object.

2. Students feel more confident and better prepared if they have already seen the beginning of the dissection in the video.
3. The quality of the essential dissection skills increases, since individual steps of the dissection are explained in detail in the video.
4. The students complete the dissection in less time, since they already know where to place the cuts and which instruments to use.
5. The students have a more positive attitude towards the dissection based on a retrospective view, since they began the dissection day with a more positive attitude.
6. The participation in WBT has a positive effect on students' learning success, since they spend more time investigating the essential structures of the organ.

2. Methods

2.1 Subjects

The schools in our investigation were all within a town with 50.000 inhabitants, which lies to the north of the metropolitan area of Frankfurt am Main, Germany. Students of both schools came from middle-class families. Both schools are co-educational. The subject group in our investigation included 48% female students. Due to the peculiarities of the German school system and to assess the effect of differences of the students' "level of performance", the two schools chosen for our study present two types of schools with different school-leaving qualifications. The 'Gymnasium' (secondary school with higher performing courses) leads to an A-Level, whereas the neighbouring 'IGS' (integrated comprehensive school with lower performing courses) offers an intermediate school-leaving qualification (GCE Ordinary level). In total 16 classes took part in this study, with 8 classes at each school.

The students showed different degrees of familiarity with the subject matter: for 7 classes the structure and function of the heart was covered several weeks before our investigation, for 5 classes the dissection coincided with their regular curriculum, and 4 classes had not yet covered this topic in their curriculum. The students' familiarity with the teacher varied: each dissection was taught by one of three teachers from the regular teaching staff of the gymnasium. These teachers alternated teaching the dissection class. The teachers

taught two of the form 5 and form 9 classes in our study as part of their regular teaching schedule. Four heart dissection classes therefore were taught to students by their regular teacher and 12 classes were taught by an unfamiliar teacher. Most dissections began at 8am, four started at 10am, two at 12noon and one at 2pm. Including the pre-test and the introduction to the dissection, teachers had 90 minutes available for each lesson, although the actual dissection required much less time.

Two grade levels were identified to assess the effect of differences of the students' "development status". Based on curriculum requirements, a heart dissection could be integrated usefully in form 5 (10-year-old students) and form 9 (14- year-old students). The fundamentals of the cardiovascular system as well as respiration and metabolism are addressed in the curriculum of form 5 students. Four years later form 9 students study blood groups and the immune system. This provides an opportunity for a renewed look at the cardiovascular system. The 401 students from 16 classes carried out the dissections in teams of mainly 3-4 persons. Of all 122 teams, 77 teams (37 teams in classes without WBT, 42 teams in classes with WBT) were monitored by teacher observers.

2.2 The self-learning course: WBT

As can be expected in Germany (Behrens & Rathgeb, 2015), all students had access to a computer at home with internet access. This was the basis for a preparatory homework task and to develop a self-learning course, which was available as WBT. The WBT was not meant to replace conventional teaching, but to support and prepare before the actual lesson in class (compare with Petko, 2014).

Self-determined learning is advantageous, since the students themselves determine when and how long they study. If they want, they can repeat elements of the course and monitor their progress themselves (Kubler LaBoskey, 2004). However, students often find it difficult to study the material themselves, since the motivation of being part of a social group is absent. By studying alone, there is a danger that the course is not completed (Kerres, 2001; 2013). These benefits and problems were weighed with the 'Autorensystem Lernbar' (Krömker, 2016), a relatively short WBT package with 15 pages, which can be completed in 20 minutes. This prevents students aged 10 to 14 from ending the course prematurely because of temporal overload. Furthermore, students

were asked to print the last page, which provided an overview to the previously completed pages, and bring it to the dissection day. This allowed teachers to check whether the preparatory homework was completed.

The core element of the self-learning course is a video (Gerhard & Wrede, 2016; DiLullo et al., 2006) which shows students the dissection cutting technique. Students can learn a lot by observing in particular: the site of applying the cuts on arteries, the distance of the cuts from the groove between the ventricles, known as the anterior interventricular sulcus, the appropriate use of knife or scissors, and the application of slicing rather than pressing cuts. The video also contains the colour and size of the heart, the size in relation to a person's hands, as well as how to handle a pig heart. A voiceover on the video describes the individual steps of the dissection, as well as introducing and explaining essential technical terms and dissection skills in detail.

Following the video, 10 diagrams were presented covering the content of the video. All questions could be answered based on the video alone. Since well-defined lower skills were required to answer all questions, answers were presented as multiple choice. There were two to six possible answers. Students had three tries to give a correct answer before viewing the next diagram: they could check their answer as the correct answer was shown, and they could use the reset button to retry the question. After each answer, feedback was shown to inform the students further. Some questions included advice on which part of the video answers the question. After viewing the video again, students could tackle the question anew. The questions overall were to increase the attention of students on the content of the video and to retain the information.

2.3 Implementing the pig heart dissection

The human heart discussed in the curriculum is similar to a pig heart in structure and size, hence the latter were chosen for this dissection. It took approximately two months for a butcher to gather the necessary amount of hearts. Approximately 140 hearts were stored in a freezer in the school until the dissection day. The hearts were thawed once before to remove the enclosing membrane, the pericardium and any adipose or fatty tissue. The veins and arteries were shortened. Pulmonary and body arteries were separated from each other by cutting the connective tissue, so they are easier to recognise. Thawed hearts, like any other meat, develop an odour over time. It is therefore recommended to place the hearts thawed the evening before the dissec-

tion in lukewarm water for a half hour and to speed up the thawing period by flushing the heart with water shortly before the dissection. Flushing the heart is also important to remove blood clots from the ventricles. Such clots are not recognised as such by students and can increase their revulsion towards a dissection.

According to the concept of problem-based learning, before initiating the dissection the question was raised with the students, "Why does each side of the heart only pump blood in one direction?" (Raschke, 2009). Often students suspect an answer to this question analogous to a technical solution, such as a bicycle valve. Such statements were kept in mind as a working hypothesis. The dissection therefore begins with an investigation into the structures which prevent the blood from flowing backwards. For this reason, the self-learning courses only covered the dissection technique and did not discuss or show the opened heart. The experimental and control groups only differed in the preparation by the WBT. Apart from that, the same applied to both groups. Every team used preparation trays and preparation instruments, including knife, scissors, glass rod and tweezers, as well as other equipment such as disposable gloves and paper towels. Both groups, experimental and control, received a dissection guide in the form of two information pages. These were meant to help the students in imagining the real structures of the heart and therefore organise the information in their minds (Schaal, 2012). This guide was designed to match the video. It contained an overview diagram with labels of the heart and a step-by-step guide with sketches of each cut.

Since the students in the control group did not take part in the WBT, a detailed description was necessary for them to accomplish the dissection properly. This is particularly important for this dissection, since completed cuts cannot be corrected, especially when interesting structures are already cut. Since the time was short, a detailed guide was necessary to complete the dissection in time (Otteni, 2010). The experimental group received the same guide and all conditions on the dissection day were kept the same. To reduce the number of hearts needed and promote co-operation, the students worked in teams. Most often this resulted in working groups of 3 or 4, seldom in pairs or larger groups. This also allowed students who did not want to participate in the dissection themselves to observe.

The dissection itself consisted of two main parts:

External inspection. A glass rod was used to trace the path of blood through the heart and to mark the nine visible parts of the heart using small labels: the arteries, the veins, the four chambers and the groove between the ventricles. The labelling was monitored by the teacher before the team was allowed to progress in their dissection.

Inner inspection. Two cuts into the heart were made. Each cut was placed into the appropriate artery to open the right or left ventricle. Structures that belonged to the atrioventricular valves, the chordae tendineae, were quickly discovered. The colourless, close membrane of the semilunar valves were discovered much later, often only with help see below. When the dissection was complete the hearts were collected in a separate container und the dissection equipment cleaned.

Before the end of the problem-based lesson students were reminded of the question asked at the beginning of the lesson: "Why does each side of the heart only pump blood in one direction?" The discussion considered the statements of the students which were kept in mind as a working hypothesis. Through this the prior knowledge about technical solutions, such as a bicycle pump valve, was added by biological solutions, such as atrioventricular and semilunar valves.

2.4 Research design

The present quasi experimental design compared students who received conventional teaching with those who received WBT. We used a two factorial design (Table 1) in a fully crossed design (Bates et al., 1996). We analysed the independent variables "preparation" (WBT yes / no) versus "level of performance" (school type) as well as "preparation" (WBT yes / no) versus "development status" (age of students).

Table 1 Number of classes of the two 2x2 factorial designs with the independent variables "preparation" versus "level of performance" as well as "preparation" versus "development status".

"preparation"	"level of performance" (school type)		"development status" (age of students)	
	Gymnasium	IGS	form 5	form 9
WBT yes	4	4	4	4
WBT no	4	4	4	4

In total, 401 students performed a dissection in 122 teams of mainly 3 or 4 students. One half of the 16 classes, the experimental group, received WBT and the other 8 classes formed the control group without WBT. Always two classes each of following combinations: Gymnasium/form 5, Gymnasium/form 9, IGS/form 5 and IGS/form 9, together 8 classes, formed one group. Initially we assessed potentially confounding variables. Most important for the performance skills and attitude are familiarity with the teacher, time point in which the dissection topic was addressed within the school year and time of day. The participating classes were distributed into experimental and control groups so that the confounding variables were present to the same degree. This was used to control the confounding variables by group matching.

To check whether the use of the WBT changed the learning success we used different methods:

Questionnaires. The students of the experimental group and the control group were given a questionnaire to determine a pre and post dissection change. This allowed us to investigate non-observable behaviour, such as opinions and attitudes, by self-evaluation. The completion of the questionnaire was done anonymously and at the same time by the whole class. We used a pre-test-post-test-control (PCC) design: shortly before their dissection (t1), one week after the dissection (t2), and half a year after the dissection (t3). During the 6 months none of the classes performed a further dissection in Biology lessons.

To focus on the dependent variable "better dissection attitudes" the pre-test (t1) contained five items the students used to estimate their attitude towards a heart dissection (Figure 1a) and a further 5 items which enquired about the level of preparation for the dissection (Figure 1b). The post-test after a week (t2) referred to the same variable. It contained five items where students needed to self-assess their attitudes due to the dissection based on a retrospective

view (Figure 1c). The "change of dissection attitudes" towards the real object (pig heart) is evident in 3 items, which were enquired about at all three points of time t1-t3 (Table 2). These were all closed questions using a four-level Likert scale, which made indecision impossible (Creswell, 2012), since there was no middle value allowing a "neither agree nor disagree" or alternative allowing a "don't know" option. For statistical analysis the chi-square test (^{2}) was used. In the figures the percentage of agreeing to the statement, i.e. levels 3 (rather agree) and 4 (fully agree) of the Likert scale are displayed as bar graphs with their vertical axis on the left side. The mean and standard deviation are displayed with their vertical axis on the right side.

The questionnaires at all three points of time t1-t3 gathered data about the dependent variable "knowledge concerning dissection skills". One item about the correct cut is depicted in Table 3. Therefore students were given three illustrations to choose from: a correct cutting procedure (from arteries along the septum), a faulty cut (from the veins along the outer sides of the heart), or a false cut (perpendicular to the septum). In the faulty case the position of the cut was wrong but the direction of the cut was correct. The possibility to choose the faulty cut revealed a student's partial understanding.

The dependent variable "acquired content knowledge" was elevated only in the delayed post-test (t3), with an open question to express students' opinions freely (Seliger & Shohamy, 2013).

Their ideas about the aim of the problem-based learning, "Why does blood only flow in one direction?" were categorized into concepts (Table 4). Five concepts differing in complexity were identified following the qualitative methodology of thematic analysis (Braun and Clarke, 2006). Furthermore, we analyzed whether technical terms were used in answers or not.

Observations: In order to identify if students after the preparation by WBT showed better dissection attitudes, additional teachers as observers monitored the dissections of 77 teams. Before the dissection day the observers were briefed and used a standardized observation sheet during the dissection. Beside some open answers there were closed questions using a four-level Likert scale, exactly like in the questionnaires. The results of five items about the general impression of attitudes during the dissection are shown in Figure 2a. Five items of the observation of the quality of the essential dissection skills

are shown in Figure 2b. Additionally, the working speed during the dissection was monitored by the observers and sometimes documented by photographs. Altogether we were able to specify the working speed of the external inspection for 92 out of 122 teams and the internal inspection for 87 teams. In the statistical analysis the null hypothesis was checked with the Student's t-test.

3. Results

3.1 Self-assessment by students

Clear and significant differences were found between the groups in their self-assessed level of preparation for the dissection (Figure 1a). Students having prepared with the WBT indicated more frequently that they knew the steps of the dissection. They also claimed to know which instruments to use and which components of the heart they would see. The students in the WBT group also indicated they knew in which direction to place the cuts. Overall the students felt more confident and better prepared, since they had already seen the beginning of the dissection in the video. Partly the participation in the WBT changed the attitude and feeling of students towards a real pig heart. The answers of the students showed ambivalent attitudes toward the dissection. On the one hand students felt more confident because of their increased knowledge due to "preparation" by WBT. They also had a clearer and more accurate idea of what a heart looks like after completing the video preparation (Figure 1b). However, on the other hand the students could not imagine any more what a dissection feels like than students from the control group. Both groups were relaxed shortly before the dissection (Figure 1b).

Surprisingly, after the dissection day the students did not have more positive attitudes towards a dissection due to the dissection based on a retrospective view, although they felt more confident because of their increased knowledge at the beginning of the dissection day. On the contrary, following the dissection, all group differences in the students' self-assessment disappeared (Figure 1c). Both groups showed the same degree of fascination towards the dissection. Most students agreed with the statement that they learned a lot during the dissection, that they would remember the structure of the heart, and that they would remember each step of the dissection. Both groups of students showed an interest in dissecting other organs in class.

The "change of dissection attitudes" towards the real object (a pig heart), was not very obvious either. The willingness to touch the pig heart was similar before (t1) and a week after (t2) the dissection among the experimental and the control group (Table 2a). After half a year (t3) students prepared by a WBT were more willing to touch a heart. The willingness strongly decreased in all control groups. It increased among students with WBT in form 9 and in the IGS. The WBT affected the groups of both school types in a different way concerning the willingness to cut the heart themselves (Table 2b). It was the willingness of the IGS students, which rose, as opposed to the willingness of students of the Gymnasium, which decreased first (t1 und t2). Shortly before the dissection (t1) the students who knew what to expect stated more often that cutting a heart was disgusting for them, as opposed to students who were not prepared.

A week after the dissection (t2) the number of students who felt disgusted increased in both groups. However, the increase of disgust of the control group was clearly stronger than the one of the experimental group.
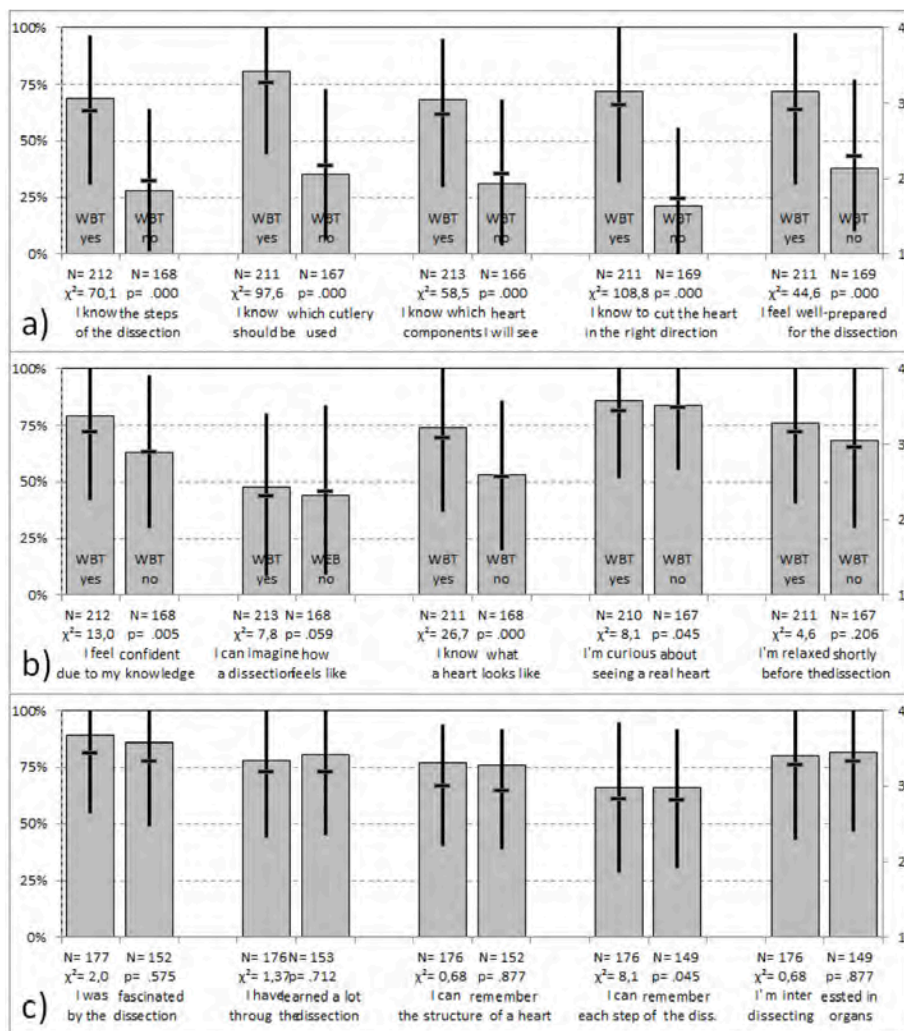
Figure 1 Attitudes of students as self-assessed in questionnaires in the experimental (WBT yes) and control (WBT no) groups, with median and standard deviation of Likert scales (right vertical axis) and degree of agreeing to the statement as a bar graph (left vertical axis). Items concerning: a) the level of preparation for the dissection-day (pre-test t1), b) a heart dissection before dissection day (pre-test t1), c) the dissection based on a retrospective view (post-test t2).

After a long period of time (t3) the disgust towards the pig heart seemed to be remembered clearly more strongly (Table 2 c). Students of the IGS accounted for the largest share of this result. The groups in this school type, both the experimental and the control groups, show a stronger difference at all times (t1-t3) than the groups in the Gymnasium. The WBT seemed to

stabilize the way form 5 students felt disgust towards the cutting of a pig heart (74% t2, t3), whereas the share of unprepared students in form 5 increased strongly due to the dissection (63% t1 to 88% t2) and afterwards decreased again back to the start value (63% t3).

Table 2 Change of attitude towards the pig heart, degree of agreeing to the statement (I agree: levels 3 and 4 of Likert scales) to the statements of students in the experimental (WBT yes) and control (WBT no) groups; date: pre-test before dissection (t1), post-test one week after the dissection day (t2) and half a year after the dissection day (t3).

WBT		Total			IGS			Gymnasium			form 5			form 9		
											10 years			14 years		
	date	t1	t2	t3	t1	t2	t3	t1	t2	t3	t1	t2	t3	t1	t2	t3
Yes	N=	211	177	145	105	83	73	106	94	72	84	62	58	127	115	87
No	N=	168	153	135	68	42	44	100	111	91	65	58	63	103	95	72
	a) Item in the questionnaires: I want to touch the pig heart.															
Yes	I agree %	83	86	89	88	90	95	78	82	83	81	79	81	84	90	94
No	I agree %	81	87	75	78	86	70	83	87	77	80	86	75	82	87	75
	b) Item in the questionnaires: I want to cut the pig heart myself.															
Yes	I agree %	82	87	79	81	90	79	83	84	78	86	82	75	80	89	82
No	I agree %	84	87	74	74	80	80	91	90	71	86	88	73	83	87	75
	c) Item in the questionnaires: Cutting the pig heart disgusts me.															
Yes	I agree %	75	81	83	78	86	92	71	78	74	70	74	74	78	85	89
No	I agree %	71	85	67	54	81	64	73	86	68	63	88	63	76	83	71

The preparation due to WBT affected the students, despite the contradiction that there was more disgust in the long run. However, it also increased the students' willingness to touch and cut the pig heart themselves. Attitudes of form 5 students varied less and those of IGS students towards the real object increased the most.

3.2 Observations in class

The observing teachers stated that the attitudes of students who were part of the experimental group were affected by the "preparation" due to WBT. As the results of the "performance level" (i.e. different school types) and the "development status" (i.e. age of students) hardly differed when compared with the overall result of all participating students, they are not depicted in

our figures.

The general impression of attitudes during the dissection was that students after the preparation due to WBT dealt with a pig heart in a more relaxed way. Following WBT, for example, students appeared less inhibited or reserved towards the pig heart (Figure 2a, left diagram). They touched and investigated the heart more frequently and sooner. This enabled them to discover the posterior venous entrances to the heart faster. The teams in the experimental WBT group in general appeared more interested, displaying visible enthusiasm and discussions focused on the biological value of the organ (Figure 2a). If students had not encountered the object in a video previously they turned away from the heart or even left the room, thereby missing further observations.

These more positive attitudes helped students to concentrate on the dissection. The teams who completed the WBT dissected more correctly (Figure 2a) and discussed the figures given in the guide more frequently (Figure 2a). The prepared students needed less help and guidance from their teacher. However, this was not the case concerning basic dissection skills. The groups hardly differed in the positioning of the heart, and both could identify the left and right ventricles correctly (Figure 2b). Nevertheless, the observers stated that the essential dissection skills were performed significantly, more autonomously, and in a better way due to the preparation after the WBT.
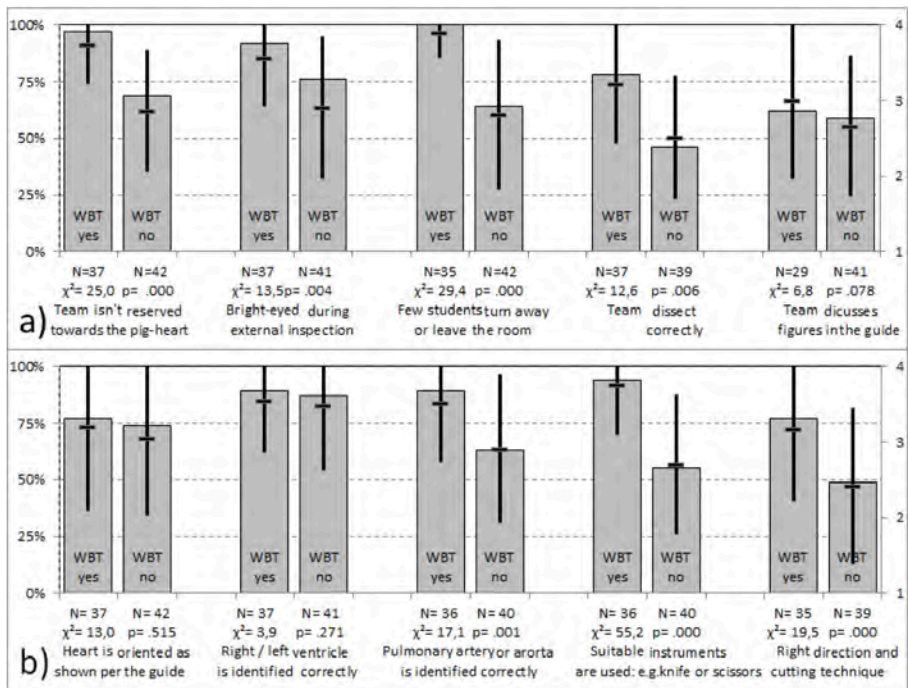
Figure 2 Dissection expertise as observed by teachers of experimental (yes WBT yes) and control (WBT no) groups, with median and standard deviation of the Likert scales (right vertical axis) and degree of agreeing to the statement displayed as a bar graph (left vertical axis). a) their general impression of students' attitudes , b) the essential dissection skills.

The WBT, however, significantly improved the speed and accuracy of students in identifying the pulmonary artery and aorta (Figure 2b). Teams from the experimental group also placed the initial cuts more appropriately: a slicing motion from the arteries into the ventricles in the direction of the cardiac septum. They also more frequently used the appropriate preparation instruments: scissors for arteries, knife for heart wall. Although both groups showed a clear distinction concerning the behaviour related to the dissection, the general school related behaviour did not differ. There were many interactions of students within all teams. Macabre interactions with the pig heart were rare and did not differ between groups. During the course of the dissection the organ engaged the students' attention more and more. For this reason, we did not observe significant differences in the number of free riders or distracting or disrupting students. We observed two students who distanced themselves clearly from the pig heart, one of which was Muslim. However, both students

were very interested in increasing their knowledge and therefore observed the dissected pig heart halves for a long time. However, we could not attribute these observations clearly to either experimental or control group.

3.3 Students' knowledge

In order to determine the positive effect on the learning process in the questionnaires there were items concerning the acquired knowledge. With regard to the students' knowledge about the cutting techniques, significantly more students in the experimental group could indicate the correct method of cutting before the dissection than in the control group (Table 3). In the pre-test the result had to be expected as only the experimental group had watched the video with the instruction for the dissection. One week after the dissection day (t2), when also the control group had learned about and applied the cutting technique, the choosing of the correct illustration increased. At this point in time, the illustration with the false cut (perpendicular to the septum) was particularly chosen less, whereas the faulty cut, which could be easily confused with the correct cut, was chosen more often.

The knowledge they acquired in the video was maintained for half a year until the delayed post-test. Obviously the WBT led to keeping the correct cutting technique in mind longer because choosing the correct illustration in the post-test (t3) decreased more strongly in the control group than in the experimental group. Table 3 indicates which students could improve their dissection skills due to the WBT most. It was the students in the IGS who improved most because the strongest differences between the experimental and the control groups were evident in the post-test (t2).

Table 3 Responses of students in experimental (WBT yes) and control (WBT no) groups regarding the correct cutting procedure (from arteries along the septum) rather than the faulty cut (from the veins along the outer sides of the heart) or false cut (perpendicular to the septum). Time points: before the dissection day (t1), one week after the dissection day (t2) and half a year after the dissection day (t3).

WBT		Total			IGS			Gymnasium			form 5 10 years			form 9 14 years		
	Date	t1	t2	t3	t1	t2	t3	t1	t2	t3	t1	t2	t3	t1	t2	t3
Yes	N=	207	177	144	104	83	71	103	94	73	83	62	59	124	115	85
	correct %	94	87	90	90	81	87	97	93	92	96	92	92	92	84	88
	faulty %	5	8	7	8	11	7	2	6	7	2	6	6	6	10	7
	false %	1	5	3	2	8	6	1	1	1	1	2	2	2	6	5
No	N=	160	150	137	63	40	45	97	110	92	61	58	63	99	92	74
	correct %	56	75	66	59	55	56	55	82	65	49	74	71	61	75	61
	faulty %	20	19	27	16	40	33	23	11	22	30	16	24	14	21	30
	false %	20	7	7	25	5	11	23	7	5	21	10	3	25	4	9

The aim of the problem-based learning during the dissection day was that students find an answer to the question, "Why does blood flow in only one direction when the heart pumps." In their answers after half a year (delayed post-test t3) students in the experimental group used more expert scientific terms (52%) than the control group (35%), such as heart valves, semilunar valves and atrioventricular valves.

Table 4 Students' ideas concerning the aim of the problem-based learning in experimental (WBT yes) and control (WBT no) groups half a year after the dissection day, answers to open questions were categorized into five concepts (a-e), ranked according to their complexity.

	WBT	total	IGS	Gymnasium	form 5 10y	form 9 14y
Concepts explaining why does blood only flow in one direction.	yes N=	163	81	82	73	90
	no N=	125	43	82	54	71
a) Students did not answer as "Because it is so". Often they expressed uncertainty or that the preparation day was too long ago.	yes %	38	46	30	42	35
	no %	56	57	55	80	37
b) Students stated an incomplete concept, such as "venous valves" / "the heart pumps"	yes %	9	6	12	10	7
	no %	8	9	7	9	7
c) Students stated a basic concept: "through / because of", often including "heart valves".	yes %	17	18	16	18	16
	no %	10	7	11	5	12
d) A commonly known concept is known, such as "something opens and closes" or "something stops the blood from flowing back".	yes %	20	15	25	19	21
	no %	16	21	14	1	28
e) Students stated a complex concept: "atrioventricular valve releases blood from atrium into ventricle", "aortic valve fills with blood".	yes %	17	17	17	10	23
	no %	9	2	13	5	13

Students in the experimental WBT-group showed more accurate and more complex understanding than the control group (Table 4, percentages of concepts b-e). Both groups, the experimental and the control group had partly inappropriate ideas concerning this question half a year after the dissection (Table 4a). However, there was 18% of inappropriate ideas across all students, and up to 38% in form 5 students due to the preparation through WBT. Therefore, Table 3 indicates which students could improve their content knowledge with WBT the most. It is the students of form 5 who improved most, with noticeably more complex ideas were expressed in the experimental group.

4. Discussion

As the results show, the student's expectations towards the dissection day did not differ significantly, regardless of whether they completed a WBT program or not. The dissection itself, specifically interacting with the real object, fascinated both groups equally, leading students to estimate their learning success similarly. This corresponds to Kinzie et al. (1993), a study in which all

students also reported a significant gain in dissection self-efficacy, and no between-group differences were found. Approximately two thirds of the students in our experimental group indicated that they were less repulsed by the heart due to the expectations formed following the video. Their objectively estimated level of disgust was reduced (cf. Randler et al., 2012), and they were more likely to want to touch the heart, both effects lasting long after the dissection itself and increasing over time. We therefore cannot fully accept the working hypothesis concerning students' more positive attitude towards the real object.

There is widespread agreement that appropriate pre-laboratory preparation is beneficial to students as it facilitates their learning and understanding (Gregory & Di Trapani, 2012; Jones & Edwards, 2010; Chittleborough et al., 2007). However, the preparation alters a number of factors significantly: the participants in the WBT felt more prepared and more often thought they knew all the necessary steps. Nearly 90% of the experimental group thought this was in connection with viewing the video clip. Our observations indicated a faster orientation and better completed dissection because of it. However, the video clip merely decreased the time students needed to view the exterior of the heart. The students mainly used the remaining piece of the aortic arch, which turns to the left-hand side, or the curved groove between the ventricles along the anterior side, to find the correct orientation of the heart. Students who did not complete the WBT needed significantly more help in this. This applies for the experiences of laboratory demonstrators who spent more time helping students resolve complex issues rather than trivialities after their students were prepared by blended learning (Gregory & Di Trapani, 2012). The teachers benefited from preparation due to WBT as well. In a typical classroom with up to 30 students, in which a single instructor is hard-pressed to lend assistance to all student teams conducting a dissection, well-prepared students could be less exhausting (Kinzie et al. 1993)

Teams in our study that showed a high degree of distance from the real object only turned the heart over at a late time point, which resulted in the side facing upwards by chance to be viewed for a long time. This occurred most for teams without WBT preparation. The time needed for the inner inspection, to cut and open the heart, was similar for both groups. Teams in the experimental group needed less help from the teacher, completed the cuts faster, had to correct fewer cuts and prepared the inner view of the heart better. This left

them more time to determine the direction of blood flow and examined vital structures, such as heart valves, for longer. Gregory & Di Trapani (2012) also stated that students using online pre-laboratory preparation were able to plan and manage experimental time more effectively. This confirms three of our working hypothesis: students felt better prepared, showed a better quality of dissection skills and dissected faster when they completed a WBT beforehand. Encounters with real objects at school often tends to be very time-consuming, especially as the dissection of living things or organs is costly in terms of preparation and teaching time (Lombardi et al., 2014). If dissection is faster, with preparation due to WBT, it could prevent some of these problems dealt with in the classroom.

In both groups the majority of students were interested in doing more dissection. This investigation found no differences in attitude towards the dissection before or after the intervention. These results lead us not to accept the fifth hypothesis. However, we regard it as an indication of the fact that dissection in schools makes learning more motivating and diverse (Elizondo Omaña et al., 2005). There are indications that the use of an interactive simulation as a preparation for dissection affects the learning of anatomy positively (Kinzie et al. 1993; Akpan & Andre, 2000). Crucial for the use of a WBT in school, however, is the clearly positive effect on learning success (our 6th hypothesis): the participation in our WBT led to better recall of the task, more scientific concepts and more frequent use of technical terms up to half a year after the WBT and dissection day.

With a view towards a successful outcome, executing the dissection correctly is very important. Cuts cannot be adjusted easily for very delicate objects and the most interesting structures may be destroyed before the investigation. In these cases a preparatory WBT is very useful. The WBT and self-learning was meant to give a better overview of the field, and thereby differ from a time-consuming and often laborious internet-based search for materials, which would require analysis and understanding (Döring, 2000). Our impression is that form 5 students, who had little experience with dissecting real objects, benefitted most from the clear guidelines and preparation at home (see Murphy (2004) Leadership for Literacy), particularly since self-evaluation was an integral part of WBT, which ensured a sufficient understanding of the video. Brennan (2003) found that in order to help ensure effective

student learning outcomes, online pedagogy needs to address a variety of factors. Among these are the approaches which are used to enable learners to build new skills based upon the ones they have already acquired, and whether a consistent level of appropriate feedback exists.

We would have liked to communicate the haptic feedback in the video, as this might have reduced the degree of irritation when encountering the organ. However, this is not possible with current technology. The use of new media still had a positive and supportive effect, since form 9 students use these in their spare time and could use them appropriately to increase their learning success (Sofos & Kron, 2010). With regards to the constructivist view of learning, this investigation succeeded in encouraging students to construct their own knowledge (Gerstenmaier & Mandl, 1995) through the intervention of a preparatory web-based training programme. The training allowed and fostered independent thought, which stimulated knowledge that was retained for a long time.

Acknowledgement

References

Akpan, J.P., & Andre, T. (2000) Using a Computer Simulation Before Dissection to Help Students Learn Anatomy. *Journal of Computers in Mathematics and Science Teaching*, 19(3), 297-313.

Albanese, M. A., & Mitchell, S. (1993) Problem-based learning: a review of literature on its outcomes and implementation issues. *Academic Medicine*, 68(1), 52-81.

Al-Khalili, SM, & Coppoc, GL (2014) 2D and 3D stereoscopic videos used as pre-anatomy lab tools improve students' examination performance in a veterinary gross anatomy course. *Journal of Veterinary Medicine Education* 41, 68–76.

Balcombe, J. (1997) Student/Teacher Conflict Regarding Animal Dissection. *The American Biology Teacher*, 53(1), 22-25.

Bates, R. A., Buck, R. J., Riccomagno, E., & Wynn H. P. (1996) Experimental design and observation

for large systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1), 77-94.

Behrens, P., & Rathgeb, T. (2015) JIM Jugend, Information, (Multi-)Media Basisstudie zum Medienumgang 12- bis 19-Jähriger in Deutschland. Medienpädagogischer Forschungsverbund Südwest, Stuttgart.

Braun V., & Clarke V (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology,* 3:77–101.

Brennan, R. (2003) *One size doesn't fit all: Pedagogy in the online environment.* Kensington Park, Australia: National Centre for Vocation Education Research, Australian National Training Authority.

Chittleborough, G., Treagust, F., & Morecino, M. (2007) Achieving greater feedback and flexibility using online pre-laboratory exercises with non-major chemistry students. *Journal of Chemical Education*, 84(5), 884-888.

Creswell, J. W. (2012) *Educational Research: Planning, Conducting and Evaluating Quantitative and Qualitative Research.* ISBN: 978-0133549584, London/New York, Pearson.

DiLullo, C., Coughlin, P., D'Angelo, M., McGuinness, M., Bandle, J., Slotkin, E. M., & Berray, S. J. (2006) Anatomy in a new curriculum: Facilitating the learning of gross anatomy using web access streaming dissection videos. *Journal of Visual Communication in Medicine*, 29(3), 99-108.

Döring, N. (2000) Lehren und Lernen im Internet. In: Batinic, Bernard (Ed.) *Internet für Psychologen.-* ISBN: 978-3801712266, Hogrefe, Göttingen, 449-508.

Elizondo Omaña, R. E., Guzmán López, S., & De Los Angeles García-Rodríguez, M. (2005) Dissection as a teaching tool: past, present, and future. *The Anatomical Record Part B: the New Anatomist*, 285(1), 11-15.

Erickson, H.H., & Clegg, V.L. (1993) Active learning in cardiovascular physiology. In: *Promoting Active Learning in the Life Science Classroom*, edited by Modell H.I., Michael J.A. New York: Annals of the New York Academy of Sciences, vol. 701, p. 107–108.

Gerhard, M., & Wrede, J. (2016) Mehr Zeit zum Experimentieren-Ein Vorschlag zur Konzeption einer Unterrichtseinheit „Enzymatik" in der Oberstufe mittels des ICM. *Biologie Lehren und Lernen–Zeitschrift für Didaktik der Biologie*, 20(1), 82-91.

Gerstenmaier, J., & Mandl, H. (1995) Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik*, 41, 867-887.

Graf, E. (2016) Biologiedidaktik: für Studium und Unterrichtspraxis (5. bis 13. Klasse) - 3. Aufl. Auer Verlag, Donauwörth.

Gregory, S. J. & Di Trapani, G. (2012) A blended learning approach to laboratory preparation. *International Journal of Innovation in Science and Mathematics Education (formerly CAL-laborate International)*, *20*(1), 56-70.

Gropengießer, H., Harms, U., & Kattmann, U. (2013) Biologie Fachdidaktik, Aulis Verlag.

Hepner, L.A. (1994). *Animals in education: the facts, issues, and implications*. Richmond Publishers.

Hommel, B., & Stränger, J. (1994) Wahrnehmung von Bewegung und Handlung. In W. Prinz & B. Bridgeman (Hrsg.). Enzyklopädie der Psychologie, Themenbereich C Theorie und Forschung, Serie II Kognition, Band 1 Wahrnehmung (S. 529-604). Göttingen: Hogrefe.

Jones, S. M. & Edwards, A. (2010) Online pre-laboratory exercises enhance student preparedness for first year biology practical classes. *International Journal of Innovation in Science and Mathematics Education* (formerly CAL-laborate International), 18(2)

Keiser, T.D. & Hamm, B.W. (1991) Forum: dissection: the case for. *The Science Teacher*, 58(1), 13,15.

Kerres, M. (2001) Multimediale und telemediale Lernumgebungen: Konzeption und Entwicklung.

ISBN-13: 978-3486250558, DeGruyter Oldenbourg, München.

Kerres, M. (2013) Mediendidaktik: Konzeption und Entwicklung mediengestützter Lernangebote. ISBN-13: 978-3486736021, DeGruyter Oldenbourg, München.

Killermann, W., Hiering, P. & Starosta, B. (2016) Biologieunterricht heute: Eine moderne Fachdidaktik (Alle Klassenstufen).- Auer Verlag, Augsburg.

Kinzie, M. B., Strauss, R. & Foss, J. (1993) The effects of an interactive dissection simulation on the performance and achievement of high school biology students. *Journal of Research in Science Teaching*, *30*(8), 989-1000.

Köhler, K. (2012) Welche Erkenntnismethoden sind für den Biologieunterricht relevant?. IN: Spörhase, U., Biologie Didaktik – Praxishandbuch für die Sekundarstufe I und II.- 5. Auflage, Cornelson Berlin.

Krömker, D. (2016) Autorensystem Lernbar. Goethe-University, Senckenberganlage 31, 60325 Frankfurt am Main, Germany URL: http://lernbar.uni-frankfurt.de/.

Kubler LaBoskey, V. (2004) The Methodology of Self-Study and Its Theoretical Underpinnings. In Loughran, J. J., Hamilton, M. L., Kubler LaBoskey, V., & Russel, T. (Ed.), *International Handbook of Self-Study of Teaching and Teacher Education Practices* (pp. 817-869). Springer International Handbooks of Education, Volume 12, ISBN: 978-1-4020-1812-1.

Leonard, W. H. (1992) A comparison of student performance following instruction by interactive videodisc versus conventional laboratory. *Journal of Research in Science Teaching*, 29(1), 93-102.

Lilienfield, L.S. & Broering, N.C. (1994) Computers as teachers: learning from animations. *Advances in Physiology Education*, 11(1).

Lombardi, S.A., Hicks, R.E., Thompson, K. V., & Marbach-Ad, G. (2014) Are all hands-on activities equally effective? Effect of using plastic models, organ dissections, and virtual dissections on student learning and perceptions. *Advances in Physiology Education*, 38(1), 80-86.

Lunsford, B.E. & Herzog, M.J. (1997) Active learning in anatomy and physiology: student reactions and outcomes in a non-traditional A&P course. *American Biology Teacher* 59, 80–84.

Mayer, V.I. & Hinton, N.K. (1990) Animals in the classroom: Considering the options. *The Science Teacher*, 57(3), 27-30.

Modak, H.K. & Borse, L.J. (2015) Use of computer-assisted teaching with animal simulator in experimental physiology for the first year MBBS students. *Group*, 2(29), 13-4.

Murphy, J. (2004) Leadership for Literacy: A framework for policy and practice. School Effectiveness & School Improvement, 15(1), 65-96.

Offner, S. (1993) The importance of dissection in biology teaching. *The American Biology Teacher*, 55(3), 147-9.

Orlans, F. B. (1988) Debating dissection. *Science Teacher*, 55(8), 36-40.

Otteni, M. (2010) Betrachten und Interpretieren. In Spörhase, U., & Ruppert, W. (Ed.), Biologie Methodik Handbuch für die Sekundarstufe I und II. Berlin, Cornelson, 84-87.

Otteni, M. (2014) Methoden zum Entdecken, Erfinden und Erarbeiten.- IN: Spörhase, U. & Ruppert, W.: Biologie-Methodik: Handbuch für die Sekundarstufe I und II. Cornelsen.

Peteroy-Kelly M. (2010) Online pre-laboratory modules enhance introductory biology students' preparedness and performance in the laboratory. *Journal of Microbiology & Biology Education*, 11:5–13.

Petko, D. (2014) Einführung in die Mediendidaktik: Lehren und Lernen mit digitalen Medien. ISBN: 978-3407256782, Beltz, Basel.

Petto, A.J. & Russell, K.D. (2003) Humane Education: The Role of Animal-based Learning. In: Arm-

strong, S.J. & Botzler, R.G., *The Animal Ethics Reader*, Routledge, London, 296-308.

Randler, C., Wüst-Ackermann, P., Vollmer, C., & Hummel, E. (2012) The relationship between disgust, state-anxiety and motivation during a dissection task. *Learning and Individual Differences*, *22*(3), 419-424.

Raschke, N. (2009) Herzpräparation. Materialien und Ideen zum problemorientierten Biologieunterricht, Kantstr. 22, 27721 Ritterhude, URL: http:// www.biologieunterricht.info/ unterrichtsmaterialien/ herzpraeparation.html, as consulted online on 31.05.2017.

Samsel R.W., Schmidt G.A., Hall J.B., Wood L.D., Shroff S.G. & Schumacker P.T. (1994) Cardiovascular physiology teaching: computer simulations vs. animal demonstrations. *Advances in Physiology Education,* 11, 36–46.

Schaal, S. (2012) Die Arbeit mit Medien planen. In Weitzel, H. & Schaal, S. (Ed.), *Biologie unterrichten: planen, durchführen, reflektieren (pp.118-131). Scriptor Praxis: Biologie.* Berlin: Cornelsen.

Seliger, H.W. & Shohamy, E. (2013) S*econd Language Research Methods. Oxford Applied Linguistics, ASIN: B00FKW2BIM*, Oxford University Press.

Sofos, A. & Kron, F. W. (2010) Erfolgreicher unterrichten mit Medien. ISBN: 978-3936172126, Logophon, Mainz.

Spörhase-Eichmann, U. (2005) Ein Schweineherz untersuchen. Unterricht Biologie, 302, 25–28.

Steed, C. (1999) *Web-based Training.* Gower, Aldershot.

Whittle S.R. & Bickerdike S.R. (2015) Online preparation resources help first year students to benefit from practical classes. *Journal of Biological Education*, 49, 139–149.

BioDiv2Go: does the location–based geogame 'FindeVielfalt Simulation' increase the valuing of local biodiversity among adolescent players?

Sonja Schaal[1], Steffen Schaal[1] and Armin Lude[1]

schaal02@ph-ludwigsburg.de

Abstract

The valuing of biodiversity is precedent to its conservation. Therefore, the aim of this study is to use mobile technology to combine game-based and location-based learning and to provide sensuous experiences discovering local biodiversity. The theoretical framework consists of biodiversity-related knowledge (BRK), situational content-related interest (SCI) as indicators for the valuing of biodiversity, game-related enjoyment (ENJ) and attitudes towards nature (INS). In the present study, a so-called geogame to foster local biodiversity was developed, and the valuing of local biodiversity as well as the role of game-related enjoyment as satisfaction of intrinsic needs were assessed. The first step of the study was to examine the consistency of the enjoyment construct using confirmatory factor analysis (CFA). Biodiversity-related knowledge, situational content-related interest and attitudes towards nature were examined via t-test and analysis of variance. The results demonstrate a significant increase of biodiversity-related knowledge in general and a stronger connection to nature from pre to post. The enjoyment data reveals some conflicting results compared to the underlying theory, and situational content-related interest increases for those who succeeded within the geogame. Implications for further research and for the implementation of geogames in environmental education are discussed.

Keywords: digital game-based learning, enjoyment, geogames, valuing of local biodiversity

[1] Ludwigsburg University of Education, Ludwigsburg, Germany

1. Introduction

The major aim of the BioDiv2Go project is to foster the valuing and the appreciation of local biodiversity using mobile devices. In general, adolescents´ interest in plants is very low and only 22% of Germans know the term *biodiversity* (BMU & BfN, 2010), and its protection is perceived to be of low importance (Menzel & Bögeholz, 2009). Discussion about its value emphasises the attitudes, value systems and interest of people as well as their willingness to protect it (cf. literature overview in Schaal, et al., 2015). There is broad agreement in the scientific community that nature experiences can increase knowledge, influence attitudes and enhance the perception of biodiversity (cf. literature review in Raith & Lude, 2014). But can nature, and especially local biodiversity, also be experienced while playing a location-based game on mobile devices? *Pokémon go* activated thousands of people to go outside catching virtual species. It also inspired national parks and environmental organisations to harness the game for real-life experiences in nature (e.g. www.modernhiker.com/2016/07/14/pokemon-go-while-hiking-outdoor-teaching-tool).

In #PokeBlitz players get help in identifying real animals while catching pokémons. So obviously there are people who believe in nature experiences moderated by technical devices. But from a more widespread perspective, digital games in general elicit different and also negative reactions by educators, teachers and parents, whereas adolescents play them regularly (Feierabend et al., 2015). Not even learning-games can fully persuade teachers, parents and students in educational settings. Parents (Bourgonjon et al., 2011) and especially experienced teachers are sceptical (Schifter & Ketelhut, 2009) and often technical obstacles are their reason for rejecting digital devices for teaching and learning. However, teachers also value the potential of digital games to promote learning (Ray & Coulter, 2010). Adolescents appreciate games because they are a form of fun, enjoyment and pleasure (Prensky, 2001), and Prensky (ibid.) shows how the role of fun in learning is controversially discussed, but he emphasizes that fun is a motivator for learning (ibid: Iten & Petko, 2016).

The geogame *FindeVielfalt Simulation* (FVS) was developed in accordance to these potential opportunities: (i) nature experiences, because the game is located in hotspot areas of local biodiversity, (ii) the potential of learning, since it is a serious game, and (iii) the experience of fun and enjoyment whilst

following the guidelines and rules of game-flow theory (Sweetser & Wyeth, 2005). So such prerequisites for it to be a promising tool for environmental education have been kept in mind. However, there is little evidence to show that a certain combination of prerequisites will enhance success in increasing the valuing of local biodiversity among adolescent players.

2. Theoretical background of the BioDiv2Go framework and the geogame *'FindeVielfalt Simulation'* (FVS)

The valuing of biodiversity contains knowledge and situational interest in local biodiversity including its perception. This definition is a result of a systematic literature review, that considers the so-called ecosystem service approaches (Sukhdev et al., 2014) respecting pedagogical-educational perceptions of the valuing of biodiversity (eg., Mayer, 2006 or Navarro-Perez & Tidball, 2012). For a detailed description of the review process see Schaal et al. (2015). The BioDiv2Go framework (Figure 1) consists of three interacting conceptual areas with the valuing of local biodiversity as a core concept and target. Likewise the game-related enjoyment is essential. The BioDiv2Go framework and its operationalization are a result of an elaborated validation process (Schaal et al., 2015) that is part of a design-based research process. It is primarily derived from the Competence Model for Environmental Education (CMEE) (Roczen et al., 2014) and it was adapted for biodiversity learning.
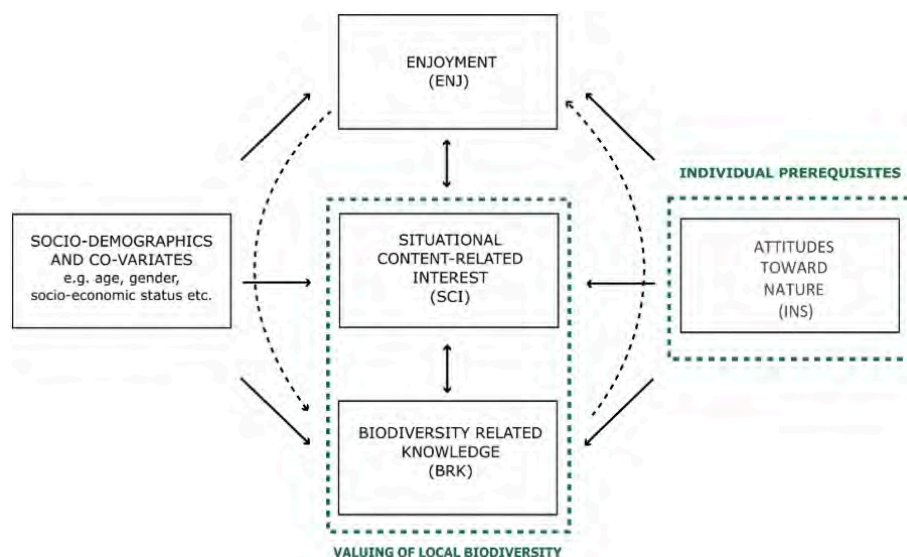
Figure 1 The BioDiv2Go framework.

Geogames are mobile, location-based and location-dependent games (Schlieder, 2014). The geographical position of the players and their locomotion are part of the game. As this type of game requires activities in the physical space, it is an exclusive possibility to combine nature experiences and gaming. In addition to conventional outdoor games like geocaching, smartphones and tablets provide audio-visual storytelling and the use of tools like integrated camera, recording and texting. A central feature of the FVS geogame is to link the real world with a virtual world (see Figures 2 and 3). The game was designed according to Schlieder′s (2014) three game-design levels. At the narrative level, FVS focuses on an adolescents′ developmental task to cope with quandaries and dilemmas between biodiversity-related and economy-related decision-making. The player takes the role of the main character in the story and has to discover a real orchard and its biodiversity, and then has to solve location-based tasks at several real places that can be located with GPS and an embedded map with flagged places (Figure 2). Within the simulation after each solved task the player manipulates and operates the orchard in the virtual world (Figure 3). With both, the game tasks and the simulation, the player gets insight into the dependencies and relationships of local biodiversity on an orchard. Helpful information to win the game and to cope with the dilemmas is provided by short texts or videos. The game is played in small groups with a maximum of three students. The goal is to balance biodiversity and economic outcomes.

Figure 2 Students playing FVS in a real orchard.

Figure 3 Simulation for orchards with different game-characters.

In environmental education, mobile technologies offer exclusive possibilities for location-based learning (Brown, 2010). The *SAMR* framework for the use of technology in learning processes (Puentedura, 2006) shows, that in reaching the levels of *Modification* and *Redefinition*, learning and teaching can be significantly transformed i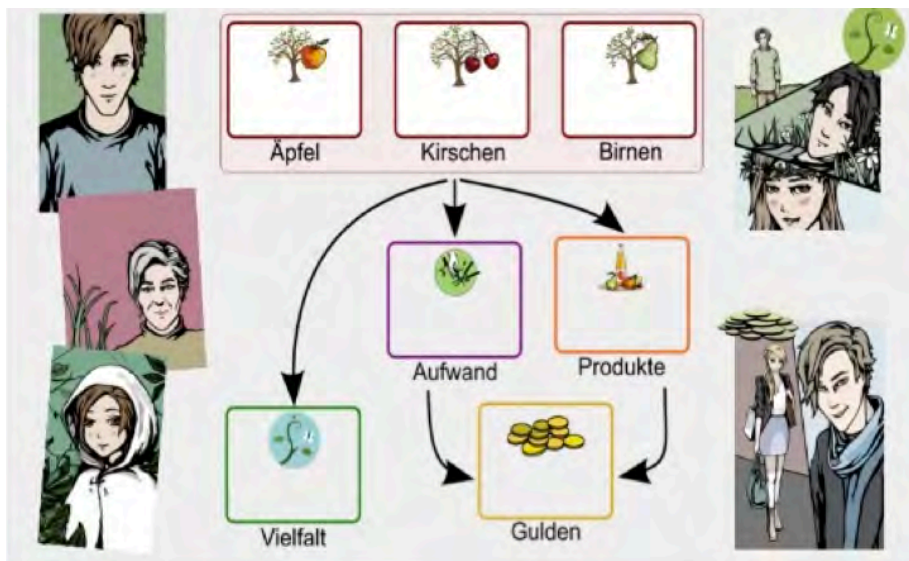n ways that would be inconceivable without technology. In the geogame for instance, the manipulation of the orchard simulation or embedded videos about animals that live in the orchard but normally cannot be observed, are prospects beyond the means of conventional educational outdoor-inquiry activities.

The FVS was designed incorporating different features and characteristics of games, as described in Garris et al.'s (2002) game dimensions and it follows the guidelines and recommendations of game-flow theory (Sweetser & Wyeth, 2005; Jegers, 2009). FVS contains *clear rules* and *goals* and *challenges* with outdoor activities combined with in-game tasks. The learners can *control* the sequence and order of their play, *skills* can improve within the game and the characters in the narrative provide help if necessary. *Feedback* is given after each task and while progressing towards the goals. The game supports *social interaction* through competition among different player groups and cooperation within a group of players.

Video and audio sequences provide information about the local biodiversity on site, which is necessary to solve the game tasks. The tasks at the real orchard site help the players to operate a virtual orchard, to raise many possibilities to discover biodiversity and to get experiences of nature.

FVS can be theoretically allocated at the interface between knowledge transfer and inquiry-based learning and playing as an active form of entertainment (Prensky, 2001; Kerres & Bormann, 2009) and thereby it can be related to a digital game-based learning framework.

The BioDiv2Go project fosters sensuous experiences and the valuing of local biodiversity using mobile technology. Numerous studies have shown, that mobile technologies used in educational settings can enhance knowledge (Perry & Klopfer, 2014; Li et al., 2013; Lai et al., 2007) and increase motivation (Tsai et al., 2012). There is also evidence that in science and environmental education, mobile technologies have a meaningful impact on the players (Perry & Klopfer, 2014; Bleck et al., 2012; Lude et al., 2013; Schaal et al., 2012; Schaal & Lude, 2015). First results using FVS underline these findings (Schaal

et al., 2015). The data analysis in the pilot study revealed an increase of bio-diversity-related knowledge. Most of the players enjoyed playing the geogame (ibid.).

A specific aim of research on educational location-based games is to identify the role of enjoyment during the playful activity and its benefit for intended outcomes. A commonly used definition of enjoyment is grounded in game-flow theory (Sweetser & Wyeth, 2005; Jegers, 2009). The game-flow model of Sweetser & Wyeth (2005) consists of eight different elements: concentration, challenge, skills, control, clear goals, feedback, social interaction and immersion. Following these recommendations, enjoyment can be predicted. But one aspect of game-flow is absolutely not compatible with location-based and location-dependent geogames like FVS: the immersion. According to Jegers (2009), players should become less aware of their surroundings to submerge completely into the game. In contrast to this, FVS intends to enhance the awareness of the environment and the local biodiversity.

Many approaches investigate only the hedonic, enjoyment part of game (e.g. Trepte & Reinecke, 2011), but according to Tamborini et al. (2010) enjoyment is not limited to a mere hedonistic pleasure response playing the FVS geogame. Connecting to self-determination theory (Deci & Ryan, 1993), the psychological basic needs fit in very well with the game design components described above: *autonomy* as a psychological need regulates one's own playing and learning process, *competence* as a need for challenge and feeling for efficacy and skill, *relatedness*, as a need for social interaction with other players. Additionally, and maybe most importantly, game-related enjoyment is not an end in itself. Its functional role in needs satisfaction could explain why gaming can add value to learning contexts.

3. Research design and methods

This study focuses on the efficacy of the geogame in relation to the particular facets of the framework. The theoretical assumptions above lead to three research questions (RQ) and hypotheses (H) for this study:

RQ1:   Can enjoyment, in its functional role as the satisfaction of the intrinsic needs for autonomy, competence and relatedness in addition to the need for pleasure seeking, be linked to the FVS geogame?

H1:	The game-related enjoyment in the FVS geogame can be explained by the intrinsic needs for autonomy, competence and relatedness and the need for pleasure seeking.
RQ2:	Does the use of the FVS geogame foster the valuing of local biodiversity (defined as a combination of biodiversity related knowledge and situational content-related interest)?
H2:	The use of the FVS geogame increases the valuing of local biodiversity.
RQ3:	Does the use of the FVS geogame foster attitudes toward nature?
H3:	The use of the FVS geogame increases the connection to nature.

3.1 Participants in the field study

Data was collected from 206 German grammar school students (T1) who played the FVS geogame during an out-of-school activity in a traditional orchard in Southwest Germany. Participants came from eleven different classes of seven different schools. The students played in small groups of 2-3 players with identical smartphones provided by the project. The game sessions took 1.5 to 2 hours. Each game session was introduced and explained in a standardised way. A control group (C) with 42 students from a secondary modern school without intervention was assessed to ensure that observed effects are not caused by the completion of the questionnaire. So they had no biodiversity teaching at all. The mean age of the players was 13.7 years old, with a standard deviation of 2.3 (C: 14.5; SD 0.8) and 52.9% of the players were female (C: 47.6%).

3.2 Research design and measures

A pre-/post-test-design was applied measuring the attitudes toward nature with the inclusion of nature in one´s self scale (INS)[pre/post], biodiversity-related knowledge (BRK)[pre/post], situational content-related interest (SII)[pre/post], enjoyment (ENJ)[post], socio-demographics[pre] and co-variates (e.g. age, sex, marks for the game and for the teamwork)[pre]. The pre-test took place 1-3 days before the intervention during a school lesson, and the post-test directly after the geogame intervention on the orchard. Additionally, log files were used to diagnose to what extent each group reached the game goals (points for biodiversity and earned money). The BioDiv2Go framework (Figure 1) shows the supposed dependencies between the variables. The game mechanics, the tasks and the

narrative were matched with the research design and the development of the BioDiv2Go framework. For more details of this educational design research process see Schaal (2017).

INS

The inclusion of nature in one's self was based on only one item (Schultz, 2002). It measures the connection to nature that represents a person's attitude toward nature. Five differentially overlapping circles labelled with "self" and "nature" can be chosen by the participant. Since this measure is a one-item measure, its reliability cannot be measured.

BRK

Biodiversity-related knowledge was measured with a 19-item multiple choice test that was piloted with an online survey and during a pilot field study in the first and second year of the project (see Schaal et al., 2015). Item examples: "If there were no bees, (a) there would be no honey, (b) there would be no grain, (c) there would be less species of fruit and vegetables", or "Biodiversity decreases if, (a) a lot of sheep or cattle graze on a meadow, (b) the meadow is fertilised, (c) the meadow is mown often". For each item 3 points can be achieved (0 = no correct answer, 3 = 3 correct answers). Observed Cronbach's alpha for the BRK scale satisfied the pre-test, $\alpha = .61$ and post-test ($\alpha = .73$). It has to be considered that knowledge scales seldom show good data especially in the pre-test, which could be explained by the bandwidth fidelity dilemma (Cronbach & Gleser, 1965).

SCI

The situational content-related interest scale is derived from a shorter scale of the Intrinsic Motivation Inventory established by Wilde et al. (2009) and the STATE scale (Randler et al., 2011). The items were adapted to the subject of biodiversity (e.g. "I want to hear more about the topic biodiversity"). Participants rated the items on a 5-point Likert scale ranging from 1 (not at all true) to 5 (totally true). Negative items were reversely coded. The scale also was tested with a confirmatory factor analysis (CFA). Observed Cronbach's alpha indicates strong internal consistency in the pre-test as well as in the post-test ($\alpha = .92$). The measures of the CFA indicate that the model fits the observed data very well (Table 1).

ENJ

Game-related enjoyment was measured with a 17-item scale including the sub-scales interest/enjoyment, perceived competence, perceived autonomy and perceived relatedness during the gameflow, following Tamborini et al. (2010) and Deci & Ryan (1993). The items were adapted to the FVS geogame ("I enjoyed playing the *Finde Vielfalt*-Game", "I think I was pretty good at playing the game", "I felt close to my teammates"). Participants also rated the items on a 5-point Likert scale. Negative items were reversely coded. Observed Cronbach's alpha for the ENJ scale is satisfying ($\alpha = .83$). The measures of the CFA indicate that the model does not fit the observed data very well (Table 1). As the enjoyment scale is a principal point in the study, the CFA results are explained in detail in the next section.

Table 1 CFA measures of global fit for the SII and ENJ scales.

	χ^2	df	p	χ^2/df	GFI	AGFI	NFI	TLI	CFI	RMSEA
Thresholds for acceptable fit			>0.05	<2	≥0.90	≥0.90	≥0.90	≥0.90	≥0.90	≤0.08
Thresholds for good fit			>0.05	<1.5	≥0.95	≥0.95	≥0.95	≥0.95	≥0.95	≤0.05
SII	8.7	7	.274	1.2	.99	.96	.99	.99	.99	.03
ENJ	204.2	111	.000	1.8	.89	.84	.86	.91	.93	.06

4. Results

4.1 Game-related enjoyment and satisfaction of intrinsic needs

H1 was tested using a confirmatory factor analyse (CFA) with AMOS 23. All 17 enjoyment items were included as respective indicators of the underlying four latent game experience constructs (interest/enjoyment, perceived competence, autonomy, and social relatedness). However, according to global-fit measures this "original CFA model" showed some poor fit indices. GFI and AGFI as well as the incremental fit measure NFI fell substantially below the critical threshold of 0.90 (Table 1). Furthermore, indices of local fit could not prove that each latent construct was reliably measured by its indicators (Figure 4). In the social relatedness or perceived autonomy sub-scales several manifest items do not reach the critical threshold of 30% for the indicator reliability (Figure 4). Finally, solely for the sub-construct interest/enjoyment the explained variance (Figure 4) is satisfying. A second model for game-relat-

ed enjoyment was tested, including only the sub-scales interest/enjoyment and perceived competence, and the marks the players gave for the whole geogame. Now the model fit is very good. Comparing the standardized regression weights and the explained variance the sub-construct interest/enjoyment represents game-related enjoyment best ($\beta= .96$; $R^2=.91$).
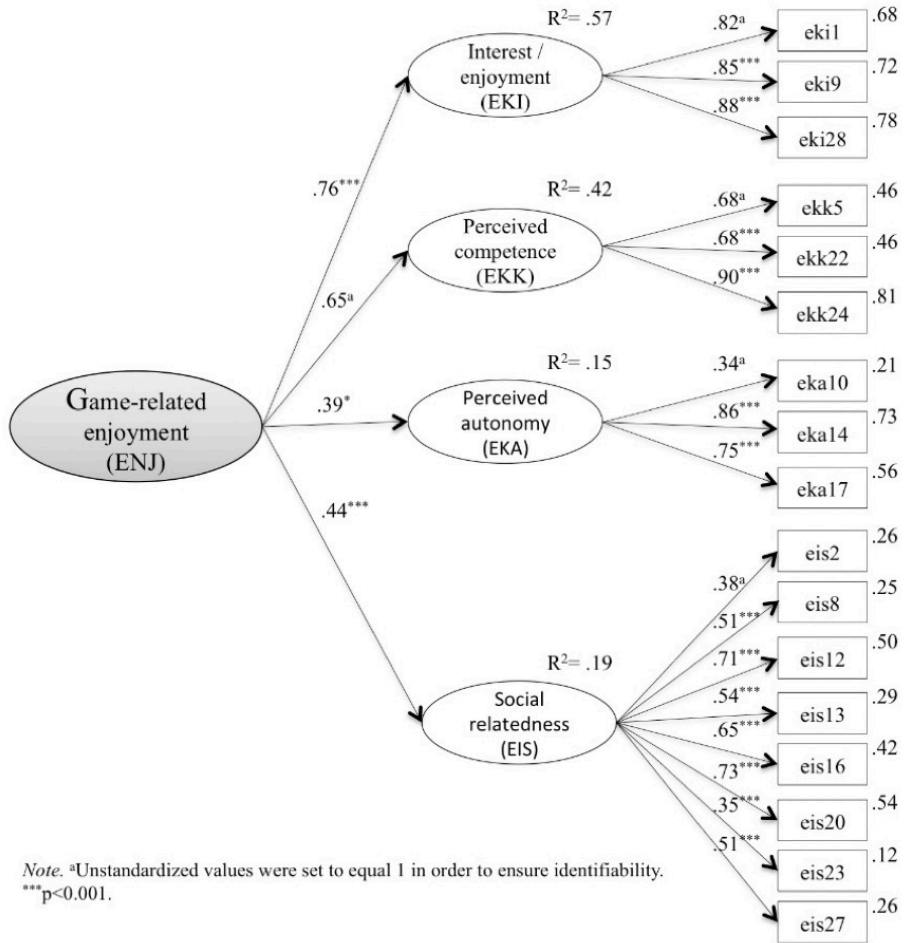

Figure 4 Measures of local fit for the enjoyment model.

These unexpected conflicting findings require a detailed discussion, but H1 is therefore not supported. The game-related enjoyment in the FVS geogame cannot be explained by the intrinsic needs for autonomy, competence

and relatedness. Only the hedonic need for pleasure seeking explains game-related enjoyment adequately.

4.2 The valuing of local biodiversity and the attitudes towards nature

The valuing of local biodiversity as the core-concept in the BioDiv2Go framework (Figure 1) is made up of situational-content-related interest and biodiversity-related knowledge. T-tests for matched pairs (Table 2) show a significant increase in the knowledge scale for the intervention group (T1) (effect size dz=.48, power: 1.0), but not for the control group (C). There is no significant increase for situational content-related interest in the intervention group (T1), but a significant decrease in the control group (C) (Table 2) (effect size dz=.51; power: 0.94). An analysis of variance (ANOVA) for the main sample T1 reveals differences between different clusters (Figure 5): groups having succeeded in both game goals, in only one of the game goals or without achieving a goal. In every cluster an increase of knowledge can be observed, even though on different levels. However, whereas the t-test does not show any significant difference in the situational content-related interest from pre to post test, the ANOVA exposes a significant increase of interest for those who succeeded in the game (Figure 5).

Table 2 Measures of the t-test.

Sample			Mean	SD	SE	t	df	Sig. (2-sided)
			T-test matched pairs paired differences					
T1	Pair 1	SUM_BRK_post - SUM_BRK_pre	2.291	4.764	.332	6.903	205	.000
	Pair 2	MEAN_SCI_post - MEAN_SCI_pre	.050	.638	.044	1.124	205	.262
	Pair 3	INS_post – INS_pre	.218	.674	.047	6.903	205	.000
C	Pair 1	SUM_BRK_post - SUM_BRK_pre	.317	4.407	.688	.461	40	.648
	Pair 2	MEAN_SCI_post - MEAN_SCI_pre	-.241	.464	.072	-3.328	40	.002
	Pair 3	INS_post – INS_pre	-.098	.539	.084	-1.160	40	.253

As implied by RQ 2 an enhancement of the valuing of local biodiversity was expected. For the knowledge scale there are distinct positive results. The players solved more multiple choice items correctly after the intervention, whereas some knowledge items picked up concrete game tasks relating to the orchard and some required an applied knowledge transfer to other habitats. Situational content-related interest was stable on average, but in relation to

the success within the geogame, over 46% of the players are more interested in local biodiversity after the intervention. Hence, the H2 can be accepted with some limitations.

The inclusion of nature in one's self scale (INS) shows a significant increase in the t-test for the main sample T1 (effect size dz=.32; power: 1.0), but not for the control group C (Table 2). The players felt 'closer to nature' after the intervention with the geogame on the orchard. Regarding the main sample, the H3 is supported by the data. The use of the FVS geogame increases the connection to nature.
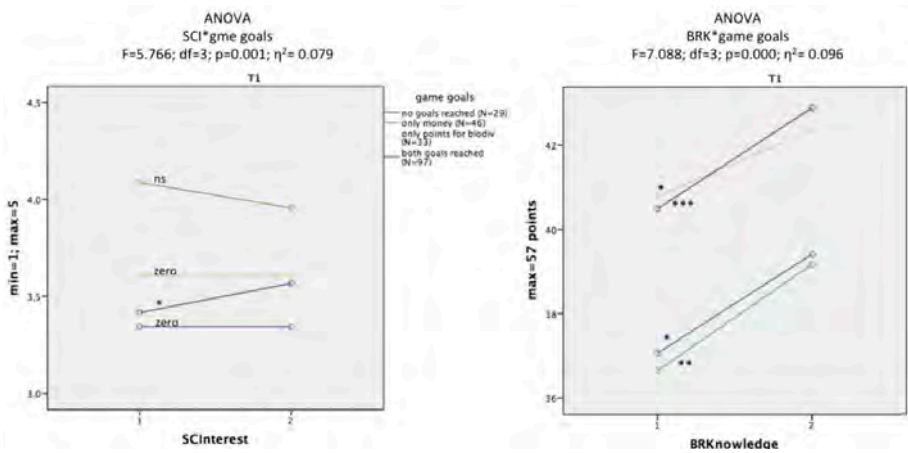
Figure 5 ANOVAs SCI and BRK in the main sample (T1).

5. Discussion

5.1 Game-related enjoyment as satisfaction of intrinsic needs

The results show that the empirical model of enjoyment rooted in the self-determination theory (Deci & Ryan, 1993) does not follow the functional role of needs satisfaction. Three psychological intrinsic needs for competence, for autonomy and for social relatedness do not explain game-related enjoyment in the FVS geogame. But it can be explained by the need for pleasure seeking. This is not in line with the findings of Tamborini et al. (2010). The needs for social relatedness and for autonomy showed a very poor model fit. This was possibly because these needs were stratified anyway in the social context of school classes, and due to the fact that the students were allowed to play in

voluntarily-selected teams. The observed cohesiveness during the game was very high and the students reported that playing in autonomous small groups was a good experience (Schaal et al., 2015). The game in Tamborini´s investigation in contrast was a classical bowling video game (Tamborini et al., 2010) tested with a computer partner or a human co-player using different types of controllers (stick movement and button press vs. a Wii controller). Hence the setting of the game is very different and not comparable to location-based games outside. It can therefore be supposed that the needs for social relatedness and autonomy can be satisfied variably in different game formats. For FVS, social relatedness and the need for autonomy seem to be important aspects (T1: mean competence = 3.7; mean social relatedness = .43; max. =5.0), but they are experienced more holistically in the setting of an out-of-school activity with classmates.

Summing up, the fit of the whole enjoyment scale and the underlying theoretical background has to be reflected for this and further studies in the BioDiv2Go project. Only the sub-scale for interest/enjoyment gives solid information about the game-related enjoyment within FVS. The FVS geogame provides a very specific kind of digital game-based learning (DGBL), this study is quite context specific and therefore the results cannot be generalized.

5.2 The valuing of local biodiversity

Since the biodiversity-related knowledge increased irrespective of the game success, the game seems to provide realistic and applied learning chances for all players. Failed answers in the game also led to more success in the knowledge-test after the intervention. Hence FVS stimulates the cognitive activation successfully. Unfortunately the research-design of the field study does not allow a follow-up test, so the sustainability of the learning process cannot be proved. The stability of the situational content-related interest in the intervention groups seems to be a good result considering that the values decreased significantly in the control group without the game. The activities on the orchard and the information about the local biodiversity capture one´s interest. But certainly more important is the result for the winning groups, whose interest even increased after the game. The situational content-related interest captured or triggered before the intervention (Mitchell, 1993), maybe influenced by the pleasant anticipation for the geogame (compare the means for SCI_pre: T1= 3.62 and C=2.93). Following Mitchell´s approach (ibid)

there might also be some successful hold-facets within FVS that stabilized or even increased the players´ interest. Making the content meaningful seems to work well within the game-design, especially when the players reached the game goals. But another interpretation of the results of the content-related interest should also be considered. The significant decrease of the interest values in the control group might also indicate that the scale is biased by the boredom the students felt filling in the test a second time. Hence the results should be interpreted carefully and further studies are necessary to investigate the validity of the interest scale.

Finally, it has to be emphasised that the connection to nature increased significantly after the intervention. That is significant for a short intervention (Bogner, 1998), because attitudes are rather stable and not easily modifiable. So this study shows that the FVS geogame manages to get adolescents in touch with local biodiversity and to connect them more closely to nature. So there is much potential for environmental education using mobile technology for game-based and location-based learning.

One aim of the BioDiv2Go project is to investigate the efficacy of the geogame itself according to the framework. A triangulation with data from another digital tool is embedded in associated studies. The more complex geogame and a simpler digital treasure hunt with an established web tool (actionbound.de) both increase the connection to nature, and show high values of game-related enjoyment (Schneider & Schaal, 2017). Crawford et al. (2016) triangulated data from digital and non-digital approaches in environmental education and showed comparable effects for the connection to nature and increased knowledge. By comparison the mobile application offered additional benefits, such as more enjoyment, compared to more traditional approaches of environmental education (ibid). So the next step within the investigation of geogames should be to estimate the impact of game-related enjoyment on the knowledge shift, and the change in the attitudes towards nature. Further game parameters like specific game results, the proportion of time the players spend on each task compared to the time spent on locomotion will be used for further analyses.

Acknowledgments

References

Bleck, S., Bullinger, M., Lude, A. & Schaal, S. (2012) Electronic mobile devices in environmental education (EE) and education for sustainable development (ESD) – Evaluation of concepts and potentials. Procedia – *Social and Behavioral Sciences,* 46, 1232-1236.

BMU & BfN (2010). *Naturbewusstsein 2009 Bevölkerungsumfrage zu Natur und biologischer Vielfalt.* Retrieved June 14, 2015, from https://www.bfn.de/fileadmin/MDB/documents/themen/ge-sellschaft/Naturbewusstsein%202009.pdf

Bogner, F. X. (1998) The influence of short-term outdoor ecology education on long-term variables of environmental perspective. *The Journal of Environmental Education*, 29(4), 17-29.

Bourgonjon, J., Valcke, M., Soetaert, R., De Wever, B., & Schellens, T. (2011) Parental acceptance of digital game-based learning. *Computers and Education*, 57(1), 1434–1444.

Brown, E. (2010) Education in the wild: contextual and location-based mobile learning in action. Nottingham: University of Nottingham.

Crawford, M.R., Holder, M.D., & O´Connor, B.P. (2016) Using mobile technology to engage children with nature. *Environment and Behaviour,* November, 3, 2016, 0013916516673870.

Cronbach, L., & Gleser, G. (1965) *Psychological tests and personnel decisions.* Urbana: Univ. of Illinois Press.

Deci, E.L. & Ryan, R.M. (1993) Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39(2), 223-238.

Deci, E.L. & Ryan, R.M. (2003) Intrinsic motivation inventory. Retrieved June 14, 2015, from http://www.selfdeterminationtheory.org/questionnaires/10-questionnaires/50

Feierabend, S.,Plankenhorn, T. & Rathgeb, T. (2015) *JIM-Studie 2015. Jugend, Information, (Multi-) Media.* Retrieved May 18, 2016, from http://www.mpfs.de/fileadmin/JIM-pdf15/JIM_2015.pdf

Garris, R., Ahlers, R., & Driskell, J. E. (2002) Games, motivation, and learning: a research and practice model. *Simulation & Gaming*, 33(4), 441–467.

Iten, N. & Petko, D. (2016) Learning with serious games: Is fun playing the game a predictor of learning success? *British Journal of Educational Technology*, 47(1), 151–163.

Jegers, K. (2009) *Pervasive GameFlow: Identifying and exploring the mechanisms of player enjoyment in perva-sive games.* (PhD thesis, Umea University Sweden), Retrieved June 14, 2015, from http://en.scien-tificcommons.org/47716935

Kerres, M. & Bormann, M. (2009) Explizites Lernen in Serious Games: Zur Einbettung von Lernauf-gaben in digitalen Spielwelten. *Zeitschrift für E-Learning*, 4(4), 23–34.

Lai, C.-H., Yang, J.-C., Chen, F.-C., Ho, C.-W. & Chan, T.-W. (2007) Affordances of mobile technol-ogies for experiential learning: the interplay of technology and pedagogical practices. *Computer Assisted Learning,* 23(4), 326–337.

Li, Z. Z., Cheng, Y. B., & Liu, C. C. (2013). A constructionism framework for designing game-like learning systems: Its effect on different learners. *British Journal of Educational Technology*, 44(2), 208–224.

Lude, A., Schaal, S., Bullinger, M. & Bleck, S. (2013) *Mobiles, ortsbezogenes Lernen in der Umweltbildung und Bildung für nachhaltige Entwicklung*. Hohengehren: Schneider Verlag.

Mayer, P. (2006) Biodiversity - The appreciation of different thought styles and values helps to clarify the term. *Restoration Ecology*, 14(1), 105–111.

Menzel, S., & Bögeholz, S. (2009) The loss of biodiversity as a challenge for sustainable development: how do pupils in Chile and Germany perceive resource dilemmas? *Research in Science Education*, 39(4), 429–447.

Mitchell, M. (1993) Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*, 85(3), 424-436.

Navarro-Perez, M. & Tidball, K. G. (2012) Challenges of biodiversity education: a review of education strategies for biodiversity education. *International Electronic Journal of Environmental Education*, 2(1), 12–30.

Perry, J., & Klopfer, E. (2014) UbiqBio: adoptions and outcomes of mobile biology games in the ecology of school. *Computers in the Schools*, 31(1/2), 43–64.

Prensky, M. (2001) *Digital game-based learning*. Mc Graw-Hill.

Puentedura, R. (2006) Transformation, technology, and education, Retrieved October 14, 2015, from http://hippasus.com/resources/tte/

Raith, A. & Lude, A. (2014) *Startkapital Natur - wie Naturerfahrung die kindliche Entwicklung fördert*. München: oekom.

Randler, C., Hummel, E., Glaser-Zikuda, M., Vollmer, C., Bogner, F. X., & Mayring, P. (2011) Reliability and validation of a short scale to measure situational emotions in science education. *International Journal of Environmental and Science Education*, 6, 359–370.

Ray, B., & Coulter, G. (2010) Perceptions of the value of digital mini-games: implications for middle school classrooms. *Journal of Digital Learning in Teacher Education*, 26(3), 92–100.

Roczen, N., Kaiser, F. G., Bogner, F. X. & Wilson, M. (2014) A competence model for environmental education. *Environment and Behavior*, 46(8), 972–992.

Schaal, S., Grübmeyer, S. & Matt, M. (2012) Outdoors and online – inquiry with mobile devices in preservice science teacher education. *Word Journal on Technology*, 4(2), 113 – 125.

Schaal, S & Lude, A. (2015) Using mobile devices in environmental education and education for sustainable development - comparing theory and practice in a nationwide survey. *Sustainability* 7(8), 10153-10170.

Schaal, So., Schaal, St. & Lude, A. (2015) Digital Geogames to foster local biodiversity. *International Journal for Transformative Research*, 3(1), 16-29

Schaal, So. (2017) *Die Wertschätzung lokaler Biodiversität mit Geogames* fördern – die Bedeutung von spielbezogenem *Enjoyment im Spiel „Finde Vielfalt Simulation"*. (PhD thesis University of Education Ludwigsburg, Germany) Retrieved June 13, 2017, from https://phbl-opus.phlb.de/frontdoor/index/index/docId/523

Schifter, C., & Ketelhut, D. (2009) *Teacher acceptance of game-based learning in K-12: The case of River City*. In Society for Information Technology & Teacher Education International Conference 2009 (pp. 3836–3842). Retrieved May 18, 2016, from http://www.editlib.org/p/31254

Schlieder, C. (2014) Geogames – Gestaltungsaufgaben und geoinformatische Lösungsansätze. *Informa-*

tik-Spektrum, 37(6), 567 - 574.

Schneider, J. & Schaal, S. (2017) Location-based Smartphone Games in the context of Environmental Education and Education for Sustainable Development - Fostering Connectedness to Nature with Geogames. *Environmental Education Research.* DOI 10.1080/13504622.2017.1383360

Schultz, P.W. (2002) Inclusion with nature: The psychology of human-nature relations. In: P Schmuck & P.W. Schultz (Eds.), *The Psychology of sustainable development* (pp. 61-78). New York: Kluwer.

Sukhdev, P., Wittmer, H. & Miller, D. (2014) The economics of ecosystems and biodiversity (TEEB): challenges and responses. In D. Helm & C. Hepburn (Eds.). *Nature in the balance: the economics of biodiversity.* Oxford: Oxford University Press.

Sweetser, P., & Wyeth, P. (2005) GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment*, 3, 1–24.

Tamborini, R., Bowman, N. D., Eden, A., Grizzard, M., & Organ, A. (2010) Defining media enjoyment as the satisfaction of intrsinsic needs. *Journal of Communication*, 60, 758–777.

Trepte, S., & Reinecke, L. (2011) The pleasures of success: game-related efficacy experiences as a mediator between player performance and game enjoyment. Cyberpsychology, *Behavior and Social Networking*, 14, 555–557.

Tsai, F.-H., Yu, K.-C., & Hsiao, H.-S. (2012) Exploring the factors influencing learning effectiveness in digital game-based learning. *Educational Technology & Society*, 15, 240–2550.

Wilde, M., Bätz, K., Kovaleva, A. & Urhahne, D. (2009) Überprüfung einer Kurzskala intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 31-45.

Butterflies & wild bees: biology teachers' PCK development through citizen science

Martin Scheuch[1,2], Tanja Panhuber[1], Silvia Winter[3], Julia Kelemen-Finan[4], Manfred Bardy-Durchhalter[1] and Suzanne Kapelari[1]

martin.scheuch@agrarumweltpaedagogik.ac.at

Abstract

Citizen science is a rapidly growing emerging field in science and it is gaining importance in education. Therefore, this study was conducted to document the pedagogical content knowledge (PCK) of biology teachers who participated in a citizen science project involving observation of wild bees and identification of butterflies. In this paper, knowledge about how these biological methods can be taught to students is presented. After two years in the project, four teachers were interviewed and their PCK was captured in the form of content representations (CoRes) and Pedagogical and Professional-Experience Repertoires (PaP-eRs). These results can help future citizen science projects to link their activities to the school curriculum. But not only success can be reported: although one of the project team's aims was to make the Nature of Science accessible to the teachers and students in the course of the project, the teachers did not take this aspect into account. This paper discusses the possible reasons and proposes various strategies for improving citizen science in the context of school biology learning.

Keywords: butterflies, citizen science, observation, PCK, species identification, wild bees

---

[1] Austrian Educational Competence Centre for Biology, University of Vienna, Austria

[2] Institute for Integrative Nature Conservation Research, University of Natural Resources and Life Sciences, Vienna, Austria

[3] Lower Austrian Academy for Nature Conservation, Austria

[4] University College for Agricultural and Environmental Education

Diagnosis of inquiry competencies using hands-on experiments with gerbils

Iris Schiffl[1] and Carina Wurdinger[1]

iris.schiffl@sbg.ac.at

Abstract

Science standards have been a focus of educational research in Austria over the last ten years. The first step was the development and evaluation of competency structure models for junior and senior classes. Subsequently, tasks for science classes were created to point out the meaning of different competencies to teachers. Later on, models of competency stages were deduced from the literature and evaluated. These models are the theoretical foundation for diagnostic instruments to measure students' competencies. So far, only diagnostic instruments limited to written answers have been used, such as the instrument for informal competency measurement in science (IKM) or the instrument for competency diagnosis in biology (IKD-Bio). As these diagnostic tools have limitations in detecting inquiry competencies, complementary methods for diagnosis are needed. This study examines whether hands-on experiments are suitable for diagnosing inquiry competencies. To this end, enquiry competencies of 8th, 10th and 12th graders were measured within a hands-on setting using gerbils. Overall, competency levels were lower if competencies were measured using hands-on experiments instead of written test designs. Problems related to handling gerbils or drawing up adequate research questions were more frequent with younger students.

Keywords: competency diagnosis, hands-on experiments, inquiry competencies

---

[1] School of Education, University of Salzburg, Austria

1. Introduction

Inquiry is the key procedure for knowledge acquisition in science. Students who use inquiry in science learning engage in the same activities and use the same cognitive strategies as scientists do when looking for answers to their research objectives (National Research Council, 2000). One method of inquiry in science classes is the use of hands-on activities (Lumpe & Oliver, 1991). A hands-on approach includes all kinds of activities that allow students to perform typical scientific work on their own (Haury & Rillero, 1994). Studies affirm the positive effects of hands-on activities on motivation (Middleton, 1995). This effect is even stronger for physical hands-on activities (such as dissection) than for virtual hands-on activities (Lombardi et al., 2014). Holstermann et al. (2010) argue in favour of hands-on experiments which are interesting for students in biology classes. Although they could not find a clear effect of experience with hands-on activities on interest, hands-on experience seems to effect science achievement positively (Stohr-Hunt, 1996). Satterthwait (2010) identified three factors which are responsible for the effectiveness of hands-on activities for student learning: peer interaction through cooperative learning, object-mediated learning and embodiment. Schwichow et al. (2016) compared the effectiveness of the hands-on approach to paper and pencil methods to find out which strategy is more fruitful in teaching the control-of-variable strategy to students. They found that students in the hands-on group outperformed students in the paper and pencil group in hands-on tests, whereas students in the paper and pencil group scored significantly higher in paper and pencil tasks.

But not all hands-on activities can be classified as inquiry-based learning. Inquiry-based learning focuses on the different steps of the inquiry cycle, starting with conceptualisation, moving on to investigation, and ending with drawing conclusions from the investigation, which in turn may lead to further discussion and questions (Pedaste et al., 2015). Hands-on activities are therefore part of performing an investigation, but without drawing up research questions, generating hypotheses, and further data processing and discussion, they do not satisfy the requirements of inquiry-based learning. The freedom to perform experiments of personal relevance in authentic contexts, for example in open inquiry laboratory sessions, is linked to the acquisition of higher-order process skills (Roth & Roychoudhurry, 1993).

As inquiry is such a crucial method in science, it takes up a central position in national science standards. In the USA, for example, the new Generation Science standards list scientific and engineering practices as one of three dimensions of the corresponding framework, and divide this dimension into eight different aspects: asking questions, developing and using models, planning and carrying out investigations, analysing and interpreting data, using mathematics and computational thinking, constructing explanations, engaging in argument from evidence, and obtaining, evaluating and communicating information (National Research Council, 2012). In Germany, inquiry is one of four competency aspects. It is described as "criterion-guided observation of biological phenomena, hypothesis-guided inquiry, criterion-guided comparison and modelling as basic methodical procedures" (KMK, 2005, p. 10). Furthermore inquiry plays a fundamental role in different standards for other countries such as Great Britain (Department for Education, 2014), Switzerland (EDK, 2011), or Singapore (Singapore Ministry of Education, 2014). As inquiry plays such a fundamental role in national science standards, and as in some of the countries listed above, students must undergo standards examinations, it is crucial for teachers to be able to diagnose their students' inquiry competencies and to use this information for planning further teaching strategies.

In Austria, the development of science standards started in 2005. These standards were subsequently transferred into competencies, including cognitive skills and abilities to solve problems in science (Weinert, 2014). Competencies are divided into three groups: competencies dealing with knowledge and its organisation, inquiry competencies and competencies for evaluating scientific findings, and engaging in actions based on these evaluations (BIFIE, 2011). So far, prototypic tasks for the different competencies have been created for use in science classes, and teachers have been trained to implement these competencies in their science teaching. In contrast to other countries like Germany, Great Britain, or the USA, no official standards testing takes place in Austria (Anand Pant et al., 2013; BGBl. II 1/2009; Schiffl, 2016). Science standards research in Austria currently focuses on the development of diagnostic instruments for teachers. Two different diagnostic tools have been developed for this purpose: the IKM (Instrument for Informational Competency Measurement) and the IKD-Bio (Instrument for Competency Diagnosis in Biology). As both tools rely on written tasks for evaluating students' competencies, they cannot cover all facets of the different competencies.

2. Objectives

As the commonly used diagnostic instruments in Austria do not include practical scientific procedures, the aim of the present study is to complement the diagnosis of inquiry competencies in biology classes by using hands-on experiments. To complete the existing data about students' inquiry competencies (Schiffl & Weiglhofer, 2016) with data from practical science procedures, the following research questions must be answered:

- To what extent do 8th grade, 10th grade, and 12th grade students demonstrate the inquiry competencies from the Austrian competency structure models when hands-on activities are used?
- How does the extent to which inquiry competencies are displayed differ between the 8th, 10th, and 12th grade?

As students are not used to doing inquiry in their biology lessons, prior studies for the IKM and the IKD-Bio generally showed lower levels of inquiry competencies than it was the case for knowledge-based tasks. Students have been found to have particular difficulties with tasks related to asking research questions and devising hypotheses (Grafendorfer & Neureiter, 2009; Schiffl & Weiglhofer, 2016). As lower-grade students have less experience in performing inquiry, their inquiry competencies are likely to be less developed than those of students in higher grades.

3. Methods

For the present study, we developed a paper and pencil diagnostic instrument. The starting point for this diagnostic instrument were the competency levels according to the inquiry cycle (Table 1; Pedaste et al, 2015). The competency levels for this study were mainly deduced from the competency levels stated and evaluated for the IKM (Schiffl & Weiglhofer, 2016), and from the empirical data on competency levels in biology presented in the IQB country comparison study (IQB, 2013; Mayer et al., 2013). For further explanations of how the individual competency levels were deduced, see Schiffl (2017).

Table 1 Competency levels for the inquiry competencies stated in the Austrian competency structure model.

Drawing up research questions

Level I   Choosing adequate research questions for completing an investigation or experiment

Level II   Drawing up adequate research questions for completing an investigation or experiment

Level III   Drawing up adequate research questions for completing an investigation or experiment, and identifying dependent and independent variables

Level IV   Drawing up adequate research questions concerning a conceptual framework in biology, and identifying dependent and independent variables

Generating hypotheses

Level I   Choosing adequate hypotheses in relation to an investigation or experiment

Level II   Generating adequate hypotheses in relation to an investigation or experiment

Level III   Generating adequate hypotheses in relation to an investigation or experiment, and providing scientific evidence for the proposed hypotheses

Level IV   Generating adequate hypotheses in relation to an investigation or experiment, and providing scientific evidence for the proposed hypotheses. Generating alternative hypotheses

Planning an investigation

Level I   Choosing an adequate experimental design

Level II   Generating an adequate experimental design

Level III   Generating an adequate experimental design with respect to confounding variables or repeat measures

Level IV   Generating an adequate experimental design with respect to confounding variables and repeat measures

Making observations

Level I   Observing without using criteria

Level II   Using given criteria for observing

Level III   Generating own criteria for observing and making observations based on these criteria

Level IV   Generating own criteria for observing, making observations based on these criteria, and explaining the results based on biological knowledge

Analysing data

Level I   Describing data from investigations or experiments

Level II   Putting data from investigations or experiments in order

Level III   Comparing different data from investigations

Level IV   Explaining data from investigations or experiments based on biological knowledge

Drawing conclusions from data

Level I   Choosing adequate conclusions for an investigation or experiment

Level II   Drawing adequate conclusions from an investigation or experiment

Level III   Drawing adequate conclusions from an investigation or experiment, and providing scientific evidence for the proposed conclusions

Level IV   Drawing conclusions from statistical data

4. Research design

The sample of this study consisted of 115 Austrian students attending the 8th, 10th, and 12th grade of schools providing general higher education. Two classes of each grade were selected for participation in the study. The sample included 39 students from grades 8 and 10, and 38 students from grade 12. The diagnostic instrument is not intended for research only, but also to be eventually available to teachers. It must therefore be applicable in normal science lessons. To ensure the applicability in science teaching, the duration of the study was limited to 50 minutes, which corresponds to the usual duration of biology lessons in Austria. As not all inquiry competencies can be surveyed in a single biology lesson, two different testing conditions were used (Table 2). Across all grades, one class was assigned to condition 1 and one class was assigned to condition 2. From 8th grade, 19 students from grade 8 completed condition 1 and 20 students completed condition 2. From 10th grade, 21 students completed condition 1 and 18 students completed condition 2. From 12th grade, 17 students completed condition 1 and 21 students completed condition 2. If a teacher was to evaluate all inquiry competencies of the Austrian competency structure model, it would take him about two biology lessons.

In both conditions, the lessons started by providing essential information about working with gerbils, including hygiene rules when working with animals and basics about handling mice. Then, some information about the inquiry process was given (Pedaste et al., 2015) and the materials available for inquiry (gerbils, observation arenas, litter, sand baths, and different toys) were presented. After the introduction, the students in condition 1 were invited to draw up a research question that could be examined by using the given materials. Where students had no idea about a question, they were given a help card from which they had to choose a suitable research question from four options. Subsequently, the research questions of all students were collected and the students were asked to generate suitable hypotheses for their questions. In the next step, the students had to plan an observation experiment. The experimentation schemes were collected and discussed in class to make sure they could be implemented in practice without any risks to students and mice. Students with similar experimentation schemes worked together in groups of four when performing the observation in practice.

In condition 2, the students had to plan an observation for answering the

following question: "How does the behaviour of a gerbil change when it is taken out of its cage (a place it knows well) and placed in an unknown territory?" The students first had to plan an observation experiment. Then, like in condition 1, the experimentation schemes were discussed in class and the students carried out the observation in groups. Subsequently, each student wrote down their observations and compared them to the observations of the other group members. Finally, each student was asked to analyse the data and write a report about the experiment.

Table 2 Research conditions.

	Condition 1	Condition 2
Basics about working with gerbils	x	x
Basics about inquiry	x	x
Drawing up research questions	x	Given research question
Generating hypotheses	x	
Planning experiments	x	x
Doing observations	x	x
Analysing data		x
Drawing conclusions from data		x
Writing reports about an experiment		x

The students were guided through the whole inquiry process by the investigator. In both conditions, worksheets containing tasks for all competencies and levels were used. The investigator had to make sure the worksheets were completed in the right order and were handed in immediately. He was also responsible for handing out help cards to students who needed them. Help cards were available for the following competencies: drawing up research questions, generating hypotheses, and planning investigations. After the investigation in class, students' written answers on the worksheets were evaluated independently by two assessors, using the competency levels from Table 1, and analysed statistically using IBM SPSS 22.0. Interrater reliability was computed using Cohen's kappa, which ranged from $k = .78$ (making observations) to $k = .95$ (drawing conclusions).

5. Results

According the inquiry cycle described by Pedaste et al. (2015), the results are divided into conceptualisation (research questions and hypotheses), investiga-

tion (planning, observing), and interpretation (interpreting data). The conceptualisation of an investigation consists of drawing up an adequate research question and devising hypotheses. Examples for answers on each competency level are provided in Table 3.

Table 3 Examples for answers on Levels I, II, III, and IV for "drawing up questions" and "generating hypotheses".

Drawing up research questions	
Level I	Help card: How does the behaviour of a gerbil change when it is taken out of its cage (a place it knows well) and placed in an unknown territory?
Level II	What does the mouse do when different toys are placed in its cage?
Level III	How does a gerbil behave on different surfaces (plain floor, litter, sand)? dependent variable: behaviour of the gerbil; independent variable: surface
Level IV	Theoretical knowledge about gerbils: gerbils are social animals which live together in families. Research question: Do gerbils react differently when placed in the observation arena on their own, with another gerbil they usually live together with, or with an unknown gerbil? dependent variable: behaviour; independent variables: alone/second gerbil; familiarity with second gerbil

Generating hypotheses	
Level I	Help card: The gerbil explores its new territory by running around and sniffing.
Level II	The gerbil is cautious at first, then explores the toys. After a while, it starts to play.
Level III	It behaves differently. On litter, it will run around as usual, because it is used to litter from its cage. On the plain floor, it is going to struggle because its feet are not made for walking on smooth surfaces.
Level IV	When the gerbil is alone in the unknown observation arena, it immediately starts to explore its territory. When a familiar gerbil is present, it first establishes contact and then explores the territory. If the second gerbil is a stranger, the gerbils will have to get to know each other before exploring the territory. Explanation: Gerbils are not as cautious with familiar mice as they are with unfamiliar ones. If they do not have to establish social hierarchies, they can start exploring earlier. Alternative hypotheses: The gerbil always starts exploring first. It does not matter if other familiar or unfamiliar gerbils are present.

Some 39% of 8th grade students (15 students) managed to draw up an adequate research question on their own. In addition, 5.6% (2 students) managed to draw up a research question and to identify dependent and independent variables. Higher-grade students reached higher competency levels. About a quarter of the students in grade 10 managed to draw up a research question and identify dependent and independent variables, whereas in grade 12, over 80% (31 students) mastered that task (Figure 1). Students of different grades differed significantly in terms of the competency level reached (χ^2=25.11; p

<.001).

For "generating hypotheses", 44.4% (17 students) from grade 8 reached Level II, meaning that they managed to generate adequate hypotheses. Another 11% (4 students) managed to explain their hypotheses. In grade 10, twice as many students reached Level III and nearly 20% (7 students) reached Level IV. In grade 12, 70.6% of students (27 students) managed to generate hypotheses (Figure 1). The difference between students from various grades is significant (χ^2=6.21; p <.05).
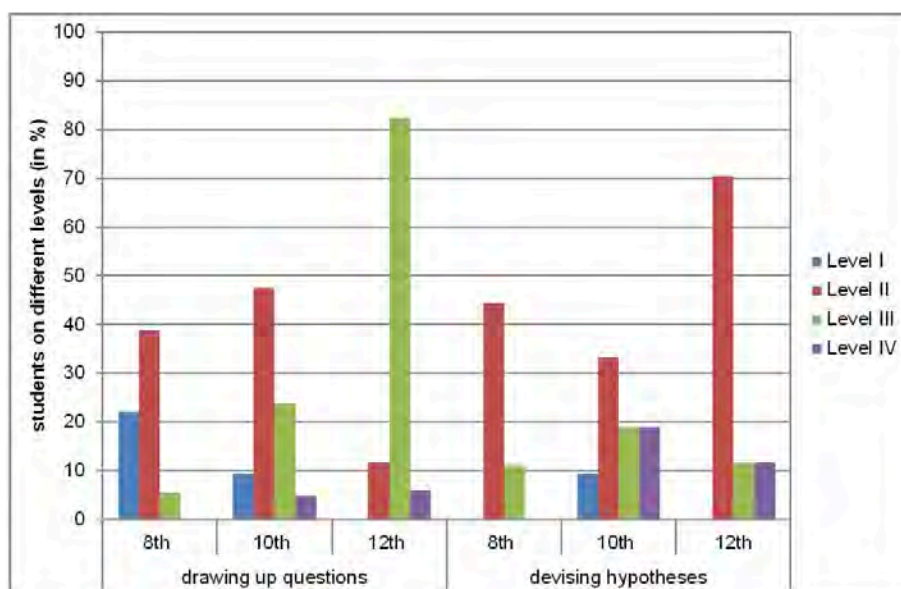
Figure 1 Percentage of students on Levels I, II, III, and IV for "drawing up questions" and "devising hypotheses".

As far as planning an investigation is concerned, students demonstrated low competency levels (Table 4). In 8th grade 28.8% of students (11 students) managed to plan an investigation which could be carried out by using the materials provided. About 10% (4 students) considered confounding variables or repeat measures, 2.6% (1 student) considered both confounding variables and repeat measures. Although about 10% more students (which means 15 students overall) than in grade 8 reached Level II in grade 10, differences in terms of Levels III and IV are negligible. Only in 12th grade, about a third of students considered either confounding variables or repeat measures. The

difference between the grades is highly significant (χ^2=8.39; p <.05). Students did not use the help cards provided.

Table 4 Examples for answers on Levels I, II, III, and IV for "planning an investigation" and "making observations".

Planning an investigation (research question: How does the behaviour of a gerbil change when it is taken out of its cage (a place it knows well) and placed in an unknown territory?)	
Level I	Help card: The behaviour of the gerbil is observed in the cage. Then, an observation arena is prepared with litter. The gerbil is put in the observation arena. The behaviour of the gerbil is observed in the observation arena.
Level II	First, I observe what the mouse does in its cage for about five minutes. Then, I take it out and put it in the arena. The arena is empty except for the mouse. Then, I observe what the mouse does in the arena for another five minutes.
Level III	First, I observe the behaviour of the mouse in its known territory. I write down everything the mouse does. Then, I prepare the arena. The arena must be clean and empty. Then, I take the mouse out and put it in the arena. Then, I observe what the mouse does in the arena for another five minutes. I have to take notes again.
Level IV	First, I observe the behaviour of the mouse in its cage. I have to count how often the mouse does certain things. I have to be quiet as not to disturb the mouse. Then, I put some litter in the arena. Then, I observe the behaviour of the mouse in the arena. Later, I must repeat the investigation with a second mouse.

Making observations	
Level I	The mouse was running around all the time. It often got on its hind legs to look out.
Level II	Using the information sheet provided and counting how often the five behaviours listed on the sheet occur
Level III	Using the information sheet provided and counting how often the five behaviours listed on the sheet occur (giving additional thoughts about how to perform the observation: time, videotaping, how to make notes)
Level IV	I would first make a list of all the behaviours of the gerbils. Then, I would draw up a table for these behaviours in each observing situation. I would count how often these behaviours occur within a 10-minute period in each situation.

Students' competency levels for "making observations" were not much higher than those for "planning an investigation" (Figure 2). In 8th grade, 23.7% of students (9 students) managed to describe observations without considering special observation criteria. About 15% of students (6 students) managed to use the observation sheet correctly and 10.5% (4 students) had additional ideas for observation criteria. In 10th grade, the percentage of students who mastered those tasks doubled, and in 12th grade, half of the students gave additional ideas for observation criteria. 5.3% of students (2 students) even planned an observation in line with the observation criteria on their own. The difference between grades is highly significant (χ^2=18.72; p <.001).
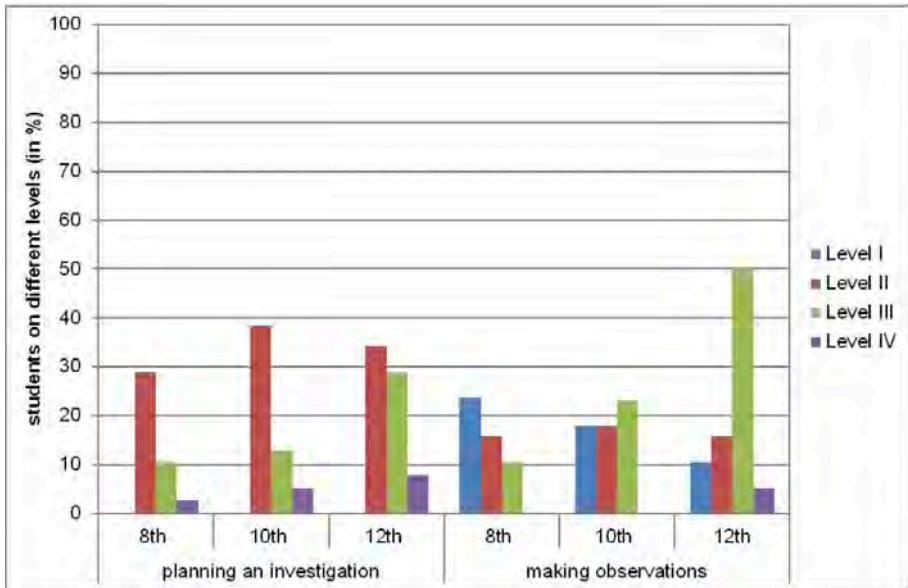
Figure 2 Percentage of students on Levels I, II, III, and IV for "planning and investigation" and "making observations".

Two competencies were investigated for handling the data of experiments and investigations: analysing data and drawing conclusion from data (Table 5).

Table 5 Examples for answers on Levels I, II, III, and IV for "analysing data" and "drawing conclusions from data".

Analysing data	
Level I	The mouse got on its hind legs about 20 times. Besides that, it did not do anything special.
Level II	Not investigated.
Level III	Description: The mouse ran around most of the time. It got on its hind legs eight times. Once it hid in the toilet paper roll. The mouse never bathed in the sand, it just ran over the sand bath. It's somewhat strange. Although we all observed the same situation, P. counted that the mouse got on its hind legs only six times. Maybe he did not watch attentively enough.
Level IV	Description: The mouse got on its hind legs 13 times in different locations. Analysis: It wanted to obtain a better picture of the place. Description: The mouse never bathed in the sand. Analysis: Mice never bathe when they are nervous, so I think the mouse was uncomfortable in its arena.

Drawing conclusions from data	
Level I	Not investigated.
Level II	Five minutes are not enough for the mouse to feel comfortable.
Level III	It was noticeable that mice are eligible prey animals, and therefore protection mechanisms like carefulness and agility are well developed.
Level IV	Not investigated.

For analysing data, Level I requires the description of data. 35% of 8th grade students (14 students) managed to complete that task. In grades 10 and 12, the percentage of students on Level I declines (Figure 3). The most common error on these levels was that students confused describing data with interpreting data. Level II was not surveyed. On Level III, students had to compare their own data descriptions with those of other students. Only in grade 12, students reported comparisons with other students' data. For reaching Level IV, students had to describe data and give adequate biological explanations for their data. In grade 8, about 25% of students (10 students) reached Level IV, whereas nearly 40% of grade 12 students (15 students) mastered that task. The differences between grades are not significant ($\chi^2=1.57$; p≥.05).
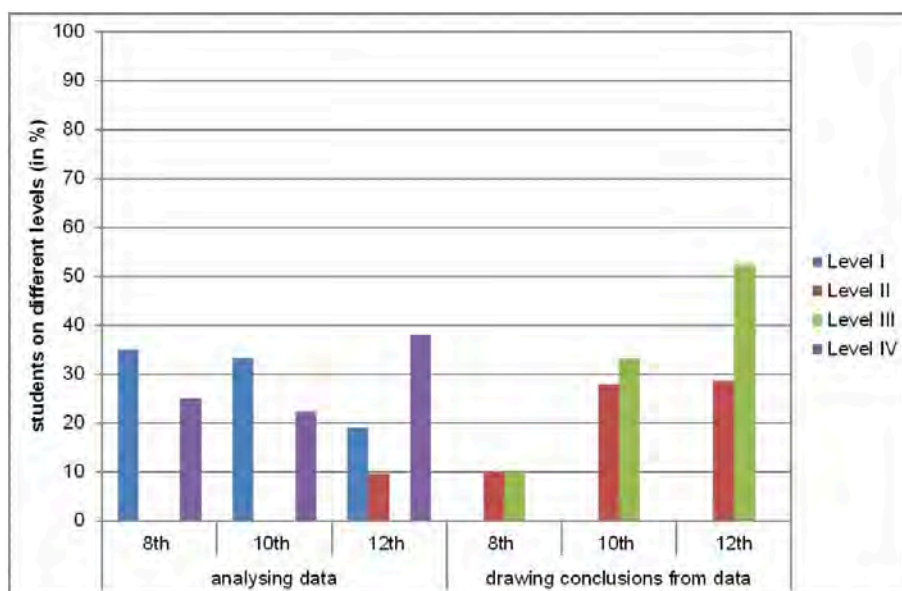
Figure 3 Percentage of students on Levels I, II, III, and IV for "analysing data" and "drawing conclusions from data"

For the competence of "drawing conclusions from data", only Levels II and III were surveyed in the study. In grade 8, about 10% of students (4 students) managed to draw their own conclusions. The number of students on Level II rose to nearly 30% in grades 10 (11 students) and 12 (11 students). The percentage of students on Level III, where biological justification for the

conclusions had to be given, rose form 10% in grade 8 (4 students) to 33.3% in grade 10 (13 students) and 52.4% in grade 12 (20 students). The difference between grades is significant (χ^2=14.77; p <.01)

6. Discussion

One aim of this study was to complement the data from large-scale paper and pencil assessments about inquiry competencies. To this end, experimental settings were created to measure students' competencies in the subject of biology. As animals are popular with students of different ages (Kellert, 1985), observations of gerbils were put in an inquiry based teaching setting. Based on the Austrian biology curriculum, gerbils or behaviour observations are taught in a couple of classes (BGBl. II 277/2004; BGBl. II 219/2016). While the focus in junior classes is on imparting knowledge of facts concerning animals and their handling, the focus shifts towards behaviour observation and the promotion of inquiry competencies in senior classes. As inquiry competencies are mentioned explicitly in the curriculum for senior classes, students of these grades are expected to demonstrate higher competency levels. On the other hand, if hands-on activities are used in biology classes, they are often teacher-guided and do not offer possibilities for students to perform inquiry their own (Grafendorfer & Neureiter, 2009). As all inquiry tasks in the present study were student-centred, students of all grades demonstrated insufficient inquiry competencies. Especially higher levels of competence were seldom reached, even in higher grades. Many students focused on finding an interesting research question and on generating corresponding hypotheses, but they neglected further tasks such as naming dependent and independent variables or stating their hypotheses. Lower-grade students frequently choose research questions which would have put the gerbils at risk or could not be investigated using the materials provided. Students of all grades had difficulties planning their own experiments and making correct observations without simultaneously giving interpretations. After collecting their data, students had problems analysing them properly. They did not compare their data to other students and, if they did, they did not engage in any discussion of why different students made different observations. In some groups, students compared their results and then made corrections to their observations.

As far as competency development is concerned, prior studies in junior

classes did not show any competency improvement between grades 5 and 8 (Schiffl et al., 2015). For senior classes, research on competency development is yet to be done. In this study, students of higher grades outperformed students in all competencies except data analysis. This suggests an acknowledgement of competency improvement in senior classes. Compared to the results of paper and pencil competency evaluations based on the same competency levels (Schiffl, 2017), students performed worse in the present study. Especially Level I was reached less often. This is due to the need to use help cards on this level. Students often did not take the initiative to request help cards and the investigator did not always notice if somebody needed a help card.

As far as the use of hands-on activities in the diagnosis of inquiry competencies is concerned, age limitations must be considered as, with decreasing age, students experienced greater problems with handling gerbils and conceptualising experiments on their own. Although mammals are a crucial topic of the 5th grade science curriculum (BGBl. II 277/2004), it is not advisable to use hands-on experiments with animals at such an early stage of science education, especially if only one teacher is available and if students have so much freedom in choosing their own experimental settings. For older students, especially for senior classes, it is an adequate tool for diagnosing inquiry competencies, in particular, if students are used to inquiry-based learning.

References

Anand Pant, H., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (Hrsg.) (2013) *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I.* Münster: Waxmann.

Bundesgesetzblatt für die Republik Österreich (2004*) BGBl II, 277/ 2004: Änderung über die Verordnung über Lehrpläne der allgemein bildenden höheren Schulen*, ausgegeben am 08. Juli 2004.

Bundesgesetzblatt für die Republik Österreich (2009) *Teil II: 1. Verordnung: Bildungsstandards im Schulwesen. Teil II,* ausgegeben am 2. Jänner 2009.

Bundesgesetzblatt für die Republik Österreich (2016) *BGBl II, 219/ 2016: Änderung über die Verordnung der Lehrpläne der allgemeinbildenden höheren Schulen*, ausgegeben am 09.08.2016.

Bundesinstitut für Forschung, Innovation und Entwicklung des österreichischen Schulsystems (BIFIE) (2011) *Kompetenzmodell Naturwissenschaften 8. Schulstufe,* Retrieved October, 24, 2016 from: https://www.bifie.at/system/files/dl/bist_nawi_kompetenzmodell-8_2011-10-21.pdf

Department for Education (2014) *The National Curriculum in England*: Framework document. Retrieved May 26th from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/381344/Master_final_national_curriculum_28_Nov.pdf

EDK (2011) *Grundkompetenzen für die Naturwissenschaften.* Retrieved May 26th from http://edudoc.ch/

record/96787/files/grundkomp_nawi_d.pdf.

Grafendorfer, A. & Neureiter, H. (2009) Unterricht in Naturwissenschaft. In C. Schreiner & U. Schwantner (Hrsg.), *PISA 2006: Österreichischer Expertenbericht zum Naturwissenschaftsschwerpunkt*. Graz: Leykam.

Haury, D. L. & Rillero, P. (1994) *Perspectives of hands-on science teaching*. Columbus, OH: ERIC Clearinghouse for Science, Mathematics and Environmental Education.

Holstermann, N., Grube, D. & Bögeholz, S. (2010) Hands-on activities and their influence on students' interest. *Journal of Research in Science Education*, 40, 743-757.

IQB (2013) *Kompetenzstufenmodelle zu den Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss*. Retrieved November, 10, 2016 from: file:///C:/Users/User/AppData/Local/Temp/KSM_Biologie.pdf

Kellert, S. R. (1985) Attitudes towards animals: Age related development among children. In M. W. Fox et al. (Eds.), *Advances in Animal Welfare Science* (pp. 43-60). Martinus Nijhoff Publishers.

KMK (2005) *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004.* München: Luchterhand.

Lombardi, S. A., Hicks, R. E., Thompson, K. V. & Marbach-Ad, G. (2014) Are all hands-on activities equally effective? Effect of using plastic models, organ dissections, and virtual dissections on student learning and perceptions. *Advances in Physiology Education*, 38, 80-86.

Lumpe, A. T. & Oliver, J. S. (1991) Dimensions of hands-on science. *The American Biology Teacher,* 53(6), 345-348.

Mayer, J., Wellnitz, J., Klebba, N. & Kampa, N. (2013) Kompetenzstufenmodelle für das Fach Biologie. In H. Anand Pant, P.   Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), IQB-Ländervergleich 2012: *Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 74-82). Münster: Waxmann.

Middleton, J. A. (1995) A study of intrinsic motivation in the mathematics classroom: A personal constructs approach. *Journal for research in mathematics education*, 254-279.

National research Council (2012) *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academy Press.

Pedaste, M. et al. (2015) Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, 14, 47 - 61.

Roth, W.-M. & Roychoudhury, A. (1993) The development of science process skills in authentic contexts. *Journal of Research in Science Teaching, 30(2),* 127-152.

Satterthwait, D. (2010) Why are "hands-on" science activities so effective for student learning? *Teaching Science*, 56(2), 7-10.

Schiffl, I. (2016) Die Situation naturwissenschaftlicher Bildungsstandards in Österreich am Beispiel der Biologie. *Zeitschrift für Didaktik der Biologie, 20. Jahrgang,* 44-61.

Schiffl (2018) *Kompetenzdiagnostik im Biologieunterricht: Entwicklung von Diagnosetools zu Erhebung biologischer Kompetenzen durch die Lehrperson.* Hamburg: Verlag Dr. Kovac.

Schiffl, I. & Weiglhofer, H. (2016) Kompetenzdiagnostik im Biologieunterricht am Beispiel des Instruments zur informellen Kompetenzmessung. In G. Maresch & J. Zumbach (Hrsg.), Didaktik der Naturwissenschaften: *Neu Horizonte in Biologie, Geometrie und Informatik* (S. 27-41). Wien: Facultas.

Schiffl, I., Weiglhofer, H. & Remplbauer, D. (2015) Development of inquiry competencies during junior classes: A replication study using longitudinal design. S*cience Education Research Letters*, Volume 2015, 16 -20. Retrieved October, 30, 2016 from http://edoc.hu-berlin.de/serl/2015/3/

PDF/2015-3.pdf

Schwichow, M., Zimmermann, C. Crove, S. & Härtig, H. (2016) What students learn from hands-on activities. *Journal of Research in Science Teaching*, 53, 980-1002.

Singapore Ministry of Education (2014) *Science Syllabus: Lower and upper secondary: normal (technical course).* Retrieved May 26th from https://www.moe.gov.sg/docs/default-source/document/education/syllabuses/sciences/files/science-lower-upper-secondary-2014.pdf

Stohr-Hunt, P. M. (1996) An analysis of frequency of hands on experience and science achievement. *Journal of Research in Science Teaching*, 33(1), 101-109.

Weinert, F. E. (2014) Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.) *Leistungsmessung in Schulen,* 3. aktualisierte Auflage (S. 17-31). Weinheim: Beltz.

Teachers' and students' opinions about the use of a motivational context-based biology teaching-learning module: a case study

Ana Valdmann[1] and Miia Rannikmäe[1]

ana.valdmann@ut.ee

Abstract

A teaching module called "Gym versus running: Is a trained person stronger, smarter, more durable and more attractive?" was developed to improve students' scientific literacy through context-based biology and higher order thinking skills, for the new Estonian Biology Curriculum. Nine teachers who taught the module and 156 students were included in the study. The research instruments used were interviews, feedback questionnaires and meeting records. Both students and teachers mentioned the essay aspect of the module as being very challenging. The teachers' main concerns were the absence of necessary evaluating, argumentation and reasoning skills, as they usually paid little attention to the argumentation in biology lessons. We concluded that context-based teaching with socio-scientific health problems and inquiry based learning are effective tools for students to acquire more knowledge in anatomy and physiology which are parts of biology syllabus.

Keywords: context-based teaching, decision-making, inquiry-based learning, socio-scientific issue, 3-stage model

---

[1] University of Tartu, Estonia

1. Introduction

In line with the European trend, which emphasises competence-based education, Estonia introduced a new curriculum in 2011. For a successful educational reform, teachers now needed to adjust their perceptions to the reform's new curriculum and strategies and to cope with new content, plus teaching and assessment strategies. Developing students' scientific literacy, through context-based biology and higher order thinking skills, was the framework for establishing a new biology curriculum for Estonian school students. Within this frame, we developed the "Gym versus running: Is a trained person stronger, smarter, more durable and more attractive?" which focused on context-based biology and higher order thinking skills. The module introduced an understanding of the circulation and respiration systems, as proposed in the biology curriculum. Learning started with a relevant socio-scientific scenario, which motivated students to learn biology. Through inquiry learning in groups, students used data-loggers to acquire more knowledge of the morphology and function of these systems. They also learned about the effect of using the gym and running on the body systems of athletes, and discussed which form of training was preferred and why. Finally discussion and socio-scientific decision-making took place.

Our research objectives were:
(a) to identify the challenges and difficulties biology teachers faced, as well as the advantages they found while teaching the "Gym versus running: Is a trained person stronger, smarter, more durable and more attractive?" module; and
(b) to investigate how teachers and students coped with context-based teaching/learning.

Our research questions were:
(a) what are teachers' views towards advantages and difficulties they experienced while teaching the context-based module?
(b) how do teachers and students cope with a context-based teaching/learning module?

2. Theoretical Framework

The major goal of science education is to develop students' scientific literacy. According to the US National Science Education Standards, students need to be able to make decisions about topics that are interdisciplinary in nature (NRC, 1996).

2.1 Learning context

In achieving this goal, teachers need to be able to account for learning science in context, and teaching concepts in relationship to real-world context is expected to make science education more meaningful, relevant, and motivating for students (Gilbert, 2006). In slightly different forms and with different purposes, context-based education has been propagated in response to alleged failures of the traditional science curriculum (Gilbert, 2006).

2.2 3-stage model

Holbrook & Rannikmäe (2010) propose a 3-stage model (Figure 1) based on the "Education through Science" philosophy (Holbrook, 2008) and Activity Theory (van Aalsvoort, 2004).
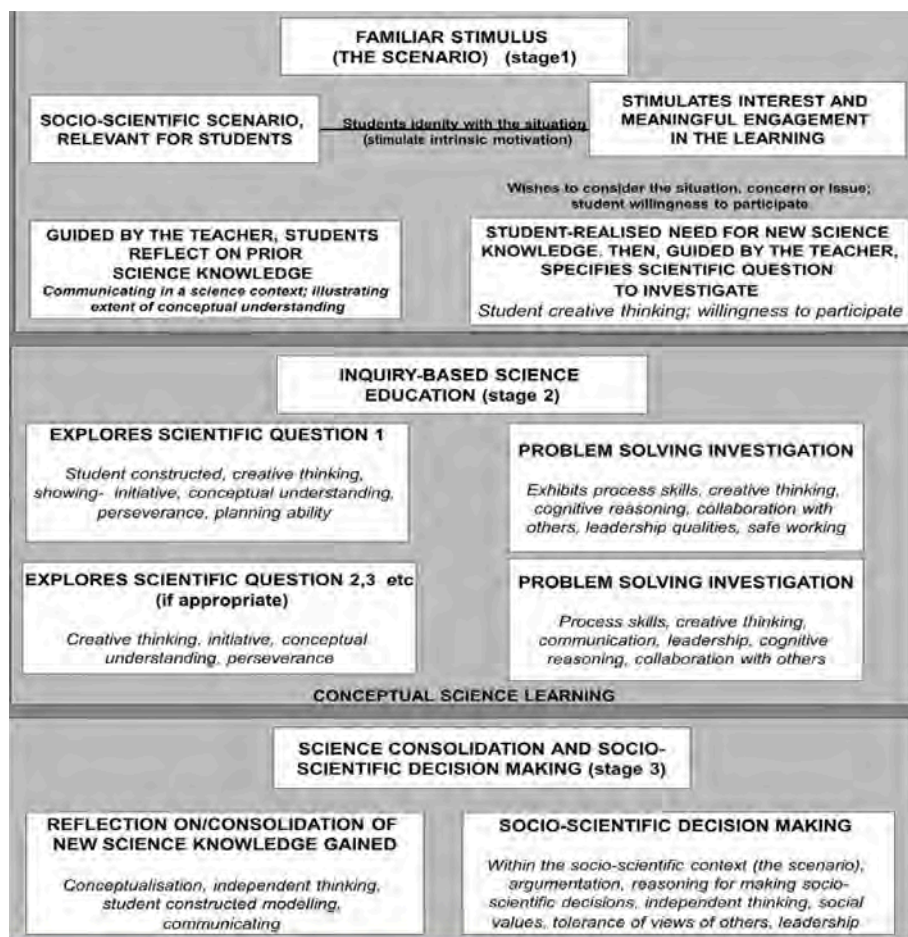
Figure 1 Description of 3-stage model of science teaching (Holbrook & Rannikmae, 2010).

2.3 Relevance

The relevance of the context is established via a socio-scientific issue or concern, presented through a motivational scenario. From this situation, a question or problem arises logically. Students answer it by performing learning activities, meanwhile gaining insights into the biological concepts needed to answer the question or solve the problem (Bennett et al., 2007). This stage also seeks to establish students' prior science background, while the follow up learning takes place within a Zone of Proximal Development (ZPD) (Vygotsky, 1978).

2.4 Inquiry-based learning

The science learning is promoted in an Inquiry-based Learning (IBL) format - as structured, guided, or open (Spronken-Smith et al., 2011), leading to the acquisition of new conceptual science. Interlinking with other previously acquired science concepts consolidates this new science knowledge. The acquired science is linked to the initial scenario issue via involving students in argumentation and reasoned decision-making, which are important skills related to future life.

The presented module is created accordingly as the 3-stage model (Holbrook, 2010; Holbrook & Rannikmäe, 2010; 2014) and the Estonian Biology Curriculum (EC, 2011), related to promoting an understanding of circulation and respiration systems. It is hypothesised that this promotes learning geared towards making a decision based on sound arguments and planning, conducting and interpreting findings from an investigation and developing career-related skills, such as communicating orally, writing conclusions and cooperating as a member of a group.

3. Research design and methodology

3.1 Design of the lesson module

Experts from Tartu University and one of the participating teachers created a context-based module compliant with the 3-stage model.

The reasons why this module was created were:

1. Students are generally interested in health and themes related to themselves. So it helps to motivate students to learn science (relevant and motivational to students) (Rannikmae et al., 2010)
2. Health and safety are cross-cutting themes in the Estonian curriculum (Estonian Curriculum, 2011)
3. The module allows the linking of subjects (biology, chemistry and physics) at the high school level, and hence promotes interdisciplinarity within basic school biology content (respiratory and circulatory systems)
4. It is in line with the development of general competencies as expressed in the Estonian curriculum (social and self-determined learning through student-centred learning and encouraging students to reflect on valuing a healthy lifestyle)

Competencies to be achieved are:
1. Being able to decide, based on sound arguments, whether people possess the responsibility to take care of their own health (competence needed in everyday life)
2. To understand that 'best for my health' can have more than one interpretation and to suggest the most appropriate meaning in this context; value healthy lifestyle (personality)
3. To design an investigation, interpret experimental instructions and carry out an experimental procedure (academic skills)
4. Being able to communicate orally and by means of a written conclusion, cooperating as a member of a group (competencies needed for a career)
5. Understanding the design and function of respiratory and circulatory systems, homeostasis, short-term and long-term training effects on human functioning, aerobic and anaerobic glycolysis, analysis of the impact of training on functions of circulatory system and health (academic knowledge)

Stage 1 Introducing the socio-scientific issues scenario (45 minutes/ one lesson)

Why the title: *Gym versus running: Is a trained person stronger, smarter, more durable and more attractive?"*

The first part draws attention to the social aspect. As using a gym is expensive, so is there a cheaper alternative? This leads to considering questions such as: Why do people prefer one to the other (gym versus running)? Do different training methods have differing effects on the body?

The second part draws attention to the science content. Is it true that a trained person is stronger, smarter, more durable and more attractive? The latter, in turn, instigates a discussion about whether beauty is in the eye of the beholder. The sentence also relates to the short-term and long-term training effects on human health and can engulf a number of sports physiology topics as well as cover aspects of anatomy.

Scenario

To initiate learning, the scenario used was a presentation with pictures and a dialogue between two boys.

Jaan and Jaak decided to run around the lake. At first, the boys were running

smoothly, allowing them to ask questions and discuss topics between each other. After two kilometres, Jaan began to pant and could no longer answer questions. Jaak started to goad him, "Hey, are you a weakling and not used to doing exercise? What sort of man are you growing into and what girl will want you?"

Jaan replied that he goes to the gym and he has powerful biceps, which girls like.

Jaak laughed about it: "Your biceps may be big, but the heart remains weak and you spend a lot of money."

Stage 2 Inquiry-base science investigation (90 minutes/ two lessons)

Investigatory work (planning an investigation to determine the impact of training on blood pressure, pulse rate and oxygen uptake, recovery rate and its dependence on training).

Stage 3 Socio-scientific decision-making (45 minutes/ one lesson)

Group discussions and writing an essay (My moving/exercise habits - to change these or not?).

The sample consisted of 9 teachers and 156 students (Table 1). All teachers were educated as biology and chemistry teachers, they all held a Master's degree and 14-21 years' experience (average 16.5). All teachers had participated in a 40-hour continuing professional development (CPD) programme, which was specifically designed, based on teacher needs to operationalization the 3-stage model (Holbrook et al., 2014; Valdmann et al., 2016) in science teaching.

Table 1 The teachers and students involved in the study.

School level	No. of Teachers	No. of students
Basic (grades 1-9) school	6	112 (from 5 schools, grade 9)
High (grades 10-12) school	3	44 (from 3 schools, grade 11)

The research tools included interviews, teacher-student feedback questionnaires and records made from a teacher meeting. The semi-structured interview (Patton, 1990) included a pre-prepared set of questions for comparing the teacher's views with the module goals. The main semi-structured interview question was the following: *What advantages did you find and what difficulties did you encounter while teaching in context?* The interviews were recorded and the first author subsequently analysed the interviews. The interview data were content analysed.

Teacher-student feedback questionnaires consisted of 13 yes/no and 9 mul-

tiple-choice questions, using a five-point scale (i.e. not at all = 1…very much = 5) questions for students. Percentages for yes/no and mean and standard deviations for five-point scale multiple-choice questions were calculated.

4. Results and discussion

4.1 Stage 1 - socio-scientific issues scenario

According to the 3-stage model, the first stage aimed to stimulate students' interest and meaningful engagement in the learning. All teachers and students found that title of module stimulating because it emphasised the controversial aspects and encouraged discussion. 57% of students found that their class peers had different perceptions. Students also found that the discussion after the scenario was interesting and important (mean = 3.7; SD=0.74). Two teachers found that the scenario was not relevant for their students and changed it:

Teacher A: *I changed the scenario because I teach in a small country school and in our village we do not have a gym. The nearest gym is 30 km away. I only focused on running and sports games outdoors. (Basic school level teacher)*

Teacher B: *I changed the scenario because I teach in a sports gymnasium and I thought that my students are more interested in training facilities (clothes and shoes suitable for different sports) and supplements for athletes. (Gymnasium level teacher)*

All teachers and 98% of the students agreed that the socio-scientific context was helping to motivate students to learn science. 78% of teachers and 65% of students found that the presented context (ie the health context) highly motivated learning. Previous studies (Rannikmäe et al., 2010) also demonstrated students' special interest in health-related issues. We concluded that students were more motivated while learning anatomy associated with health problems. Difficulties in guiding student discussion were noted as a main concern by most teachers. According to the questionnaire we found that the discussion weakly supported the need to look for new information to solve the scientific problem which arose from the scenario.

4.2 Stage 2 - the inquiry-based science education

The aim of the second stage is to obtain a new conceptual understanding through inquiry learning. Students use data-loggers to acquire more knowledge of the morphology and function of respiration and circulation systems and analyse the impact of training on cardiovascular functions of the body.

Generally, students preferred working in groups (m=4,1; SD=0,74) rather than inquiry-based learning (m= 3,1 SD= 0,99). Previous studies (Kask & Rannikmae, 2006) demonstrated low readiness of Estonian teachers to promote inquiry skills among students, but in our study all nine teachers were well prepared and confident to use inquiry teaching. Teachers used structured or guided inquiry but rarely an open format. At the same time, 56% teachers noted that students had difficulties to set hypotheses and 44% noted difficulties in formulating research questions. This stemmed from the students low process skills and the need to use more inquiry in Estonian schools. All teachers evaluated using data loggers as a challenge.

4.3 Stage 3 - socio-scientific decision-making

The aim of the third stage was to present arguments about why the gym or running is good for health and society. As homework, after the class discussion, students were required to write an essay on "My movements and habits - change or not?" The most frequent response from teachers was that there was little discussion of the science problems, which were detected in the scenario with limited attention to economic, social and ethical aspects. 11% of teachers and 36% of students claimed that learning usually ended with presentation of research results without discussion. Only 6% of students and 33% of teachers agreed that teaching/learning should lead to socio-scientific decision-making. Teachers also admitted their limited ability to guide socio-scientific decision-making and difficulties to reach back to the socio-scientific issue presented in the scenario. At the same time, students found the problem solving interesting (mean = 3.6; SD=0.86) and important (mean =4.0; SD=0.83).

With respect to students' writing essays, their ability to use acquired knowledge in changing their personal physical habits according to socio-scientific decision-making process was found to be quite limited. Also, teachers noted difficulties in assessing essays and generally found it time consuming.

4.4 The overall assessment of the module

The overall assessment questions on the module were connected with different skills in line with the competence-based curriculum in Estonia (see Table 2 and 3). Eight teachers felt that students were more involved in the thinking process and in developing higher order thinking skills (T89%, S64%). The most significant differences between teachers' and students' opinions were in

creativity (T 56%, S 18%) and argumentation skills (T 56%, S 18%). In Stage 2, students clearly preferred working in groups. In this context, the expected results were that, by their evaluation, the module developed collaboration skills (70%), but unexpectedly low were presentation skills (42%) (each group made a presentation at the end). Students found that module-based lessons were more interesting (mean =4.2; SD=0.86) and differed from general lessons, because they had more possibilities for discussion (mean = 4.3; SD=0.86). Module-based lessons were more related to everyday life (mean = 4.0; SD = 0.93), more interesting (mean = 4.2; SD = 0.86) and not too difficult (but students felt they had to work harder to understand mean = 2.6; SD = 0.91).

Table 2 Teacher and student feedback on the 3-stage module (frequency of yes responses).

Stages and statements	Teachers (N=9)	Students (N=156)
Stage 1 (yes/no question)		
Scenario was connected to everyday life	78%	65%
Socio-scientific context motivated learning	100%	98%
The discussion revealed the need for information to solve the problem	67%	45%
The title of the module stimulated me/ my students to learn biology	100%	100%
Stage 2 (yes/no question):		
I liked the inquiry part of the module	89%	63%
Students formulated the research questions themselves	44%	37%
The hypotheses were difficult to set	56%	38%
Experiments were planned by students	56%	44%
Group work was the best learning part of the module	56%	79%
Stage 3 (multiple choice):		
Module ended by:		
a) a common discussion of the science problems which were detected in the scenario	56%	58%
b) Presentation of the experimental results	11%	36%
c) Socio-scientific decision-making activity	33%	6%
By my assessment, the module developed (yes/no question):		
Collaboration skills	89%	70%
Presentation skills	78%	42%
Argumentation skills	56%	19%
High order thinking skills	89%	64%
Creativity	56%	18%

Table 3 Student feedback on 5-point scale questions (1- not at all; 5-very much)

	Mean	Standard deviation.
Discussion after scenario was interesting and important for me		
I like to work in groups	3.7	0.74
I like inquiry-based learning	4.1	0.74
This problem (problem solving) was interesting for me	3.1	0.99
This problem (problem solving) was important for me	3.6	0.86
Module based lessons are more interesting	4.0	0.83
More possibility to discuss	4.2	0.86
Module-based lesson was more related with everyday life	4.3	0.86
Had to work harder to understand the science content	4.0	0.93
	2.6	0.91

5. Conclusion

According to our results, a socio-scientific context is useful in motivating students to learn biology, especially in the context of health-related issues. Both students and teachers found that inquiry learning and using data loggers helps to understand the science content, and makes learning more interesting. Students value group work and collaboration, considering that it promotes

their learning process. The training of teachers and students in science classes should pay more attention to promoting argumentation skills, values and attitudes.

References

Bennett, J., Lubben, F. & Hogarth, S. (2007) Bringing science to life: A synthesis of the research evidence on the effects of context based and STS approaches to science teaching. *Science Education*, *91*(3), 347-370.

Estonian Curriculum (2011) *National Curriculum for Basic schools and upper secondary schools*. Regulation of the Government of the Republic of Estonia, Rti, 14.01.2011.

Gilbert, J.K. (2006) On the nature of "context" in chemical education. *International Journal of Science Education,* 28(9), 957 -976.

Holbrook, J. (2008) Promoting Valid Assessment of Learning through Standardised Testing. In: J. Holbrook, M. Rannikmäe, P. Riiska & P. Isley (eds.). *The Need for a Paradigm Shift in Science Education for Post-Soviet Countries, 216-231. Frankfurt: Peter Lang.*

Holbrook, J. (2010) Education through science as a motivational innovation for science education for all. *Science Education International*, 21(2), 80-91.

Holbrook, J. & Rannikmae, M. (2010) Contextualisation, de-contextualisation, re-contextualisation – A science teaching approach to enhance meaningful learning for scientific literacy. In: I. Eilks & B. Ralle (Eds.). *Contemporary Science Education – Implications from Science Education Research about Orientations, Strategies and Assessment, pp 69-82. Aachen: Shaker Verlag.*

Holbrook, J. & Rannikmäe, M. (2014) The Philosophy and Approach on which the PROFILES Project is based. *CEPS journal, 4(1), 9-29.*

Holbrook, J., Rannikmäe, M. & Valdmann, A. (2014) Identifying teacher needs for promoting Education through Science as a paradigm shift in Science Education. *Science Education International, 25(2), 4-42.*

Kask, K. & Rannikmäe, M. (2006) Estonian Teachers' Readiness to Promote Inquiry Skills Among Students. *Journal of Baltic Science Education, 1(9), 5-15*

National Research Council. (1996) *National science education standards*. Washington, D, C.: U.S. Department of Education. NRC, 2010.

Patton, M,Q. (1990) *Qualitative evaluation and research methods*. Sage, Newbury Park, CA

Rannikmäe, M., Teppo, M. & Holbrook, J. (2010) Popularity and Relevance of Science Education Literacy: Using a Context-based Approach. *Science Education International*, *21(2), 116-125.*

Spronken-Smith, R., Walker, R., Batchelor, J., O'Steen, B. & Angelo, T. (2011) Enablers and constraints to the use of inquiry-based learning in undergraduate education. *Teaching in Higher Education*, *16*(1), *15-28.*

Valdmann, A., Rannikmae, M. & Holbrook, J. (2016) Determining the effectiveness of the CPD programme for enhancing science teacher's self-efficacy towards motivational context-based teaching. *Journal of Baltic Science Education*, *15(3), 284 – 297.*

Van Aalsvoort, J. (2004) Activity theory as a tool to address the problem of chemistry's lack of relevance in secondary school chemical education. *International Journal of Science Education, 26, 1635-1651.*

Vygotsky, L. (1978) Problems of Method. In: *M. Cole (trans.). Mind in Society. Cambridge, MA: Harvard University Press.*

SECTION TWO –

Papers based on the panel debate: *Considering grand challenges in biology education: rationales and proposals for future investigations to guide instruction and enhance student understanding*

Based on a panel debate about challenges for biology education at the ERIDOB conference in Karlstad the following section of short papers have been written by a selection of researchers representing four continents. The panel debate was chaired by William F. McComas.

William F. McComas (University of Arkansas, USA), **Michael J. Reiss** (University College London, UK), **Edith Dempster** (University of KwaZulu–Natal, South Africa), **Yeung Chung Lee** (The Education University, Hong Kong), **Clas Olander** (Malmö University, Sweden), **Pierre Clément** (University Aix-Marseille, ENS de Lyon-IFE, France), **Dirk Jan Boerwinkel** and **Arend Jan Waarlo** (Utrecht University, Netherlands)

Introduction

William F. McComas[1]

mccomas@uark.edu

The educational research enterprise is alive, well, and thriving. Each year increasing numbers of researchers enter the field, and each year substantial numbers of papers on all facets of education are generated. We are awash in a sea of publications and data focused on educational matters, but this tidal wave of studies, opinions, and recommendations has done less than we would hope to change the daily practices of teachers or result in enhanced learning outcomes in truly fundamental ways.

One reason for this unsettling conclusion is that the results of our research endeavours are too often unknown to those policymakers, teacher educators, administrators – and, most importantly, teachers – who are in a position to take action based on the flood of conclusions and suggestions for enhancement. Sadly, some even expressly ignore research findings because of contextual misunderstanding; some teachers, for instance, think that because the study was not conducted in their school, it cannot possibly pertain to them. More understandably, the demands on teachers are frequently so severe that one might well question where they would find the time and resources to read, assimilate, and put into action even those recommendations that seem promising. However, there is at least one more reason why much of the work engaged in by educational researchers is ignored. Many of the studies done are tiny, conducted by one or just a few researchers on odd or marginal topics that result in non-generalizable results that were never necessarily topics of interest to the community of practice anyway.

If true, what this means to those in the educational research community is that we must take some responsibility for our own marginalization – this is true both in general and within the field of science education specifically. How many among us, when beginning a new project, ask the question "So, what is the most important question that I can attack next that educators really need to have addressed?" Rather, we often work on small projects of

---

[1] University of Arkansas, USA

personal interest, with little consideration for the importance of the underlying question or the ways in which the results will be disseminated to influence practice. We share our results primarily with each other in *our* journals and then move on. To validate this point, it might be interesting to examine the titles in any research publication to see how many seem to report results from truly transformative or even modestly important avenues of investigation with big enough sample sizes in diverse settings such that generalization is possible.

Despite this dire situation, some within science education have suggested another approach. For instance, consider the thoughts of authors assembled for a special issue of the journal *Science* (19 April, 2013) who have called for a consideration of "grand challenges" that, in turn, could guide research to address large problems in science learning. The Grand Challenges in Science Education offered by this group include issues such as the proper role for technology, an understanding of individual differences in brain development and how that affects learning, professional development for teachers, and what skills are necessary for teachers to implement high-quality laboratory teaching. They also include issues related to establishing personal relevance for learners and developing students' understanding of how science creates knowledge, along with a suite of suggestions about how teachers can engage in research at the school level and how assessment results can guide instruction. Articles in this issue address these questions from a broad science education perspective. The question remains as to who will assume the responsibility for conducting the research necessary to explore these challenges and how best the results can be shared with those who might take appropriate action.

Thinking about grand challenges in biology education?

Certainly, the challenges discussed in the special issue of *Science* are worthy of consideration. However, as Zogza (2016) reminds us, biology education (or biology didactics, as it is sometimes called)[2] is a unique discipline with special teaching and learning contexts. Biology education research therefore must be "aimed at highlighting and facilitating the process of teaching and learning about the biological world" (p. 181). She states that biology didactics is not

---

[2] Bayrhuber (2016) has explained that there is a distinction between didactics and teaching, but Gericke & Ottander (2016), who have also explored that issue, have reached the conclusion that it is acceptable to see the two as similar enough to consider them synonms. Therefore, it seems reasonable to use the terms "Research in the Didactics of Biology" and "Biology Education Research" interchangeably.

just a part of science education, because the science of biology is distinct from sciences such as physics and chemistry, and that, as such, this would dictate and nuance avenues and needs for research. Perhaps the best way to think about this is that biology education is a subset of science education with its own context and, to a degree, its own questions. It is reasonable to consider that all the grand challenges in science education might be explored through the lens of their implications to biology instruction, but it is likely that biology didactics will have its own set of issues worthy of focused exploration.

In past decades, many researchers have investigated programs in biology teaching and learning and have published their results in prestigious journals, and occasionally they have offered teachers implications stemming from those investigations. Groups that focus on biology education regularly endeavour to improve the state of affairs with respect to instruction, but rarely have there been focused attempts to identify the big issues on which the research community should focus its energies. Even more infrequently do we see a team approach to addressing big programmes in this domain. One group that, more than any other, has accepted some of the burden of focusing research on problems in biology instruction has been the European Researchers in Didactics of Biology (ERIDOB), even dedicating its 10th conference, in June 2014 in Haifa, Israel, to "The Future of Biology Education Research." This important question was revisited in September 2017 with a panel discussion at the 11th conference in Karlstad, Sweden, resulting in this set of short contributions.

Of course, until a much larger and representative group of scholars is convened for the purpose of suggesting avenues of future research, it would be impossible to suggest that this article offer a definitive list of the most important problems worthy of investigation. However, those assembled in Sweden and represented here offer interesting thoughts about necessary research on topics such as teacher education, the role of language, sociocultural issues, socio-scientific considerations, and other ideas related to biology didactics. So, perhaps we might call these suggestions *emerging* Grand Challenges in Biology Education and offer them in preparation for a larger explication of where work in biology didactics should be focused.

The suggestions that comprise the bulk of this article are contributed by a group of international scholars with their own areas of specialization, who were asked to provide some background in a specific area of investigation and

then offer specific questions or areas of research based on their own domain of interest. Thus, we begin with thoughts about a framework for biology education research that looks across the three domains: biology, education, and research itself. Again, while there is no hope of offering a list of all areas of necessary future research, we hope that you find these suggestions and their rationales compelling.

A proposed framework for biology education research

Michael J. Reiss[1]

m.reiss@ucl.ac.uk

There are several possible frameworks that might guide biology education research. One approach is to see such research as the point of intersection of three overlapping domains – biology, education, and research – as they relate in a Venn diagram-like fashion (Reiss, 2016).

If we start with *biology*, we might take an approach that derives from Hirst's (1965) "forms of knowledge," focusing on the distinctiveness of biology itself. For a start, biology sits within the natural sciences, which have a methodology that emphasizes knowledge as objective, universal, and amenable to rational inquiry. Within the natural sciences, of course, biology is the study of life. In a sense, our choice of subjects is vast because there are perhaps some 10 million extant species, each of which could, in principle, be investigated in any number of ways. The most important biology research often proceeds by studying a range of phenomena and/or species, which then permits making conclusions or constructing new models that are both widely applicable and amenable to such local variation. This approach is widely demonstrated in the work of Darwin, Mendel, the discoverers of the structure of DNA, and ecologists such as E. O. Wilson. There is a lesson here for biology education research: we surely want to engage in fine-grained research that is true to the particularities of a situation; we also want to be able to extrapolate to broader horizons.

If we start with *education*, we are beginning with what has sometimes been described not as a single discipline (like history, mathematics, or biology itself) but as a field. Like medicine and engineering, education draws on a wide range of more fundamental disciplines (e.g., psychology, sociology, and philosophy) to make its advances. This approach makes an epistemological point about knowledge production in education.

However, there is another way of starting with education, not from an epistemological standpoint, but from a normative one related to values and what we *want* from a good education. It has been argued that there are two funda-

---

[1] University College London, UK

mental aims of education: to equip each learner to lead a life that is personally flourishing, and to help others to do so too (Reiss & White, 2013). In this approach, biology education research contributes to such flourishing. Indeed, if "others" is understood to include nonhumans, this argument can be seen as manifesting an inclusive environmental education. At present, this argument for the fundamental aims of education seems powerful. We are so used to arguments about extinctions, climate change, and other threats to our continued existence that it can be difficult to keep in mind how exceptional, from a biological perspective, is the age in which we live. Indeed, the recent coining of the term "Anthropocene" is an attempt to remind us of this very fact and how unusual are the current times, when seen from an historical perspective.

We have deliberately started with biology and education because, along with many who would read this article, our shared experience when supervising doctoral students and researchers is that such individuals often start with *research*. To be sure, that is yet another way to begin. Researchers are expected to identify a gap in the literature, formulate research questions, and then derive a methodology and associated set of methods that permit a way to address these questions. However, while such an approach enables findings that add to the literature, such findings are unduly constrained by the accidents of history – since what has previously been investigated drives the identification of gaps in the literature and, thus, our own research. A better starting point is to combine the personal interests of the researcher(s) with an analysis of what *ought* to be researched. We need to keep in mind the purpose of our research, or what is needed – as argued by Steinberg & Kincheloe (2004), who encourage researchers to ask research questions that will make a difference in students' lives. Or, as Karl Marx said, "The point is not merely to understand the world, but to change it."

As you will see in the contributions from my colleagues elsewhere in this article, we have endeavoured to suggest some of the areas in which biology education research might make such a contribution – whether in the education of the next generation of biology teachers, determining how to teach biology, increasing the accessibility of biology to all learners, ensuring that biology is taught authentically, or in many other ways. Our shared hope is that biology education research can indeed make an increasingly valuable contribution to what needs to be done for the benefit of learners, for human society generally, and for the planet as a whole.

What teacher education programmes produce informed and effective biology teachers?

Edith R. Dempster[1]

dempstere@ukzn.ac.za

Among the many factors that influence effective learning of biology, few are as important as the nature of the biology teacher. Teachers are the intermediaries between the content and processes of biology and the students themselves; biology teachers must be both informed and effective. Therefore, a major research goal for the future is to define the optimal teacher education programme necessary to develop biology instructors who are informed and effective. The topic is conceptualized here in relation to preservice teacher education, although answers may apply equally to in-service teacher education.

A variety of research questions have been proposed with regard to the preparation of biology teachers. Such key questions are related to the following domains.

Determining the qualities of an effective biology teacher

The qualities of an effective biology teacher, once identified, must inform the structure of preservice teacher education programmes. Lederman & Lederman (2015) reviewed the history of attempts to determine the qualities of a good teacher and describe a transition from initially asking the opinions of students to engaging the opinions of experts later. However, opinions about the effectiveness of science teachers are context dependent. For example, students, peers, administrators, and parents all will have different perspectives on what makes a biology teacher good. An educational system that is strongly examination-oriented may judge an effective teacher to be one whose students perform very well in those examinations. By contrast, a schooling system that aims to produce critical thinkers will judge a successful biology teacher differently.

---

[1] University of KwaZulu – Natal, South Africa

Knowledge and skills of effective biology teachers

From the early 1900s to the 1930s, effective teachers were described by several general attributes, including good judgment, magnetism, considerateness, and leadership (Lederman & Lederman, 2015). Later, an effective teacher was considered to be one who developed critical-thinking skills and tolerance of a diversity of viewpoints and opinions (Lederman & Lederman, 2016). Matthews (2015) describes a good teacher as a person who knows his/her subject, is interested in children and teaching, can use technology effectively, and teaches engagingly.

One quality of effective biology teachers frequently identified is the breadth and depth of subject-matter knowledge. A Norwegian colleague, Peter van Marion, who has had many years of experience in teaching the didactics of biology, reports that his preservice students identified strong subject knowledge, commitment, and enthusiasm as the most important qualities of a good teacher. The examples given here support the view that there are many answers to questions related to a determination of the knowledge and skills of effective biology teachers.

Preservice biology teacher education

Lederman & Lederman (2015) point out that there is no single best way to educate future science teachers. Matthews (2015) concurs, referring to countries that require no or minimal preservice teacher education. Simultaneously, some developed countries require a minimum of a Master's degree in biology in order to teach at the secondary level. Contextual and political issues in different countries will influence decisions about the structure of teacher education.

It is useful to remind ourselves of the important work of Lee Shulman (1987), who proposed a framework for teacher education that identified three major components: subject matter knowledge (SMK), general pedagogical knowledge (PK), and pedagogical content knowledge (PCK). These three components were subsequently expanded to include knowledge of the curriculum, knowledge of students, and contextual knowledge.

SMK is defined as a deep understanding of the fundamental concepts of a subject, knowledge of the research methods of that discipline, and knowledge of the nature of science (Groβschedl et al., 2015). Studies have shown that SMK alone is insufficient for effective teaching, hence the necessity for PCK (Zeidler, 2002; Groβschedl et al., 2015). PCK is knowledge of what

makes subject matter comprehensible to students (Großschedl et al., 2015). PK includes the foundational education disciplines, such as philosophy, sociology, history, and psychology of education.

Zeidler (2002) argues for a model of preservice teacher education that integrates SMK, PCK, and PK. He states that philosophical incompatibility between science faculty and education faculty precludes the possibility of achieving integration if teacher education is co-located with science instruction. In other words, Zeidler appears to argue against a science-content-only focus in preservice science teacher education. Others disagree. For example, Großschedl et al. (2015) found a positive relationship between the SMK and PCK of preservice biology teachers, in a programme where SMK was developed in the science department. Their study did not address the question of whether good SMK and PCK scores translated into effective classroom teaching. Zeidler (2002) and Großschedl et al. (2015) agree that SMK, PK, and PCK must be included in teacher education programmes.

Considering different models of teacher preparation

As Matthews (2015) has pointed out, there is little agreement among and within countries about how best to educate science teachers. Developing countries with a shortage of qualified teachers may accept shorter teacher-education programmes with less depth in SMK than first-world countries. Some developed countries are reducing the requirement for a formal teacher education qualification following a science degree.

In the United States, students with degrees in biology are increasingly encouraged to skip traditional teacher preparation, take a few summer seminars, and move directly to the classroom. A similar trend is evident in the United Kingdom, where the government encourages graduates to move directly into schools, where they undergo an apprenticeship mode of training (Matthews, 2015). It remains to be seen whether this produces effective teachers or just more individuals who serve briefly in teaching roles.

With such diversity of preservice programmes in operation worldwide, there is ample scope for research on the relative effectiveness of different programmes for biology teacher preparation. Such research will provide evidence that informs decision-makers about the structure of preservice teacher education. Clearly, there is much work to be done in the field of research into effective and efficient biology teacher preparation.

A summary of potentially fruitful avenues for research in biology teacher education might include the following:

- Identification of criteria linked to evaluating teaching effectiveness.

- Identification of the necessary knowledge and skills related to teaching effectiveness within different contexts and from different stakeholders' perspectives. In turn, the answer to this question should inform the design of both preservice teacher education and in-service teacher development.

- Determination of the form each of these components (SMK, PCK, and PK) takes, and how much of each component makes an effective biology teacher.

- Evaluation of a programme of teacher education must include evaluation of the classroom effectiveness of the teachers it produces. Such research will provide useful evidence informing the curriculum for preservice teacher education.

Making biology education accessible to students is a daunting quest. Challenges in teaching and learning biology have been studied extensively in terms of the topics that students regard as difficult. Students' interest in biology in general, along with gender- and age-related differences in interest, and students' understanding of the structure of the discipline have also been investigated. In addition, underlying factors that may contribute to learning difficulties include the complicated levels of organization and the abstract concepts involved (Lazarowitz & Penso, 1992), the amount of content that students often find overwhelming and difficult (Çimer, 2012), and the requirement that students must switch thought levels from tangible to molecular to symbolic or mathematical (e.g., in genetics; Bahar et al., 1999). Further impediments in learning biology also include interest. The Relevance of Science Education (ROSE) Project (roseproject.no) has indicated a lack of interest in biology topics, particularly among boys in developed countries (Sjøberg & Schreiner, 2010), which suggests a gender and cultural gap in biology learning. Students' interest in biology also appears to diminish with age (Prokop et al., 2007), which may be linked to perceptions that biology is a difficult subject (Çimer, 2012).

All these difficulties suggest a need to rethink biology education. This reconsideration might be characterized in three perspectives: epistemological, metacognitive, and motivational. *Epistemologically*, if we accept biology learning as a personal construction rather than a transmission process and if we hope that students will "own" the knowledge acquired, more active student-centred learning approaches should be considered. At the same time, a more student-centred approach must account for potential conflicts with teachers' beliefs in a more traditional approach to teaching (Kinchin, 2001),

Making biology education accessible to students is a daunting quest. Challenges in teaching and learning biology have been studied extensively in terms of the topics that students regard as difficult. Students' interest in biology in general, along with gender- and age-related differences in interest, and students' understanding of the structure of the discipline have also been investigated. In addition, underlying factors that may contribute to learning difficulties include the complicated levels of organization and the abstract concepts involved (Lazarowitz & Penso, 1992), the amount of content that students often find overwhelming and difficult (Çimer, 2012), and the requirement that students must switch thought levels from tangible to molecular to symbolic or mathematical (e.g., in genetics; Bahar et al., 1999). Further impediments in learning biology also include interest. The Relevance of Science Education (ROSE) Project (roseproject.no) has indicated a lack of interest in biology topics, particularly among boys in developed countries (Sjøberg & Schreiner, 2010), which suggests a gender and cultural gap in biology learning. Students' interest in biology also appears to diminish with age (Prokop et al., 2007), which may be linked to perceptions that biology is a difficult subject (Çimer, 2012).

All these difficulties suggest a need to rethink biology education. This reconsideration might be characterized in three perspectives: epistemological, metacognitive, and motivational. *Epistemologically*, if we accept biology learning as a personal construction rather than a transmission process and if we hope that students will "own" the knowledge acquired, more active student-centred learning approaches should be considered. At the same time, a more student-centred approach must account for potential conflicts with teachers' beliefs in a more traditional approach to teaching (Kinchin, 2001),

[1] The Education University of Hong Kong

such as the conception that teaching biology should be based on lectures (Subramaniam, 2014). Teachers' conceptions about biology and biology teaching are often intertwined with barriers such as time and resource constraints, and these barriers need to be resolved if students are to be guided through the knowledge construction process (Kinchin, 2001).

From a *metacognitive perspective*, research evidence shows that learning could be improved by encouraging students to reflect on how they learn and how effective their learning strategies are (Thomas, 2012). This would be helpful for students who perceive biology as difficult and hence lose interest. However, it is a tall order for teachers who are preoccupied with cognition rather than metacognition. Therefore, future research should help teachers develop pedagogical skills for enhancing students' cognition and metacognition simultaneously through providing more opportunities for self-directed learning.

The *motivational perspective* is also worth considering. Here, different strategies have been suggested, including out-of-school experiences such as farming and experience in science and technology (Uitto et al., 2006) as well as practical work and fieldwork, particularly for enhancing the interest of males (Prokop et al., 2007). However, one must not lose sight of the intricate relationship between the cognitive and motivational aspects. Hence, research studies would have to focus on both cognition and motivation at the same time.

Finally, it would be difficult to envisage that any teaching or learning approach, no matter how promising it appears to be, can fit the needs of all students. Future research from the three perspectives discussed here should also address the differences across gender, age, ability levels, and culture, as informed by research findings to date.

Therefore, from the perspectives discussed here, several future research questions and approaches might be suggested to break down barriers:

- How can teachers be encouraged to reflect on their conceptions about teaching biology to bring them more in line with more active learning approaches based on constructivist learning?
- How can biology teacher educators model these teaching approaches effectively in methods courses?

- How can beginning teachers be helped to reconcile the apparent conflict between student-centred approaches and school curriculum contexts that are not conducive to the implementation of these approaches?
- More research is needed to establish the interaction between cognition and metacognition in the context of biology learning, and the teaching strategies for enhancing students' metacognition, such as awareness and evaluation of their own learning processes.
- How can biology teachers be encouraged to reflect on their beliefs about instruction, how these beliefs have come into existence, what impact they have on teaching processes, and how these processes could be improved?

The challenge of developing teachers' pedagogical knowledge with respect to language in biology education

Clas Olander[1]

clas.olander@mah.se

Investigations into the didactics of biology focus on investigating questions like *what*, *how*, and *why* biology is taught and learned. The aim is to encourage the examination of unanswered questions related to the development of teachers' pedagogical knowledge. Here, I will first suggest reasons why we should do this, next mention some important and under-researched questions, and end by discussing how this might be accomplished through professional development.

Students' learning depends on many background factors, such as the learners' social, cultural, and language background, but if the aim of schooling is to balance inequalities, research endeavours should investigate impacts from school itself, and that would logically centre on the teacher factor. The skills possessed by teachers are the main element. Hattie (2012) reminds us that enhancing teacher competence is rooted not in increasing subject matter knowledge, but instead in the ways that teachers introduce, organize, and scaffold learning experience related to biology content. In other words, teachers' pedagogical knowledge and competence should be targeted. This aligns nicely with how Shulman (1986, p. 13) describes the professional teacher as one who is "capable not only of practicing and understanding his or her craft, but of communicating the reasons for professional decisions and actions to others." This implies that the quality of pedagogical knowledge is best perceived as *enacted* competence in classroom practice.

The role of language in biology instruction

When it comes to *what* kind of knowledge should therefore be developed, I suggest alignment with the idea of "disciplinary literacy in biology", which

---

[1] Malmö University, Sweden

means investigations that focus more on "content-based language teaching" (Dalton-Puffer, 2011), since learning biology involves learning to master and appropriate the specific language of school biology. Language in biology classrooms is a particularly challenging issue and is characterized by multi-modality (e.g. representations, models, metaphors, formulas) and the use of specific words and semantic patterns (Lemke, 1990). According to Brown & Ryoo (2008), the combination of content and language components together enhance students' conceptual understanding.

Words in biology can be grouped in three categories: biology-exclusive terms, words found in biology and elsewhere but with different meanings, and general language. Biology-exclusivity implies words used only in the science of biology (i.e., *allopatric*, *genotype*, and *stroma*). Understanding these concepts is important, and misunderstanding can block the making of meaning, but both the students and the teachers know that they are important and invest both time and effort on them. Second, we have terms in biology that also have other connotations. Terms such as *adapt*, *cycle*, and *energy* can confuse learners because these terms have different meanings in everyday language. For example, students could arrive at school by "cycling," but in the biology classroom cycling is also associated with "life-cycle" or the cycling of matter. The third group of expressions includes general academic ones like *converted*, *proceeds*, and *originates*. These words too are important, and biology will make little sense if they are not understood. All three word types may certainly cause problems for all learners and particularly for second-language students (Gibbons, 2003). Teachers must understand how language influences learning and develop strategies to enhance students' successful appreciation of appropriate scientific language as a continuum between daily and scientific use (Schleppegrell, 2016).

Hattie (2012) implies that research into these areas might occur in a collegial learning environment, leading to the proposal of school-based "design research" agendas (Anderson & Shattuck, 2012) wherein teachers, in their own authentic practice, are engaged in iterative cycles of planning, enactment, and evaluation of teaching and learning (for examples in genetics and nature of science, see Olander & Holmqvist, 2013; Holmqvist & Olander, 2017). This is not unlike the lesson-study model that is commonly used in Japan and frequently practiced elsewhere.

Research questions related to this domain of language might be investigated through design research with teachers and include the following:

- What is the specific character of the language in biology classrooms?
- In what ways can content and language be integrated in order to scaffold student learning?
- In what ways do multimodal resources afford and hinder learning in biology?
- How can pedagogical knowledge be described as enacted competence in biology classrooms?
- In what ways can teachers scaffold student learning progressions in biology within a continuum between every day and scientific languages and contexts?

Three possible foundations for research in biology education: sociocultural contexts; consideration of knowledge, values, and practices; and Didactic Transposition Delay

Pierre Clément[1]

clement.grave@free.fr

In 2005, UNESCO proposed linking biology education with the promotion of fundamental values, such as human rights (i.e., gender equality and the struggle against racism, sexism, and homophobia) while encouraging environmental education for sustainable development, sex education, and health promotion. However, doing this in practice is a challenge. Increasing evidence (Carvalho et al., 2008; Castéra & Clément, 2014; Clément & Caravita, 2014; Clément, 2015) shows that teachers from diverse countries have vastly different views on these issues.

For instance, consider the responses to one statement about the possible biological justification of women's roles ("It is for biological reasons that women more often than men take care of housekeeping") in a comparison study of teachers from 33 countries (>11,000 teachers). Figure 1, which includes just data from biology teachers in nations that most agreed and most disagreed with the statement, shows that ~70% of biology teachers totally or rather agreed with the statement in Algeria and Georgia, while ~10% agreed or rather agreed in Italy, Spain, and Serbia, and just 3% in France.

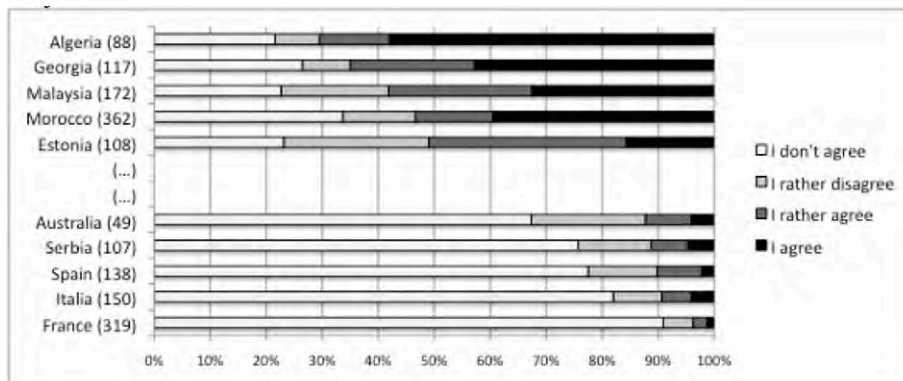[1] University Aix-Marseille, ENS de Lyon-IFE, France

Figure 1. Section of a larger data set featuring the responses of 5024 biology teachers (grouped by country in descending order of most agreement to most disagreement) to the statement "It is for biological reasons that women more often than men take care of housekeeping."

Similar differences in the opinions of biology teachers between nations were also found in many other questions. For instance, in response to a statement about the environment, "Our planet has unlimited natural resources," more than 95% of teachers disagreed in Germany and Finland, while only 20% held that same view in Morocco and Lebanon (Clément & Caravita, 2014). With respect to the origin of life, more than 80% of teachers in Algeria and Morocco chose a creationist response, compared with just 1% in Estonia, France, and Sweden. Many teachers (e.g., 60% in Malta) who believe in God revealed their evolutionist and creationist views when they indicated that the processes of evolution are controlled by God (Clément, 2015). Other important examples of national differences can be found in Clément and Castéra (2013) and Castéra and Clément (2014). These findings show that there are differences of conception about these issues among those charged with teaching science content. This knowledge certainly implies a future domain of research related to the sociocultural influences on what is taught in each country. To extend this kind of analysis, future research might use two key concepts: KVP (knowledge, values, and practices) and DTD (didactic transposition delay), which I briefly present here.

Using KVP as a foundation for future research in biology education

In considering the possible interactions KVP between scientific knowledge (K), values (V), and social practices (P) (Clément, 2006; 2013), we can start with an example such as the results of an analysis of the images of identical twins included in biology textbooks of 18 countries (Clément & Castéra, 2013). This revealed that in all cases, the twins had the same clothes and hairstyle. Science knowledge (K: "genotype → phenotype") is linked to implicit values (V: sociocultural features being determined by genes, innate ideas, fatalism) and to social practices (P: the way parents dress their children; the way publishers choose the images for their textbooks). This example illustrates interactions between the three poles K, V, and P.

More generally, any conceptions can be analysed as possible interactions between these three poles of KVP (Figure 2). For instance, since the studies of Broca at the end of the 19th century, some assumed that women were less intelligent than men because of their smaller brain size. Here, knowledge (K = brain size) was interacting with sexist values (V) and social practices (P). Of course, there is no correlation between intelligence and size of brain (new K), and gender equality is a citizenship value promoting more equality in social practices. Even if there are gender differences, such biological differences cannot justify gender inequality in action.

Figure 1 illustrates another sexist KVP interaction: in several countries, many biology teachers can justify by (outdated) biological reasons (K) that women should do more housekeeping than men, knowledge linked with local social practices (P) that are rooted in more or less sexist values (V).
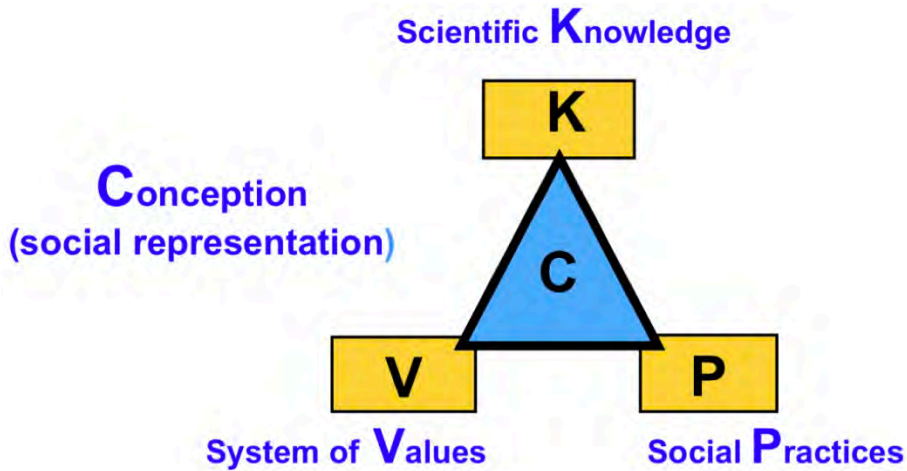
Figure 2. In the KVP model, any conception (C) can be analysed as a possible interaction between the three poles K (knowledge), V (values), and P (social practices) (Clément, 2006, 2013).

The conceptions of the different actors within the educational system can be analysed as possible KVP interactions at all the levels of the didactic transposition: learners' conceptions but also the conceptions of teachers; of authors of curricula, textbooks, and other documents; and even the conceptions of researchers who published the scientific references of the didactic transposition.

Considering DTD in biology education research

DTD is defined as a measure of the delay between the publication of a new scientific concept and its introduction in instruction (in syllabi, curricula, or textbooks; Quessada & Clément, 2007). Not surprisingly, scientific knowledge is updated frequently, sometimes in substantial ways, but often what is taught changes only slowly, with delays differing from one country to another. Therefore, the measure and interpretation of DTD could be an important approach in studying the sociocultural and economic influence on the content of taught biology across nations.

For instance, the issue of human origins has not yet been included in the textbooks of some countries (such as Algeria) and was recently suppressed in others (such as Lebanon). Important new biological concepts such as epigenetics, cerebral epigenesis, and transposons are not yet introduced in the secondary school curricula of several countries. There is also the challenge

that ideas that are taught may be partially outdated. For instance, in countries that still refer to the "genetic programme" and not yet "genetic information" (Clément & Castéra, 2013), the choice of the word *programme* may be ideological, suggesting that all our traits, competences, and performances are already written in our DNA. Consider another example of a biological fact: the number of human genes was estimated at 100,000 to 150,000 in the 1970s and '80s, yet today this estimate is ~23,000. However, this new reality is not yet reflected in all biology textbooks. Thus, DTD can be an interesting indicator of sociocultural influences on the nature of what is taught in biology classes in each country.

With respect to the foundation discussed here, future research questions might involve

- development of international comparisons of biology education, and historical approaches in each country, to identify the influences of different sociocultural and economical contexts (once differences in the biology curriculum or way of teaching are seen between one country and another, it is important to try to understand how and why biology instruction differs in these cases);

- use of KVP to analyse the conceptions of the main actors of the educational system related to each topic of biology, health, or environment – conceptions of students and of teachers as well as identifiable conceptions inside curricula, syllabi, textbooks, and other resources; and

- use of DTD to analyse the speed of changes within syllabi, within textbooks, or even within teachers' conceptions and to suggest possible interpretations of the differences seen.

Empowering students to cope with scientific innovations: lessons from genomics education

Dirk Jan Boerwinkel[1] and Arend Jan Waarlo[1]

D.J.Boerwinkel@uu.nl

Biological research has not only changed our views of life, disease, and behaviour, but has also generated applications in many areas vital to humans, including food production, medical diagnosis and therapy, and forensics. The positive outcomes are many, but with these have come important dilemmas. Consider, for example, whether we should encourage or avoid using genetically modified organisms, or whether we should use medication for children with behavioural problems. Socio-scientific issues such as these cannot be addressed solely through more biological research. Personal reflection and societal dialogue on these practices are needed to clarify the values and interests at stake and to explore possible scenarios and regulations.

One justification for biology education is to support citizenship, with the aim of empowering students for decision-making by bringing both the findings of biology research and related implications into the classroom. Both the risks and benefits of recent technologies and findings should be addressed, but also the so called "soft impacts" (Boerwinkel et al., 2014).

In 2002, the Dutch government started funding the Netherlands Genomics Initiative, channelling large funds for fundamental and applied genomics research while including humanities and social science research and societal dialogue activities. Our institute and the Cancer Genomics Centre collaborated in designing, implementing, and studying genomics education and communication. Our educational output consisted of mobile DNA labs (van Mil et al., 2010), teacher education workshops, teaching materials, and strategies for discussing ethical dilemmas. Our research, which we offer as a model in other science domains, consisted of addressing student problems in understanding the cellular and molecular mechanisms of disease and focused on molecular mechanistic reasoning (van Mil et al., 2016), examining teachers' challenges when discussing ethical dilemmas in genetics (van der Zande et al., 2012) and finding international consensus on which genetic knowledge is required by scientifically literate citizens (Boerwinkel et al., in preparation).

---

[1] Utrecht University, The Netherlands

Our involvement in research on cancer genomics informed our curriculum designs and suggested future work. We learned about how cancer genomics research reveals the ways in which cells are regulated, and how these new findings change our views on traditional and basic biological concepts such as gene, phenotype, and trait. Our meetings with genetic counsellors showed how these professionals deal with statistics and what questions their patients and their families have, and the choices their clients must make. All these experiences were immensely fruitful in developing our thinking about biology education within a contextualized approach. The involvement of humanities scholars and social scientists made us aware of political aspects such as the regulation of diagnostic testing of embryos for the presence of BRCA gene variants (Robertson, 2003). Analysing dilemmas such as those related to informing relatives about the possibility of carrying a high-risk gene variant taught us about the different conflicting moral principles.

Some recommendations for future research in biology education can be derived from this work. First, studying the personal and societal impacts of new scientific practices implies that we should analyse how the meanings of relevant biological concepts change (e.g., the concepts of gene and trait). Second, cooperation between biological researchers and with experts on related ethical, legal, and sociological issues is vital. Finally, the most generalizable conclusion we can offer is that this framework of involving a variety of experts and stakeholders should be considered, no matter what biology content is the focus of instruction. No one method and no one group of experts can be sufficient.

In conclusion, we offer the following biology education research questions related to the work discussed here:

- How are biological concepts used contextually within innovative scientific and professional practices, and what are the differences with traditional meanings of these concepts in biology education?

- In what kinds of decisions should students be prepared to participate as citizens, and what do they need for informed decision making and acting?

- What are the consequences of adding concepts and skills to, or removing them from, the biology core curriculum?

- What are effective instructional strategies that stimulate reflection and argumentation in biology classrooms, and how can we prepare biology teachers to implement them?

Grand challenges in biology education research: some conclusions

William F. McComas[1]

mccomas@uark.edu

There was no expectation among the authors that any article like this could provide all or even most of the recommendations that should be made about future directions in biology education research. However, the notion that those with expertise in biology education could suggest targeted research questions was intriguing to all concerned. This work has influenced my own thinking about research in biology education, so let me offer a few personal thoughts about potentially fruitful avenues for research. We could

- examine potentially promising – but small – studies reported in the literature and encourage researchers to engage in larger versions of those studies in wider contexts to permit inclusive and comprehensive conclusions;

- consider the most effective organizational plan for biology instruction (should the study of cells come first, or would students respond better to a "big picture" environmental approach?);

- add to recommended biological pedagogical content knowledge (B-PCK) by examining and reporting prior studies of students' alternative ideas and beliefs about biological phenomena, which can confound instruction;

- include in B-PCK the determination and dissemination of useful analogies and examples;

- examine how best to weave nature of science (NOS) ideas into the biology curriculum and teacher education plans; and

---

[1] University of Arkansas, USA

- most importantly, determine how to link the biology research community with practitioners and other stakeholders so that we can work to explore actual problems of interest and share research findings with those who can put recommendations into action.

No matter the specific perspectives each of us has offered, we all share the goal that larger groups of researchers across wider educational and social contexts should be engaged if there is any hope of gaining conclusive answers to any questions of interest. We must have a more embracing conversation between a more diverse and multinational group of biology education researchers in cooperation with other stakeholders such as teachers, textbook authors, scientists, and policymakers. Only then can we identify and attack truly significant questions and propose ways in which our answers can inform biology education in meaningful ways. However, for now, we are pleased to offer these suggestions as small steps on the long road toward the determination of the Grand Challenges in Biology Education. Once these big questions are identified and their solutions considered in the widest settings and contexts possible, we will be positioned to enhance biology teaching and learning in the most effective and generalizable ways possible.

References

Anderson, T., & Shattuck, J. (2012) Design-based research: A decade of progress in education research? *Educational Researcher, 41,* 16–25.

Bahar, M., Johnstone, A. H., & Hansell, M. H. (1999) Revisiting learning difficulties in biology. *Journal of Biological Education, 33,* 84–86.

Bayrhuber, H. (2016) Our double helix: ERIDOB in the face of the two strands of biology "didaktik." In T. Tal & A. Yarden (Eds.), *The Future of Biology Education Research: Proceedings of the 10th Conference of European Researchers in Didactics of Biology* (pp. 149–154). Technion, Haifa, Israel.

Boerwinkel, D. J., Swierstra, T., & Waarlo, A. J. (2014) Reframing and articulating socio-scientific classroom discourses on genetic testing from an STS perspective. *Science & Education, 23,* 485–507.

Boerwinkel, D. J., Yarden, A., & Waarlo, A. J. (in preparation) Reaching consensus on the definition of genetics literacy that is required by a 21st century citizen. *Science & Education.*

Brown, B. A., & Ryoo, K. (2008) Teaching science as a language: A "content-first" approach to science teaching. *Journal of Research in Science Teaching, 45,* 529–553.

Carvalho, G., Clément, P., Bogner, F., & Caravita, S. (2008) BIOHEAD-Citizen: Biology, Health and Environmental Education for Better Citizenship, Final Report. Brussels: FP6, Priority 7, Project no. CITC-CT-2004-506015.

Castéra, J., & Clément, P. (2014) Teachers' conceptions about the genetic determinism of human behaviour: A survey in 23 countries. *Science & Education, 23,* 417–443.

Çimer, A. (2012) What makes biology learning difficult and effective: Students' views. *Educational Research and Reviews, 7*(3), 61–71.

Clément, P. (2006) Didactic transposition and the KVP model: Conceptions as interactions between scientific knowledge, values and social practices. *Proceedings of ESERA Summer School 2006* (pp. 9–18). Braga, Portugal: IEC.

Clément, P. (2013) Values in science and in science education. In M. Abrougui et al., *Science & Technology Education for Development, Citizenship and Social Justice (IOSTE-14). Journal INEDP* (Tunisia) vol. 1, no. 1.

Clément, P. (2015) Creationism, science and religion: A survey of teachers' conceptions in 30 countries. *Procedia: Social and Behavioral Sciences, 167,* 279–287.

Clément, P., & Caravita, S. (2014) Education for Sustainable Development: An international survey on teachers' conceptions. In C. Bruguière, A. Tiberghien, & P. Clément (Eds.), *Topics and Trends in Current Science Education* (pp. 175–192). Dordrecht, Netherlands: Springer.

Clément, P., & Castéra, J. (2013) Multiple representations of human genetics in biology textbooks. In D. F. Treagust and C.-Y. Tsui (Eds.), *Multiple Representations in Biological Education* (pp. 147–164). Dordrecht, Netherlands: Springer Science+Business Media.

Dalton-Puffer, C. (2011) Content-and-language integrated learning: From practice to principles? *Annual Review of Applied Linguistics, 31,* 182–204.

Gericke, N., & Ottander, C. (2016) On the issue of "research in the didactics of biology": Definitions and demarcations. In T. Tal & A. Yarden (Eds.), *The Future of Biology Education Research: Proceedings of the 10th Conference of European Researchers in Didactics of Biology* (pp. 155–162). Technion, Haifa, Israel.

Gibbons, P. (2003) Mediating language learning: Teacher interactions with ESL students in a content based classroom. *Tesol Quarterly, 37,* 247–273.

Großschedl, J., Harms, U., Kleickmann, T., & Glowinski, I. (2015) Preservice biology teachers' professional knowledge: Structure and learning opportunities. *Journal of Science Teacher Education, 26,* 291–318.

Hattie, J. (2012) *Visible Learning for Teachers: Maximizing Impact on Learning.* New York: Routledge.

Hirst, P. H. (1965) Liberal education and the nature of knowledge. In R. E. D. Archambault (Ed.), *Philosophical Analysis and Education* (pp. 113–138). London: Routledge & Kegan Paul.

Holmqvist M.O. & Olander, C. (2017) Analysing teachers' operations when teaching: What constitutes scientific theories? *International Journal of Science Education, 39.* Doi: 10.1-080/09500693.2017.1310407.

Kinchin, I. M. (2001) If concept mapping is so helpful to learning biology, why aren't we all doing it? *International Journal of Science Education, 23,* 1257–1269.

Lazarowitz, R., & Penso, S. (1992) High school students' difficulties in learning biology concepts. *Journal of Biological Education, 26,* 215–223.

Lederman, N. G., & Lederman, J. S. (2015) The status of preservice science teacher education: A global perspective. *Journal of Science Teacher Education, 26,* 1–6.

Lederman, N. G., & Lederman, J. S. (2016) The functions of a teacher. *Journal of Science Teacher Education, 27,* 693–696.

Lemke, J. L. (1990) *Talking Science: Language, Learning, and Values.* Norwood, NJ: Ablex.

Matthews, M. R. (2015) *Science Teaching: The Contribution of History and Philosophy of Science, 20th Anniversary Revised and Expanded Edition.* Oxon, UK: Routledge.

Olander, C., & Holmqvist, M.O. (2013) Professional development through the use of learning study: Contributions to pedagogical content knowledge in biology. *Procedia: Social and Behavioral Sciences, 89,* 205–212.

Prokop, P., Prokop, M., & Tunnicliffe, S. D. (2007) Is biology boring? Student attitudes toward biology. *Journal of Biological Education, 42,* 36–39.

Quessada, M.-P., & Clément, P. (2007) An epistemological approach to French syllabi on human origins during the 19th and 20th centuries. *Science & Education, 16,* 991–1006.

Reiss, M. J. (2016) The future of biology education research. In T. Tal & A. Yarden (Eds.), *The Future of Biology Education Research: Proceedings of the 10th Conference of European Researchers in Didactics of Biology* (pp. 177–180). Technion, Haifa, Israel.

Reiss, M. J., & White, J. (2013) *An Aims-Based Curriculum: The Significance of Human Flourishing for Schools.* London: IOE Press.

Robertson, J. A. (2003) Extending preimplantation genetic diagnosis: The ethical debate. Ethical issues in new uses of preimplantation genetic diagnosis. *Human Reproduction,* 18*,* 465–471.

Schleppegrell, M. J. (2016) Content-based language teaching with functional grammar in the elementary school. *Language Teaching, 49,* 116–128.

Shulman, L. S. (1986) Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14.

Shulman, L. S. (1987) Knowledge and teaching. *Harvard Educational Review, 57,* 1–22.

Sjøberg, S., & Schreiner, C. (2010) The ROSE project: An overview and key findings. http://roseproject.no/network/countries/norway/eng/nor-Sjoberg-Schreiner-overview-2010.pdf.

Steinberg, S. R., & Kincheloe, J. (Eds.) (2004) *19 Urban Questions: Teaching in the City.* New York: Peter Lang.

Subramaniam, K. (2014) Student teachers' conceptions of teaching biology. *Journal of Biological Education, 48,* 91–97.

Thomas, G. P. (2012) Metacognition in science education: Past, present and future considerations. In B. J. Fraser, K. G. Tobin, & C. J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 131–144). Dordrecht, Netherlands: SpringerScience+Business Media.

Uitto, A., Juuti, K., Lavonen, J., & Meisalo, V. (2006) Students' interest in biology and their out-of-school experiences. *Journal of Biological Education, 40,* 124–129.

UNESCO (2005). United Nations Decade of Education for Sustainable Development (2005–2014) Guidelines and recommendations for reorienting teacher education to address sustainability. UNESCO Education Sector, Technical Paper no. 2.

van der Zande, P., Akkerman, S. F., Brekelmans, M., Waarlo, A. J. & Vermunt, J. D. (2012) Expertise for teaching biology situated in the context of genetic testing. *International Journal of Science Education, 34,* 1741–1767.

van Mil, M. H., Boerwinkel, D. J., Buizer-Voskamp, J. E., Speksnijder, A., & Waarlo, A. J. (2010) Genomics education in practice: Evaluation of a mobile lab design. *Biochemistry and Molecular Biology Education, 38,* 224–229.

van Mil, M. H., Postma, P. A., Boerwinkel, D. J., Klaassen, K., & Waarlo, A. J. (2016) Molecular mechanistic reasoning: Toward bridging the gap between the molecular and cellular levels in life science education. *Science Education, 100,* 517–585.

Zeidler, D. L. (2002) Dancing with maggots and saints: Visions for subject matter knowledge, pedagogical knowledge, and pedagogical content knowledge in science teacher education reform. *Journal of*

Science Teacher Education, 13, 27–42.

Zogza, V. (2016) "Biology didactics": A distinct domain of educational research. In T. Tal & A. Yarden (Eds.), *The Future of Biology Education Research: Proceedings of the 10th Conference of European Researchers in Didactics of Biology* (pp. 181–187). Technion, Haifa, Israel.

Challenges in Biology Education Research

This volume consists of 24 original papers related to biology education research. The papers were first presented at the 11th Conference of European Researchers in Didactics of Biology (ERIDOB) organized by the Academic Committee of ERIDOB and the Centre of Science, Mathematics and Engineering Education Research (SMEER) at Karlstad University in Sweden. The conference took place on 5-9 September 2016 with 165 participants representing 24 countries. There were 77 oral presentations, including four symposia, and 52 poster presentations. After the conference the presenters were asked to send in extended papers, which all then went through a rigorous peer review process and these 24 were selected for this volume. They are presented in section one.

The theme for the 11th ERIDOB conference was Challenges of Biology Education Research – the same as the title of this book of collected papers, and a panel debate around this issue was arranged with William McComas as organizer. Included in the debate were contributions from seven scholars in the field of biology education research representing seven countries from four continents. In section two of this volume, we include the peer-reviewed versions of these short papers as a joint article commenting on the current challenges, trajectories and opportunities for biology education.

We hope that this volume will find its way to biology educators as well as biology education researchers and make a useful contribution to the development of biology education in Europe and around the world.

eridob
european researchers
in didactics of biology