Deliverable number: D7.2.6

Deliverable title:

Evaluation of DynaLearn - first phase insights

Delivery date:	n/a
Submission date:	2011/02/27
Leading beneficiary:	WP7
Status:	Final
Dissemination level:	PU (public)
Authors:	D. Mioduser, R. Zuzovsky, P. Salles, A. Souza, A. Portella, E. Braga, L. Wilhelms, M. Pereira Resende, G. Morison, P. Assumpção da Costa e Silva, I. Gontijo de Sá, L. Pires, A. Simões R. Noble P. Borisova, Y. Uzunov D. Zurel, M. Leiba, H. Benayahu, R. Nachmias, J. Ram A. Zitek, M. Poppe, M. Stelzhammer

Project number:	231526
Project acronym:	DynaLearn
Project title:	DynaLearn - Engaging and informed tools for learning conceptual system knowledge
Starting date:	February 1st, 2009
Duration:	36 Months
Call identifier:	FP7-ICT-2007-3
Funding scheme:	Collaborative project (STREP)

Abstract

This deliverable conveys an integrative overview of the results of the evaluation activities conducted by WP7 partners in phase 1 of the evaluation of DynaLearn. It refers to 22 evaluation activities conducted during 2010 with 442 participants. In the deliverable, the overall evaluation framework is presented, along with the questions addressed in phase 1, the main insights gained and the main conclusions drawn in this phase, and the plans for evaluation phase 2.

Internal review

WP7 partners authors of deliverables D7.2.X, Bert Bredeweg

Acknowledgements

The authors would like to thank our partners for the insightful discussions on the different evaluation issues and the development of evaluation plans and instruments. We would also like to thank the participants in our evaluation activities, students and teachers in all participating countries.

Document History

Version	Modification(s)	Date	Author(s)
01	Draft for TAU group internal review	10.1.11	Mioduser, Zuzovsky, WP7 partners
02	Draft for WP internal review	31.1.11	Mioduser, Zuzovsky, WP7 partners
03	Complete version for internal review	21.2.11	Mioduser, Zuzovsky, WP7 partners
04	Final version	27.2.11	Mioduser, Zuzovsky, WP7 partners

Contents

1. Introduction

This document reports the results of the first phase of evaluation activities of DynaLearn by WP7 partners. The evaluation activities were designed with the aim to address the evaluation questions posed in the DOW, in correspondence with the functionalities and features afforded by the modelling environment -the prototype of DynaLearn software- in this first phase. The main questions were operationalized and decomposed into specific questions and integrated in an evaluation framework guiding the development of the activities, methods, procedures and instruments.

In the following section the evaluation framework is briefly described, as well as an overview of the first phase's evaluation activities. A brief account of the activities conducted is presented in chapter 2. A summary of the results of this phase and the main insights gained appears in Chapter 3 followed by their discussion in Chapter 4. Chapter 5 presents the foci and guidelines for the next evaluation phase.

1.1. Evaluation framework

The evaluation framework is built as the inquiry space for the formulation of the evaluation questions and the design of the evaluation activities aiming to answer to these questions. Our main goal for the evaluation is:

To assess the contribution of learning by conceptual modelling with DynaLearn on students' understanding of environmental systems.

In specific terms, we aim to assess the effect of Dynalearn's key features and the process of modelling on students':

- Conceptual understanding (CU) - their learning of content knowledge related to the behaviour of complex ecosystems.
- Scientific reasoning, Qualitative reasoning (QR), and System Thinking (SQS) - their acquisition of scientific reasoning skills and ability to cope with complexity, through QR approach and language.
- Motivation and attitudes (M/A) - towards learning science and learning by modelling.
- Self-directed learning (SDL) - supporting the growth of Independent learning skills and practices
- Learning (L) - general aspects of learning enhanced by DynaLearn integrated functionalities.

The above, as function of learning with an environment encompassing:

- Conceptual Modelling (CM) - in terms of DynaLearn's specific modelling language, modelling process and 6 modelling levels. - the Learning Spaces.
- Conversational virtual agents (VC) - as these act in various functions and roles while interacting with the learner.
- Semantic Technologies (ST) - individualization of learning via DynaLearn tools for ontology mapping, diagnostic procedures, and the semantic repository.

The evaluation inquiry space is thus depicted in Figure 1.1. The shaded cells in the Figure indicate the issues covered in the first phase evaluation activities. Most questions and data collected relate to the effect of conceptual modelling with DynaLearn on students conceptual understanding, acquisition and application of scientific skills and system thinking, and on motivational aspects Questions of types 1a, 1b and 1c). In the last stage of this evaluation phase, with the availability of features related to DynaLearn's semantic technologies, we also conducted initial activities about their effect on learning (Questions of types 3a and 3b).

In the DOW the overall set of questions to be addressed is presented, formulated in general terms (pp. 20-21). These questions, and the cells in the inquiry space into which they are inscribed, are:

- Does the diagrammatic approach (as organised in the DynaLearn setting) actually allow learners to address more complex problems? **[1a, 1b]**
- Does the meta-vocabulary from which a conceptual interpretation is built, provide learners a domain independent analytic instrument that enables them to construct more fine grained and thorough analyses of how systems work? **[3a, 3b]**
- Do the embodied conversational agents establish the 'involvement momentum' required for learners to actually benefit from the added value provided by the software for handling conceptual knowledge? Which agents work best? And why or why not? **[2a, 2b, 2c]**
- Do the instruments to individualise learning (ontology mapping, diagnostic procedures, and semantic repository) adequately steer learners in acquiring the target subject matter? **[3a]**
- Does the personal autonomy cause learners to be more motivated? **[1cd, 2cd, 3cd]**
- Do learners actually learn better when using the full set of DynaLearn results? **[general - L]**
- And are students more motivated to take on science curricula? **[1c, 2c, 3c]**

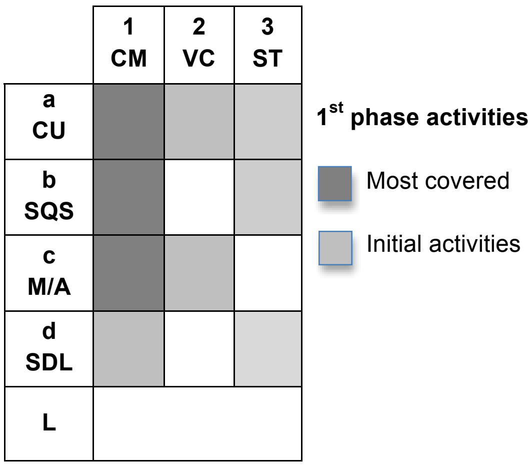

Figure 1.1: Evaluation inquiry space

For the 6th question relating to the most general issues and aims of the DynaLearn project, the answers will be: (a) gradually constructed upon the cumulative results for the various sub-questions along the project, and (b) addressed in the next phase's evaluation activities when DynaLearn as integrated learning environment will be available.

As mentioned above, the questions in the DOW were operationalized and specific sub-questions were generated for planning and conducting the evaluation activities. For example, while dealing with question 1 it was required to specify the aspects of the diagrammatic approach faced by the students, and the variables by which students' understandings of complex phenomena were assessed. The specific evaluation questions addressed, and variables and interpretative schemes for the results are summarized for each evaluation activity in the following chapters.

2. 1st phase evaluation activities - general

2.1. Introduction

Phase one WP7 evaluation activities were conducted mostly during 2010. A total of 22 evaluation activities were conducted, varying in evaluation goals, sample size, research design, and duration (details are provided in the methods section).

Overall, the first phase of evaluation was of **exploratory** nature. The activities were built to examine the very first "educational encounter" of DynaLearn with learners in real learning situations and in real learning settings (e.g., school, college, and university classes). The activities were conducted within the possibilities and constraints dictated by the conditions in which they were implemented, such as:

- They were conducted while components of the software were gradually completed and released, affecting decisions about the features on which they were focused at each stage.
- Instruments and analyses schemes had to be developed mostly from scratch, due to the novelty of the learning environment and its features. Previous research on Learning by Modelling with other tools and approaches could only serve as starting points or reference. As a result, much effort had to be put in instruments' development.
- Environmental Science is not defined as independent subject in most countries. At the school level its topics and concepts are embedded in the curricula of different subjects (e.g., Earth or Life Sciences, Social Sciences or History). At the University level there are Academic programs on the subject (e.g., as in TAU's Porter School of Environmental Science) however, in most cases the individual courses are focused on specific disciplinary themes. As a consequence, to consolidate appropriate settings for the evaluation activities and reach reasonable samples demanded the solution of administrative (and even ethical) procedures. In some activities even affected the possibility to control continuous participation and even impede abandonment.

In the following sections, various methodological aspects of the activities conducted are described.

2.2. Questions addressed

The general evaluation goal addressed in the 1st phase of evaluation activities was:

> **To assess whether DynaLearn, by its different technologies, in its different learning spaces, and in various pedagogical modalities - contributed to students' conceptual understanding of complex systems, to their system thinking and scientific inquiry skills, and their motivation for learning by modelling and self-directed learning.**

This general question was operationalized in the studies conducted by all WP7 partners, in categories of questions and sub-questions dealing with the following issues:

- Conceptual understanding of Environmental Science concepts and knowledge
  - Learning of scientific contents - changes in knowledge state along the learning cycle (working in the different Learning Spaces) and/or at its end
  - Ability to apply the gained knowledge in new contexts and situations
- System thinking and scientific reasoning skills
  - Growth of system thinking approach and skills

  - Coping with complexity
  - Causality - understanding causal relationships and processes and change along the Learning-by-Modelling activities
  - Growth of Scientific Reasoning skills
  - Support for the formulation of scientific arguments

- Motivation - towards Learning by Modelling (LbM), learning Science by means of LbM and learning with DynaLearn

- Modelling capability and behaviour
  - Understanding and learning the QM approach and language
  - Modelling capability in different LS
  - From naïve to expert modelling

- Contribution of semantic technologies and VC to students' learning
  - Grounding
    - Effect on the quality of student models
    - Facilitation of self-directed learning
    - Functionality issues
  - Ontology-based feedback
    - Effect on quality of student models
    - Extent of (correct) equivalence between terms in student and reference models
  - Teachable Agent
    - Contribution to better understanding of the systems behaviour
    - Contribution to students' construction of better models
    - If/Then questions effect on causal arguments

- Social interaction during the LbM process

- Software usability
  - Observed usability aspects
  - Observed problems and difficulties

Additional questions were explored with groups of teachers as target population, focusing on:

  - Teachers' perception of DynaLearn, and its role and potential contribution in teaching/learning processes
  - Teachers' motivation to work with qualitative models and to integrate DynaLearn in their teaching

In addition to these categories of questions, two additional pedagogical issues were addressed during the evaluation activities:

- Characterization and definition of a repertoire of pedagogical modalities for integrating DynaLearn in teaching and learning processes.

- The development of assessment instruments - first used as measurement and data-collection instruments during the evaluation then refined as pedagogical instruments for assessing students' learning.

In the next sections we summarize methodological aspects of the activities conducted. The main results of phase 1 concerning the above questions are presented in Chapter 3 of this document, and discussed in Chapter 4.

2.3. Methodological aspects

Most evaluation activities were primarily targeted at assessing the impact of LbM with DynaLearn on five targets: the acquisition of domain content knowledge (conceptual understanding); understanding of causal relations (causal understanding); understanding the complexity of ecological systems (complexity understanding); acquiring modelling capabilities and scientific reasoning skills, and strengthening the motivation to learn by modelling using DynaLearn.

These targets required the design of non-conventional evaluation instruments and the bottom-up construction of evaluation criteria to assess the various cognitive constructs addressed in the evaluation questions.

In that sense, evaluation studies conducted were of an exploratory nature. During the progressions from one evaluation activity to the following, the instruments were revised and evaluation criteria refined.

2.3.1. Instruments and evaluation criteria

The instruments used by the different partners included intuitive concept maps drawn by the learners themselves, or cognitive maps produced on the basis of primary documents (textual) that were written by the learners. These instruments were used to obtain initial, pre-modelling data on the levels of "conceptual understanding" and "complexity understanding" of the students. These data were compared to data obtained from the analyses of the final models' students constructed using DynaLearn. The comparison required to decide upon comparable evaluation criteria (as in TAU and UHULL).

Two analytical tools were used for the analysis of video data and textual content analysis. For textual content analysis, Atlas TI software was used. This software enables identification and counting of students quotations. The quotations were then used to detect causal understanding, conceptual understanding (fit with expert answers) and for building mental maps. The other tool used was University of Wisconsin's Transana software that allows researchers to transcribe and analyze video data. Short video clips are organized into meaningful categories. These clips are transformed into wave files that are the basis for time sensitive coding aspects. Atlas TI and Transana were only used by BOKU and enabled to identify differences between pre- and post-tests (Atlas TI) and the the mix of behaviours students exhibited while working in different learning spaces as assessed using video analysis (Transana).

For other analyses of qualitative data such as data obtained through the drawing of concept maps (Novak, Gowin & Johansen 1983; Ruiz-Primo & Shavelson, 1996) and the diagrams produced by DynaLearn we adopted cross-case display techniques (Miles & Huberman, 1994). The data gathered in this way was later on summarized for better description and interpretation (TAU comparison for concept maps and final models). Another way to assess models or concept maps was by comparing them to a reference (norm model) that was created (as was done by UHULL).

Another common instrument that was used in the different evaluation studies was a domain-specific test usually containing open-ended questions that detected conceptual understanding. These tests and their scoring guides were created by all partners.

Similarly, a motivation/attitudes questionnaire has been developed and used by all partners. An initial set of questions was developed and subsequently refined for its implementation in the different evaluation activities.

2.3.2. Participants

The sample size in most studies was small. This stemmed from the fact that the software was in its experimental phase and getting permission to test it in classrooms was problematic. At secondary school levels, conducting the evaluation activity interfered with the tight learning schedule and at the university level it interfered with other programs.

A summary of the samples in each study is shown in Table 2.1.

2.3.3. Design, data collection instruments and analyses procedures

The most common evaluation design used was the one Group Pre-test-Post-test Design. In cases where information was gathered during a sequence of lessons, the design was a "Single Group Interrupted Repeated Measure Design" – in these cases, the researcher recorded measures for the same group after each intervention.

In some cases when the effects of the software features were evaluated, the design of the study was a quasi-experimental one with non-equivalent experimental and control groups. Randomly assigning treatment to groups was hard to obtain in the context of most evaluations studies.

When the unit of analysis was individual students, the small sample size required non-parametric statistical procedures. However, when the analysis was done on a large number of qualitative narratives or video segments regular statistical procedures were applied.

A summary of the designs implemented by each WP7 partner (alphabetical order) and their reference code used in the remaining of this deliverable is shown in Table 2.1.

Table 2.1: Summary of methodological features of the evaluation activities in phase 1

Activity	Sample	Duration	Design	Code
BOKU1	2 (HS)	4 meetings / 24 hours	Case study / intensive data collection	CS
BOKU2	29 (GS)	4 hours	One group pre-post-measurement	1G/Pr-Po
BOKU3	2	Multiple sessions	Case Study w/repeated measures	CS
FUB1	60 (HS)	8 meetings / 12 hours	Experimental/Control	E/C
FUB2	60 (HS)	8 meetings / 12 hours	Experimental/Control	E/C
FUB3	27 (Teachers)	12 meetings / 60 hours	One group w/repeated measures	1G/Rm
FUB4	4 (HS)	6 meetings / 8 hours	One group (small) post-measurement	1G/Po
FUB5	10 (GS)	30 hours	One group post-measurement	1G/Po
FUB6	13 (Teachers)	4 meetings / 12 hours	One group pre-post-measurement	1G/Pr-Po
FUB7	21 (HS)	18 hours	Experimental/Control	E/C
FUB8	49 (HS)	18 hours	Experimental/Control	E/C

FUB9	35 (HS)	12 hours	Experimental/Control	E/C
FUB10	5 (HS)	3 meetings / 9 hours	One group (small) w/repeated measures	1G/Rm
IBER1	5 (GS)	6 meetings / 12 hours	One group (small) pre-post-measurement	1G/Pr-Po
IBER2	10 (S)	6 meetings / 12 hours	One group pre-post-measurement	1G/Pr-Po
TAU1	10 (GS)	7 meetings / 14 hours	One group w/repeated measures	1G/Rm
TAU2	23 (HS)	6 meetings / 24 hours	Experimental/Control	E/C
TAU3	14 (GS)	1 meetings / 3 hours	One group pre-post-measurement	1G/Pr-Po
UHULL1	4 (US)	4 meetings (12 hours)	Case Study w/repeated measures	CS
UHULL2	23 (US)	3 hours	E/C - incomplete due to technical failure	-
UHULL3	18 (GS)	3 hours	One group pre- post-measurement	1G/Pr-Po
UHULL4	18 (GS)	3 hours	Experimental/Control	E/C
Participants	442			

HS: High School; **US**: Undergraduate students; **GS**: Graduate students; **S**: mixed-degrees University students

3. Main insights from phase 1 of evaluation of DynaLearn

A succinct account of the main insights from phase 1 of the evaluation is presented in the following sections by the main categories of questions as detailed in section 2.2. Each section includes a set of main claims, followed by examples of observations from the whole set of evaluation activities. Examples and observations from each evaluation activity are preceded by indication of the study, e.g., BOKU2, and its design's reference code from Table 2.1, e.g., 1G/Pr-Po.

3.1. Conceptual understanding of Environmental Science contents

An essential focus for our examinations was the contribution of DynaLearn to students' conceptual understanding of ecosystems, and of concepts in Environmental Science. Evidences collected in the evaluation activities conducted by all WP7 partners are clearly indicative of this contribution.

Overall, two main conclusions can be drawn from the observations in the evaluation activities. The first is that the main gain at the conceptual understanding level relates to students' acquisition of a systemic perspective in learning about ecological systems. The second is that previous content-knowledge plays a crucial role in the model construction activity. Alternatively, the acquisition of the new knowledge and concepts required for complex modelling demands an extended learning process. This was facilitated by long-term interventions, in contrast with the limited effect of short-term interventions.

Data on the effectiveness of LbM showed a wide range of results, from almost no effect of LbM to a significant effect - on students learning. Examples of observations collected:

- [UHULL2] Analysis of written assignments did not reveal any significant differences in knowledge gains between control groups and experimental (DynaLearn) group.
- [FUB1,2,8,9 - E/C] Significant difference was observed in conceptual understanding gain in the experimental groups in [FUB1,2,8,9], however, post-test differences between experimental and control groups were not significant in [FUB8,9].
- [BOKU2 - 1G/Pr-Po] The use of DynaLearn led to significant increase in content knowledge and conceptual understanding. Topics explored with DynaLearn were graded highly in the final exam that also contained topics not covered through DynaLearn intervention.
- [TAU2 - E/C] Students' explanations to a set of twenty key concepts in ecology were improved from pre- to post-testing (large effect sizes in both groups). However, both groups obtained low average scores.
- [IBER2 - 1G/Pr-Po] All students performed very well on knowledge questions related to topics explored with DynaLearn compared to topics explored without DynaLearn.
- [FUB1, 2, E/C] Significant difference was observed in students' gain of conceptual understanding in the experimental group, and in comparison with the control group.

It is suggested [UHULL D7.2.2] that students learning by modelling will potentially move through three phases as they are introduced to the DynaLearn software from novice to apprentice and master phase. In these three phases students could be expected to:

- Fail to gain a greater conceptual understanding of a topic because they are focussed on understanding (or failing to understand) conceptual modelling (Novice).

- Gain a greater conceptual understanding of a topic by learning how to model directly using domain knowledge provided to them (Apprentice).
- Gain a greater conceptual understanding of a new topic by refining and consolidating information using a modelling/systems thinking approach (Master).

Therefore, it is probably unrealistic to expect a rapid response with students who are working at the novice level.

In addition, differences in the results could additionally be explained by motivational factors transmitted by the teacher, well known to have the potential to significantly influence learning success in general.

Focal observations indicating changes in conceptual understanding were obtained in the evaluation activities - these follow.

- [BOKU] Significant increase of the abstraction level of representing knowledge was observed.
- [TAU1,2 - 1G/Rm and E/C] Comparing student concept maps (previous to the modelling process) and final models yielded the following findings: Student final models contained fewer entities (only the relevant ones); quantities that were ignored in the concept maps were properly addressed in the models; differentiated representations of causal relationship were included in the models. The models were built around specific research questions and hypotheses that were tested through simulations and led in some cases to new questions and insights.
- [TAU2 - E/C] Students' capability to apply the gained knowledge in new contexts and situations was observed.

Generally, gains in conceptual understanding were assessed in most cases using tests of short answer type or longer written assignment. Although there were mostly indications that students who used DynaLearn exhibited better conceptual understanding (TAU, BOKU, IBER), there were also cases that did not show this gain (UHULL).

Two explanations are given to the limited gains in conceptual understanding: The first one relates to the fact that in many activities DynaLearn intervention was a short term one, not sufficient enough to alter or improve students' understanding of the topic. The other explanation relates to the modelling activity itself that shifts students' efforts from mastering the knowledge domain to mastering the modelling capabilities.

3.2. System thinking and scientific reasoning skills

3.2.1. On system thinking and coping with complexity

Students' system thinking and ability to represent a system's structural and behavioural features were clearly contributed by the work with DynaLearn. Along the learning processes, growth of skills and abilities was observed.

Overall, the students acquired rapid mastery of the skills and procedures required for constructing complex models with DynaLearn. As the modelling sessions advanced, their products reached high levels of complexity. At the end of the learning cycles, an increase in students' ability to represent a system's structural and behavioural features was observed. Analyses of models and explanations given to them along the course showed clear advances toward systemic view and understanding of the complexity and causal relationships (chain and loops) that stand behind the system behaviours.

Specific insights from the different evaluation activities follow.

Examples of observed performances indicative of growth in System Thinking are students' [TAU2 - E/C]:

- Progressive ability to define and refine the foci (the essential properties) of the model to be constructed
- Defining criteria for reducing the amount of model ingredients while preserving its meaningfulness
- Evolving ability hypotheses formation and testing
- Generation of new questions and inquiry processes beyond the original information used for generating the model
- Perceiving the value of the models constructed not only in terms of the specific phenomena modelled, but as paradigmatic examples of complex systems in other areas of study

Also observed was growth in students' skills such as the referred in the literature as comprising system thinking (e.g., Draper, 1993; Assaraf & Orion, 2005), e.g., the ability to identify components of a system and processes within a system; the ability to organize the system's components within a framework of relationships; understanding the hidden dimensions of a system; or thinking temporally - retrospection and prediction.

- [TAU2 - E/C] Pre- post-data showed changes in the experimental group's perceptions of the system, as reflected in their representations and models: increases in web-type representations, in adopting ecological organizing principles, and in the complexity of relationships among entities were observed. In contrast none of these changes were observed in the control group.
- [IBER1,2 - 1G/Pr-Po] At the end of the course, the students seemed to have acquired the skills to correctly distinguish different concepts in a scientific paper as corresponding to a particular model ingredient type in DynaLearn. That is, which are part of the structure of the system, and which aspects are dynamic (the quantities). Furthermore, they seemed to be able to choose the correct causal relationship between the quantities. This is, they can successfully identify the processes that are important in the system. This suggests that the students acquire a new set of analytical skills with which they can analyze topics.

The software definitely allows students to tackle complex problems. However, the development of system thinking implies an essential transition in students' learning: a transition in paradigmatic approach towards the inquiry of phenomena in the world. The vast majority of science teaching curricula and learning materials (particularly at the school level) aims to teach "classical science" and its methodology. Novel theoretical and methodological approaches in development in the sciences community for several decades, that view phenomena from the perspective of complexity and systems theories, are almost absent from existing curricula. This transition is not a trivial one, and demands the development of novel pedagogical approaches and essentially - it requires time, or long-term involvement in learning processes involving system thinking.

This also requires that the learning by conceptual modelling pedagogy and curriculum identifies suitable structures and learning activities through which students can move from novice to mastery in application of qualitative modelling and systems' thinking to learning.

3.2.2. Understanding causality

Understanding causal relationships within a system and between it and its environment is critical for understanding processes at different levels of the system's behaviour and for predicting or hypothesizing about its behaviour under changing circumstances. DynaLearn allows representing causal relationships in increasing level of complexity along its different Learning Spaces. In the evaluation activities conducted observations were collected about students' modelling work in all LS's. Insights gained follow.

About student ways to perceive causality, distinction can be drawn between two main layers. The first relate to students' ways of expressing "explicit" causal relationships, namely, those that are attached to the components as represented in the model (e.g., direct relationships between quantities or causal chains). The other is the "hidden" layer, related to their understanding of how the overall behaviour of the system results (emerges) from the causal configuration of its components.

- [TAU1,2 - 1G/Rm, E/C] For the first layer students represented causal relationships ranging from single/unidirectional relationships, through different configurations of one-to-many and many-to-one relationships, causal chains, to simple and complex feedback loops.

  For the second layer, our main observation was that along the modelling activities, students' perceptions (hence representations) of causal patterns and configurations became more complex. Even in the school students group, most of them succeeded in expressing long causal chains in their models (in contrast with a non-modellers control group). Also in most cases students' showed high ability to predict causal chains and loops in alternative scenarios for the systems modelled.

  For the higher layer - understanding of the overall behaviour of the system- we found ample evidence in students explanations. For example, undergraduate students' generation of inquiry questions beyond the questions in the scientific paper used by them as reference to build their models, is clear evidence not only of their understanding of the causal configurations among factors in the original experiment, but also of their ability to suggest and explore alternative causal configurations. One school student's comment is illustrative of the insights gained: "*The modelling activity taught me that some changes have long-term and far effects – If you touch one thing, everything can change*".

Evidence for students' understanding was noticeable in observations collected about their representations and verbal and written explanations along the learning cycles. These performances were also affected by the level of experience and previous knowledge of the participants - whether considering differences between younger/older students or novice/expert students.

- Significant increase in the use of causal relations was observed, especially graphical ones [BOKU2 - 1G/Pr-Po] and verbal expressions [BOKU1 - CS]. Data from 2 case studies extensively analyzed showed a significant increase in use of causal verbal expressions between pre- and post-tests, whereas wrong causal relations did not occur in post-tests.

- [BOKU2 - 1G/Pr-Po] An increased use of causal verbal expressions can be expected with younger/inexperienced students. The increased use of written causal words instead of using graphical means at younger students might be a result of their prior learning, while learning to use graphical and more abstracted representations to communicate knowledge might be a mid-to long-term goal of education.

- Increased match between students' models and expert models are indicative of the advantage of LbM for representing causal relationships, and of growth in causal understanding. [BOKU]

- [TAU2 - E/C; IBER2 - 1G/Pr-Po] At the end of the learning cycles, an increase in students' perceptions and representations of multiple-variables causal relationships, causal chains and feedback loops was observed. This is indicative of students' evolving understanding of the complexity of a system and of the type of causal configurations provoking its behaviour.

Difficulties in students' appropriate definition of causal relationships could be attributed to various aspects, some of them related to students' understandings (or missing- or miss-conceptions), and other to the affordances of DynaLearn in its different Learning Spaces.

- [UHULL3 - 1G/Pr-Po] In many cases students missed many of the key concepts required to build appropriate models. The majority of errors by students in causal dependencies were due to them not being implemented, either through missing the quantity out in the first place or through just missing out the correct dependency. Additional errors came from implementing either conceptually incorrect dependencies or dependencies that although correct at a certain level of granularity would be considered non-sequential in the teacher’s model.

- [UHULL3 - 1G/Pr-Po] Although a fair amount of the errors in causal implementation were due to the students' naivety in modelling, some issues remain with the implementation of causality in LS2. Although the positive and negative relationships available can be viewed as general causal relations at this level, they have very specific meanings (that of proportionality) when it comes to simulation and causal explanation of the model. Given this care must be taken when building and simulating basic causal models (LS2 and LS3) such as not to introduce inconsistencies in causal explanations. However, this also provides an opportunity for learning activities to introduce and explain causal differentiation to students.

- A set of evaluation activities at FUB included questions focusing on students understanding of causal relationships, measured as their explicit reference in written texts to causal Inferences, trivial and non-trivial conclusions. [FUB1,2,7,8,9]. In two studies with High School students [FUB1,2 - E/C] evident gains in understanding was observed (significant difference in results within the experimental group and compared with the control group). Interpretation of the whole set of results suggest that while DynaLearn appears to foster the growth of significant scientific reasoning skills, short time interventions are not enough to support their development and consolidation.

There is a major transition in the software between the notion of basic causality when students merely represent notions of positive and negative relations between quantities (although it should be noted that these relations are specifically proportionalities within the internal reasoning of the software) and causal differentiations where students implement notions of processes, direct influences and proportionalities between quantities. Therefore, the appropriate use of the different learning spaces and the activities used as transitions between these will have great importance for students developing reasoning skills based around qualitative process theory and systems thinking. Of particular importance will be the development of learning activities and use-cases for the implementation of models using basic causality in LS2 and LS3 and the transition from that to causal differentiation at LS4.

The development of reasoning skills around causality and qualitative process theory for novices working in the DynaLearn software will come mostly from the simulation and exploration of model behaviour (exploration modelling mode). As such the development of the “diagnosis” mode and the feedback given by interactive virtual characters, such as the teachable agent or other means, will have a great contribution to teaching the modeller about causal reasoning within the software. At the moment there are no easy avenues for a student to explore within the software to learn about the built in reasoning logic. This is something that the inbuilt help and diagnosis modes will address. Additionally, this is an area where well defined use-cases and appropriate modelling activities (advanced lesson plans and modelling curricula) need to be developed to optimise learning whilst using the software.

3.3. Motivation towards LbM and towards learning Science by LbM

Data on motivational aspects and students and teachers attitudes towards LbM and towards learning Environmental Science with DynaLearn were collected by all partners using questionnaires sharing common contents and structure.

In general students feedback concerning DynaLearn and the contribution of LbM was rated positively, indicating the general acceptance of the whole DL approach by students. Specific observations by WP7 partners follow. In general, students' answers to motivation and attitudes questions in all questionnaires administered showed similar trend, thus a succinct overview of the observations by each partner is presented here.

- [IBER] The motivation questionnaire results indicate that students find DynaLearn software easy to learn, and they think that the software can be applied more widely in other curricula. Al students indicated that they would use DynaLearn in other subjects. Students also indicate that they would like to have better learning materials so they can work more independently. These results suggest that students are motivated to take on more science curricula with the DynaLearn software.
- [UHULL] The answers given by the students indicate that they found it an interesting and challenging activity and some of them indicated that they found modelling a motivating activity (even without any of the added value technology such as the virtual character interaction). From the verbal feedback given by the PGCE students it was clear that many of them struggled to see how the approach might be applicable to them in their current teaching practice. However, those that experienced the teachable agent mode and virtual characters quickly identified how interactive modelling activities could be developed. However, many pointed out learning activities would need to be captivating, engaging and flexible so that students of different ability could be handled within the same activity.
- [TAU] Overall, the marks for motivational issues related to learning with DynaLearn were high. Looking at specific aspects, the higher scores were related to the students' perception of the software's motivational value for building ecological models (in the course and in the future as well), and to the contribution of the work with the software to their system thinking. At a lesser extent the modelling work was perceived by the students as contributing to their conceptual understanding and learning subject topics.
- [BOKU] The motivation questionnaires yielded in general only positive feedback to all questions asked. In detail, case-studies students very much liked the lesson and learning activities, indicating that modelling with the software led them to better understandings, and they highly agreed that modelling with the software could be also used in other learning topics. They less agreed that the software provides a very comfortable way of learning (modelling with DynaLearn was experienced as being challenging). Motivation data collected at [BOKU2] with more experienced students shows more heterogenic results. The highest agreement was documented for the applicability of the software to other learning topics. Furthermore they liked the lesson and learning activity supported by DynaLearn and they found it very interesting to work with DynaLearn. Also the importance of building models in different LS's was ranked as high. As the students were already well informed about the issue that was explored by DynaLearn, the activity did not much contribute to a new understanding of the system.
- [FUB] Students found it interesting and motivating to work with DynaLearn. In the case of deaf students they commented on how the use of qualitative models could help the deaf to learn concepts and to improve their writing skills. Students noted explicitly their perception of the contribution of DynaLearn to their learning.

- [FUB] Teachers recognized the high potential of the modelling activities for the development of a number of competences and skills, including the ability to make inferences, analogies and deductions while analysing the behaviour of a system; formulate hypotheses and predict results; analyse and compare possible solutions to the same problem. One of the teachers said: “To me, qualitative models refine the scientific method, allowing the student to formulate hypothesis and predict results in a consciously way.” Difficulties were attributed mainly to modelling at the more complex levels, and to particular functional aspects of the software (see details in sections 3.4 and 3.6).

One of the aims of DynaLearn as constructivist learning environment is to foster students' self-directed learning (SDL, Gibbons, 2002). It is clear that appropriate balance should be achieved between SDL and teacher-directed learning (TDL), aiming to create the necessary motivation and perception of self-confidence for the student to take control of her learning. Numerous observations of the gradual transition towards a more independent learning modality were collected by all partners. However, these situations should be formally and systematically devised as pedagogical process to be used in the future as models for the pedagogical implementation of DynaLearn.

Still in the dimension of the interaction between pedagogical approaches and students' motivation and attitudes towards learning, an important transition is required. For years the main teaching modality in which students are involved is the lecture-based or information-delivery modality. Learning about systems with DynaLearn demands a clearly different approach: learning by constructing models, and inquiry-based learning of phenomena represented in models. This transition implies once again a pedagogical model in which the "model constructor" responsible for her own learning stands at the centre of the scene (Papert, 1991). The constructionist idea stating that the students construct their inner world by constructing in the outer world demands a supporting pedagogy. Our preliminary observations of learning processes with DynaLearn serve as promising background for the development of these pedagogies.

3.4. Modelling capability and skills

Data on the gradual construction of modelling skills and capabilities were collected in many of the evaluation activities conducted. The observations are related to two main levels: (a) the conceptual level, focusing on students' gradual development of the QM approach and their ability to express phenomena in terms of qualitative models; and (b) the software level, related to students' work with the specific tools and features of DynaLearn, and the difficulties encountered. Naturally, the most detailed observations were obtained in the evaluation activities conducted as small-group case-studies, in which the learning process was intensively analyzed in detail (e.g., in BOKU or UHULL). However, substantial data were collected as well in activities in which repeated-measures were implemented, allowing comparison of stages in the modelling process.

At the conceptual level, it was evident that entering the realm of QM demanded a change in students' perceptions and approaches towards the inquiry of phenomena, contrasting with the perspective characterizing most science teaching and learning in educational systems. Evident as well is the fact that this transition demands time and involvement in recurring opportunities "to do the work" - to experience modelling tasks of varying types and complexity. Observations in the different activities unveiled different characteristics of this process, as in the following examples.

- Changes in perspective were reflected in the way students reformulated the phenomena under study for its representation in the modelling process. In [TAU2 - E/C] most students described the phenomenon to be modelled in the first session of the course in terms of a general question (e.g., "How do the wind and the waves affect the patella attachment to the rock?"). At the advanced modelling sessions they shifted to a language more focused on systemic and causal relationships (e.g., "The relationship between crabs, barnacles and the limpet"). The changes in the description of the aim of their modelling activity imply a change in perspective: from a focus on the local and specific aspects to be modelled to a more systemic view of the phenomena. The more generic descriptions imply also that the students were able to view the phenomena as particular instance of broader categories, in which multiple-variables causal relationships take place. Similar perceptions of modelling as novel approach were also observed in [FUB4 - 1G/Po].

- [BOKU1 - CS] While modelling, as students were not familiar with the ecological topic, they spent a lot of time in LS1 for picking up information (asking the teacher/other-student, looking in the internet, into additional materials). Topic-related questions decreased significantly with the progression to higher LS's, where the work focused on single processes. In addition, conversation on the modelling process and aspects of the activities (between students especially in LS2, and with the teacher especially in LS4) increased from LS1 to LS4.

- [BOKU1 - CS] In LS2 students worked very self-dependent on their models which implicated more time for thinking. LS2 also allowed students to easily translate their mental model into a dynamic model instead of having to invest too much effort in identifying relevant variables and relationships between them.

- The acquisition of QM language and different expression tools showed uneven patterns. The basic ideas and procedures behind the diagrammatic approach were rapidly grasped by most "naïve modellers" [e.g., in UHULL3 - 1G/Pr-Po; FUB8 - E/C]. In contrast, difficulties were observed in mastering more complex ideas in the modelling language, such as those related to "direct influences" and "proportionalities", causal dependencies that respectively represent processes, the initial cause of changes in the system, and the modelling element used to propagate the effects of processes [e.g., UHULL1 - CS; TAU1 - 1G/Rm; FUB7 - E/C]. It is evident that there is a need to plan appropriate pedagogical interventions to overcome these difficulties.

- [UHULL1, 3 - CS, 1G/Pr-Po] Additional insights were gained on the characteristics of naïve modellers. Their models showed great levels of variability in terms of complexity, mainly relating to the numbers of configurations and causal relations used rather than the number of entities and quantities used. Growth in confidence and understanding about how the modelling environment works resulted in better disposition to enter more complexity in the models. As well, crucial variables that appear to affect the modelling ability and pace are previous knowledge and individual cognitive styles (e.g., for solving problems). The ways these variables affect the modelling process should be systematically examined in future research.

- In general, activities lasting at least several weeks allowed observation of the gradual acquisition of modelling skills and methods. At the beginning of the learning cycle students had difficulties in discerning the set of necessary ingredients to construct the model, thus including a wide scope of components; along the modelling process, students' modelling became more focused and clear-cut concerning the distinction between necessary and unnecessary components [e.g., IBER1,2 - 1G/PrPo; TAU1 - 1G/Rm; UHULL3 - 1GPr-Po].

- Additionally, students adopted a range of strategies for coping with modelling tasks [TAU1,2 - 1G/Rm, E/C], as in the following examples: balancing between "trial-and-error" and "goal-oriented" modelling; "modelling-to-get-the expected-result" according to disciplinary knowledge - aiming to probe that meaningful (and known) results can be achieved with the software prior to using it as inquiry tool for new hypotheses; or differential attitudes towards "following-formal-instructions" while modelling [UHULL3 - 1G/PrPo].

Concerning the software level, we refer here to the contribution of its features to the modelling process - usability and technical issues will be briefly considered in section 3.6. One important set of observations relate to the contribution of the different earning paces to students' learning.

- [BOKU1 - CS] A detailed mapping of students modelling behaviour in each LS was depicted. In brief: Meaningful questions about modelling with the software increased along the work in the LS's, reaching levels of complexity that corresponded with the complexity of the features at each LS. As well, main mistakes (software related) made by the students in each LS were mapped. This mapping might well serve as basis for devising appropriate pedagogical solutions for supporting students' work.

- [FUB5 - 1G/Po; IBER2 - 1G/PrPo] LS4 was considered by many students as the learning space that most contributed to the understanding of the concepts represented by the models, probably because this is the first level at which causality differentiation can be applied and the consequences can be observed in action.

Overall, the evaluation activities in phase 1 allowed the composition of a valuable picture of students gradual mastery of modelling capabilities, of difficulties encountered, and of affordances of the software at each LS, as basis for the development of more systematic pedagogical solutions and more refined questions for the next evaluation phase.

3.5. Contribution of semantic technologies and VC to students' learning

As mentioned in section 1.1, questions related to the additional technologies integrated in DynaLearn, namely the Semantic Technologies and the Virtual Characters were only initially examined. These features were not available by the time most evaluation activities were conducted, and only in the last part of phase 1 a few preliminary activities could be implemented. Moreover, evaluation activities on these features were conducted only by some of WP7 partners: UHULL (teachable agent), TAU (grounding), BOKU (OBF). In consequence, the following are preliminary observations which will be surely continued in more systematic and comprehensive ways in the next evaluation phase.

The grounding feature

The "grounding" feature, although examined in its stage of development at the time of the evaluation, clearly supported students acquaintance with unknown concepts, and affected the quality of their models [TAU3 - 1G/PrPo]. For the unknown concepts, students' inquiry of their meaning using the grounding feature allowed their appropriate integration into the models. Before the grounding task about 40% of the students created quality models of the highest scores. After the grounding, all but one model got the highest score.

Teachable Agent

The Teachable Agent (TA) was implemented in [UHULL4 - E/C]. The main goals were to examine the contribution of students interactions with the TA to their understanding of the system's behaviour and to the quality of their models. In a more specific issue, the contribution of a particular question-format to students' causal reasoning was examined. Overall, no significant difference was observed between the experimental and control groups in the different aspects examined:

- The analysis of the students' written tests did not reveal any significant difference between the treatment and control group. A probably reason for this result is the previous knowledge which both groups already had on the topic under study.

- No difference in favour of being scaffolded by "if, then" type questions for building causal arguments was observed. It is most likely that a single exposure to this learning mode is not effective for supporting the development of the target skills.

- The models created by both control and experimental did not show significant differences. Students from both groups left out quantities and had errors in building their model. TA scaffolding did not improve the quality of the models. It can be concluded that the short term intervention is insufficient to affect in significant manner the model construction process.

As mentioned above, these short term activities can be seen only as preliminary exploration of the kinds of situations and opportunities in which the semantic technologies, and the VC, can be implemented as pedagogical means for supporting the modelling process.

3.6. Software usability

Issues related to students' interaction with the software such as the learning process of its features and functionalities or difficulties encountered, were addressed explicitly in some evaluation activities, and in others observations were collected as additional data alongside the main data collection.

It should be noted also that the evaluations were spread over several months, and during this time successive versions of the software were released in which previously lacking functionalities were completed or bugs were solved. Thus, focal (and temporary) technical issues are not mentioned here.

Overall, students in all activities conducted reached rapidly mastery of the procedures and methods required to work with the software. As well, they indicated that they see the potential of the modelling software for learning other subject areas and curricular topics besides Environmental Science contents. General observations from the partners' reports follow.

- [FUB] Students are very enthusiastic about the software, even saying that it is easy to model concepts in a model and indicating that they think the software can be widely applied different scientific disciplines.
- [IBER 1,2 - 1G/PrPo] Some students indicated that the software is difficult to use initially, but becomes easier to use in time. Other students indicated that they have no trouble using the software at all. Students indicated that they will use the DynaLearn software for the rest of their education.
- [TAU2 - E/C] High school students reached immediate mastery of the software's features and functions, even in shorter time than the observed in the undergraduate course.
- [BOKU] Students commented on what they liked most - "learning-by-doing".
- [UHULL1 - CS] Overall most students indicated that once they had identified the appropriate icons the software was fairly easy to use.

Difficulties encountered while working with the software, or requests for missing components, pertain to different categories. The first is related to language issues, particularly in activities done with the younger students:

- [IBER] the students missed the option to use Cyrillic in the DynaLearn software. As a workaround the students wrote Bulgarian using the Latin alphabet, which is suboptimal. Furthermore, they would also like the menus and vocabulary to be in Bulgarian so that it is easier to understand.
- Similar comments were found in the reports by [FUB[and [TAU]. In the last case, the problem becomes substantial concerning the interactions with the VC, due to the requirements of the Hebrew language (e.g., writing direction, text-to-speech production).

Other comments by the students relate to the need to fix existing functions or to ad additional functions, such as:

- [UHULL1 - CS] Two key usability issues identified by the students were: Firstly, most students asked for an “undo” button, as found in most other software packages; secondly, the implementation of the “add quantity” and “add configuration” functions should not automatically set you up to edit an existing model ingredient. On numerous occasions students found themselves inadvertently overwriting an ingredient they had previously added.
- FUB5 - 1G/Po] Students commented on the lack of commands that are found in other software packages commonly used in Brazil which allow, in general, the actions of "undo" and "redo" things and "copy" fragments of the models built.
- [BOKU1,2 - CS, 1G/PrPo] BOKU's report includes detailed lists of suggested modifications and improvements to the software as a result of the observed learning processes (see D7.2.5). These suggestions range from technical aspects (e.g., naming rules, display constraints, outputs arrangement), via the addition of buttons and functions (e.g., back button, view of models opened and changed since last session), to functional support of the modelling work (e.g., a sketch environment for causal models in LS6, importing information already defined in one LS into other LS's).

Finally, conclusions were drawn as to the situation of conducting the evaluations vis-à-vis the development of the software. In [UHULL D7.2.2] is noted that: "As the software is still a beta prototype with constant changes and updates of functionality it suffers from some level of instability. As such the evaluation exercise was also used to identify bugs and issues that caused the software to crash. Although using software that risks crashing during an evaluation exercise is undesirable in terms of collecting data and promoting the software to potential stakeholders (e.g. the group of trainee teachers) it is unavoidable in this development phase. What is of note is that some students seemed to suffer from more inexplicable crashes than others and it is likely that some naïve users attempt to do things in the software that a developer would never try to do (mostly through knowing how something is meant to be done). Therefore, testing of this sort with naive users is probably essentially for testing software functionality and stability".

3.7. Pedagogical outcomes

In addition to the evaluation results, two important outcomes were produced during the evaluations planning and implementation process. The first is a preliminary repertoire of pedagogical modalities and solutions for integrating DynaLearn in teaching and learning processes. The second is a collection of evaluation instruments which in turn might become meaningful assessment tools.

The vast majority of the evaluation activities were conducted in "real-life" learning settings (as opposed to "lab-like" settings), being these regular or specially developed courses for High School or University students. The immediate implication is that pedagogical procedures and processes had to be planned for running the activities. For this purpose, lesson plans including a range of pedagogical ingredients were devised, among others:

- Ways to introduce students to the approach and main concepts of QM.
- Support for the gradual activation and consolidation of system thinking skills.
- Pedagogical sequences supporting the gradual acquisition of modelling capabilities using DynaLearn.
- Strategies and methods for helping the students in their modelling processes.

- Repertoire of prototypical examples for introducing DynaLearn features at each LS.
- Alternative ways to trigger the modelling process (e.g., working on questions described in a scientific paper, or on a situated dilemma).
- Ad-hoc activities for working with specific features (e.g., grounding or TA).

A pedagogical mode of use in many of the evaluation activities was based on a collaborative configuration, in which students worked in dyads. The modeling process provided collaboration artifacts, serving as anchors for discussing, justifying and explaining the models and simulations. In addition, the idea of implementing structured collaboration was explored in a study conducted in the University of Amsterdam, using the Pair Modelling technique (by this technique, the partners are assigned roles -i.e., "modeller" and "reviewer"- changing them alternatively during a session). The study showed encouraging results based on assignments scores, observations and a questionnaire. Students' attitudes were neutral on the average, but the average score of the group that employed Pair Modeling was significantly higher than the average score of the control group that employed unstructured pair collaboration.

Another important pedagogical outcome of this evaluation phase is the set of instruments first used as measurement and data-collection instruments during the evaluation, then refined as pedagogical instruments for assessing students' learning. It should be noted that in many aspects the development of evaluation activities and instruments had to be done "from scratch", as previous research literature on many of the issues examined is neither abundant nor consistent. While the measurement of learning gains in terms of content knowledge using structured instruments is of common practice for decades, the measurement of the gradual construction of System Thinking, QM skills, and scientific skills required to cope with complexity still demands significant research effort.

During the first phase of evaluation activities we have developed a set of instruments and scoring guides for analysing students concepts maps, models at various levels of complexity, video-data, open texts and structured questionnaires that comprise a valuable methodological infrastructure. On this basis we will continue to refine the existing instruments and develop new ones for the next phase of evaluation. Eventually, many of the instruments and scoring guides will be reformulated as assessment tools and offered to teachers as pedagogical resources.

3.8. Concluding remarks

We conclude this chapter on the main insights gained from the 1st phase of evaluation of DynaLearn with a general overview on our main conclusions. Overall we have learned that:

- DynaLearn has proven to be of great potential for supporting: the growth of causal System Thinking; the acquisition of scientific reasoning skills; the ability to learn about complex ecosystems; the gradual construction of content knowledge; the gradual development of QM approach and skills.
- At the same time it became evident that changes in perceptions and approaches towards systemic phenomena and achievement of significant learning gains demand time: short term interventions or one-time activities are of modest impact on students' conceptual understanding, conceptual change and skills acquisition.
- Equally important, the acquisition of QM skills and mastery of DynaLearn's language and features is time demanding. Although students grasped in very short time the essentials of the work with the software and its basic features, mastery of more complex modelling tasks and of the higher Learning Spaces' features requires a continuous learning process and appropriate pedagogical support.

- Across most evaluation activities, students perceived as main learning gain the change in approach towards complex phenomena, and the acquisition of intellectual tools (QM skills and methods) for addressing these phenomena from a systems perspective. At the motivational level, most students' perceptions towards LbM with DynaLearn were generally positive.

- Numerous observations of the gradual transition towards a more independent learning modality were collected. This transition implies once again a pedagogical model in which the "model constructor" responsible for her own learning stands at the centre of the scene. The constructionist idea stating that the students construct their inner world by constructing in the outer world demands a supporting pedagogy. Our preliminary observations of learning processes with DynaLearn serve as promising background for the development of these pedagogies.

- Finally, difficulties were faced during the evaluation activities due to the fact that these were conducted while components of the software were under development. As the successive versions are being released, we expect these situations to diminish considerably during phase 2 evaluation activities.

4. Towards phase 2 of the evaluation of DynaLearn

For the next round of evaluation we will be able to design activities based on the implementation of a more complete configuration of the software, making possible to address the remaining questions in our evaluation framework (see Figure 1.1 and list of questions in page 9). More specifically, we expect to address in the next round key questions related to features of the semantic technologies and to the motivational and learning roles of the Virtual Characters. As well, we expect to be able to address the more general learning questions, and to deepen our questions on motivational aspects and self-directed learning. Finally, on the basis of the methodological approaches and instruments developed, we will be able to deepen our examination of students thinking about complexity.

The evaluation inquiry space for phase 2 is shown in Figure 4.1. Questions related to features of the technologies which were only initially examined in phase 1 will become the main foci of the evaluation activities (supplying evaluation data for questions 3 and 4 in the DOW). Questions on learners self-directed-learning will be explicitly addressed in some of the activities, however a more comprehensive elaboration on this will be based on the whole set of materials obtained during both evaluation cycles. Along similar lines, answers for the most general question on the overall contribution of DynaLearn to learning (question 6 in the DOW), will be composed upon all evaluation data collected during the project.

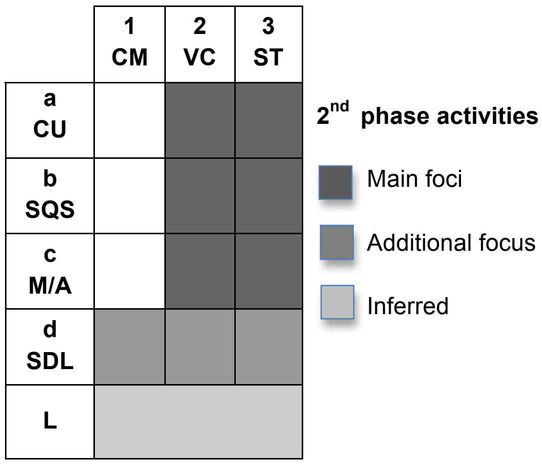

Figure 4.1: Evaluation inquiry space for phase 2

Detailed plans for phase 2 of evaluation are currently under preparation by all partners. About 20 evaluation activities are planned. These are expected to be conducted between mid March and July 2011. In brief, the overall plan for phase 2 is as follows (see also Figure 4.2):

- **Main questions**: Our main questions in this phase will relate to the contribution of OBF and help functions to learning; to the learning and motivational value of the interaction with the conversational agents; to a fine-grained characterization of what aspects of these features contribute best to learning; and to teachers perception of the educational value of DynaLearn's various features in teaching and learning of science.

- **Timeline**: most evaluation activities will be conducted between mid March and the end of July.

- **Target populations**: High-School students, University undergraduate and graduate students, teachers - ranging from small groups, focal groups, to classes (25-30 students)

- **DynaLearn features**: feedback and help functions; grounding; VC in various functions; CM environment.

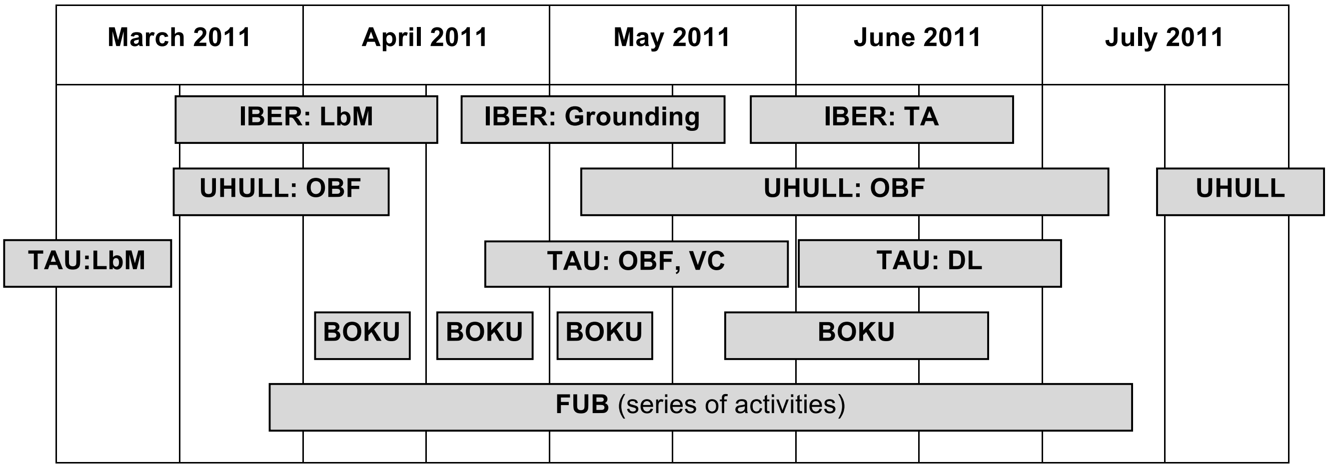

Figure 4.2: Preliminary timetable of DynaLearn phase 2 evaluation plans

References

Assaraf, O., & Orion, N. (2005). Development of System Thinking Skills in the Context of Earth System Education. *Journal of Research in Science Teaching, 42*(5), 518-560.

Bredeweg, B. *et al.* (2008a) DynaLearn – Engaging and informed tools for learning conceptual system knowledge. DynaLearn, EC FP7, STREP Project no. 231526, Description of Work (DOW).

Draper, F. (1993). A proposed sequence for developing systems thinking in a grades 4—12 curriculum. *System Dynamics Review, 9*(2), 207-214.

Gibbons, M. (2002). The self-directed learning handbook. San Francisco, CA: Jossey-Bass

Miles, M., & Huberman, A. (1994). *Qualitative data analysis (*2nd ed.). Thousand Oaks, CA: Sage.

Novak, J.D., Gowin, D.B., & Johansen, G.T. (1983). The use of concept mapping and knowledge vee mapping with junior high school science students. *Science Education, 67* (5), 625-645.

Papert, S. (1991). Situating constructionism. In I. Harel & S. Papert. (Ed.) *Constructionism*. Norwood. NJ: Ablex.

Ruiz-Primo, M.A., & Shavelson, R.J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching, 33* (6). 569-600.

Appendix A:

Summary of each evaluation activity by each WP7 partner

BOKU Evaluation Activities

<table>
<tr><th>Activity Title</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>Activity I</td><td>LS1 – LS2
LS4</td><td>Wind energy production</td><td>Upper secondary students in technical school</td><td>2</td><td>
<table>
<tr><td></td><td>1
CM</td><td>2
VC</td><td>3
ST</td></tr>
<tr><td>a
CU</td><td>2</td><td></td><td></td></tr>
<tr><td>b
SR/M</td><td>1</td><td></td><td></td></tr>
<tr><td>c
M/A</td><td>3</td><td></td><td></td></tr>
<tr><td>D
SRL</td><td></td><td></td><td></td></tr>
</table>
</td><td>
<table>
<tr><td>O_1</td><td>X_1</td><td>O_2</td><td>X_2</td><td>O_3</td><td>X_3</td><td>O_4</td></tr>
<tr><td>Pre-test</td><td>Introduction lecture</td><td>Concept-map</td><td>Lecture hands-on model-ing (LS2</td><td>Student's LS2 Model</td><td>Demon-stration of modelling (LS4 model)</td><td>Stu-dent LS4 mo-del post-test</td></tr>
</table>
</td></tr>
</table>

Evaluation Questions	Results & Conclusions
1. Modelling behaviour and social interaction (usability) How was the behaviour and social interaction during the modelling work?	Student behaviour differed significantly between the different LS - more picking information, more thinking time in LS_2 - increased conversation time from LS_1 - LS_2 - LS_4 especially with the teacher - increased number of questions related to modelling from LS_2 – LS_4
2. Causal understanding and content knowledge. How did the causal relations change during modelling (graphical and verbal expressions)?	The use of causal relations increased by 91% from pre- to post-test Wrong causal relation did not occur in the post-test Increased match between students' models and expert models and indicating acknowledging the advantage of LBM for representing causal relationship and indicating growth in causal understanding.
3. Motivation	Positive feedback to all questions asked. Better conceptual understanding; perceiving modelling with the software applicable for learning other topics; challenging; modelling LS4 models contributed mostly to their conceptual understanding

<table>
<tr><th>Activity Title II</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>BOKU1</td><td>LS2
LS4</td><td>Aquatic ecology and river manage-ment</td><td>Post graduate Students at BOKU University</td><td>21</td><td>
<table>
<tr><th></th><th>1
CM</th><th>2
VC</th><th>3
ST</th></tr>
<tr><td>A
CU</td><td>1</td><td></td><td></td></tr>
<tr><td>b
SR/M</td><td>2</td><td></td><td></td></tr>
<tr><td>c
M/A</td><td>3</td><td></td><td></td></tr>
<tr><td>D
SRL</td><td></td><td></td><td></td></tr>
</table>
</td><td>
<table>
<tr><th>O_1</th><th>X_1</th><th>O_2</th></tr>
<tr><td>Pretesting</td><td>Modelling activity</td><td>Post testing final content exam</td></tr>
</table>
</td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">1. Feedback on usability
2. Causal understanding
3. Motivation</td><td colspan="2">Decreased wrong causal relations
Increased graphical causal relations
Slight decrease in the use of verbal causal relations
The topic explored with DL was amongst the bet graded question in the final exam.
Highest agreement for applicability of the software for other learning topics as well.
Liking, interesting, the modelling activity needs to build models in different LS.
The activity did not contribute to better conceptual understanding, but students developed a better focus and used more causal relations.</td></tr>
</table>

<table>
<tr><th>Activity Title II</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>BOKU II</td><td>LS1-LS4</td><td>River continuum and river catchment</td><td>Master Students</td><td>2</td><td>
<table>
<tr><th></th><th>1 CM</th><th>2 VC</th><th>3 ST</th></tr>
<tr><td>a CU</td><td>1,2</td><td></td><td></td></tr>
<tr><td>b SR/ M</td><td></td><td></td><td></td></tr>
<tr><td>c M/A</td><td>3</td><td></td><td></td></tr>
<tr><td>d SRL</td><td></td><td></td><td></td></tr>
</table>
</td><td>Continuous documentation of modelling behaviour problems, attitudes and motivation.</td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">1. Usability issues
2. What problems learners encounter using the software
3. Motivational issues</td><td colspan="2">Comments on what students liked – learning by doing
Did not like – Lack of guiding manual for different LS
Comments on LS</td></tr>
</table>

<table>
<tr><th>Activity Title I</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>BOKU III</td><td>LS6</td><td>-</td><td>Master students (2)
Post-Doc (1)</td><td>3</td><td>
<table>
<tr><th></th><th>1 CM</th><th>2 VC</th><th>3 ST</th></tr>
<tr><td>a CU</td><td>1,2,3</td><td></td><td></td></tr>
<tr><td>b SR/M</td><td>4</td><td></td><td></td></tr>
<tr><td>c M/A</td><td>5</td><td></td><td></td></tr>
<tr><td>d SRL</td><td></td><td></td><td></td></tr>
</table>
</td><td></td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">1. Usability of OBF
Correctness (equivalence) between terms in student model and the reference model
2. Quality of the models supported by OBF</td><td colspan="2">Difficulties in using OBF window (list of problems).
Alternative tools (Text, Metadata on models) can compete with the OBF functionalities
Lack of well developed guidance for the whole process of model comparison.</td></tr>
</table>

IBER Evaluation Activities

<table>
<tr><th>Activity Title
I</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>Pilot Study</td><td>LS1

LS6</td><td>Water cycle
Nutrient cycle</td><td>First degree students</td><td>5 students</td><td>
<table>
<tr><th></th><th>1
CM</th><th>2
VC</th><th>3
ST</th></tr>
<tr><td>a
CU</td><td>2</td><td></td><td></td></tr>
<tr><td>b
SR/
M</td><td>1</td><td></td><td></td></tr>
<tr><td>c
M/A</td><td>3</td><td></td><td></td></tr>
<tr><td>d
SRL</td><td></td><td></td><td></td></tr>
</table>
</td><td>
<table>
<tr><th>O_1
Concept map
(LS1)</th><th>X
Course activities</th><th>O_2
Final model
(LS6)</th></tr>
</table>
</td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">1. Does the diagrammatic approach (as organized in the DynaLearn setting) actually allow learners to address more complex problems? (1b)

2. Does the meta-vocabulary from which a conceptual interpretation is built, enable them to construct more fine-grained analysis of how students work?(1a)
3. Are students more motivated to take on science curricula? (1c)</td><td colspan="2">At the end of the course, the students seemed to have acquired the skills to correctly distinguish different concepts in a scientific paper as corresponding to a particular model ingredient type in DynaLearn. That is which are part of the structure of the system, and which aspects are dynamic (the quantities). Furthermore, they seemed to be able to choose the correct causal relationship between the quantities. This is, they can successfully identify the processes that are important in the system.
-

Students are very enthusiastic about the software, even saying that it is easy to model concepts in a model and indicating that they think the software can be widely applied different scientific disciplines (and the education thereof).
Some students indicated that the software is difficult to use initially, but becomes easier to use in time. Other students indicated that they have no trouble using the software at al. Students indicated that they will use the DynaLearn software for the rest of their education.</td></tr>
</table>

Activity Title II	Learning Spaces	Content Domain	Target Population	Sample Size
	LS1, LS4	Intensive agriculture	First, second and third degree students	10

Evaluation Spaces

	1 CM	2 VC	3 ST
a CU	2		
b SR/M	1		
c M/A	3		
d SRL			

Evaluation Design & Instruments

O_1	X	O_2
Concept map	Course activities	Final model (LS2, LS4) Motivation measures Grades on final exam

Evaluation Questions	Results & Conclusions
1. Does the diagrammatic approach (as organized in the DynaLearn setting) actually allow learners to address more complex problems? (1b)	All students performed very well on questions related to the topic explored with DynaLearn compared to other topics explored without DynaLearn.
2. Does the meta-vocabulary from which a conceptual interpretation is built, enable them to construct more fine-grained analysis of how students work? (1a)	At the beginning of the modelling activity students tend to use all entities and relationships they identified. At the end of the activity, they committed themselves to only those essential for the purpose of modelling. Seventy percent of the students considered modelling at Learning space 4 to be very informative, but risky, since it required to differentiate between direct and indirect causal relationships. Fifty percent of the students preferred to work at LS3.
3. Are students more motivated to take on science curricula? (1c)	Students responded positively to the motivation questionnaire. Found the software interesting and learning by models useful. The software can easily be implemented for learning other topics.

HULL Evaluation Activities

<table>
<tr><th>Activity Title
I</th><th>Learning
Spaces</th><th>Content
Domain</th><th>Target
Population</th><th>Sample
Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>Pilot Study

Activity I</td><td>LS! - Concept Maps
LS2 – basic causal model
LS4 – causal differentiation model</td><td>Global Climate Change</td><td>University Biological sciences Students</td><td>4 - 7</td><td>
<table>
<tr><th></th><th>1
CM</th><th>2
VC</th><th>3
ST</th></tr>
<tr><td>a
CU</td><td>1</td><td></td><td></td></tr>
<tr><td>b
SR/M</td><td>3,2</td><td></td><td></td></tr>
<tr><td>c
M/A</td><td>4</td><td></td><td></td></tr>
<tr><td>d
SRL</td><td></td><td></td><td></td></tr>
</table>
</td><td>
<table>
<tr><th>X_1
Stimu-lus material</th><th>O_1
Con-cept map written assign-ment</th><th>X_2
Lectur demo hands on modelling norm CM</th><th>O_2
LS2 Model writ-ten assignment</th><th>X_3
Demo hands on modelling normal basic model</th><th>O_3
LS4 Model written assignment</th></tr>
</table>
</td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">1. Do students easily understand the DynaLearn modelling language and approach? (1b)
2. Do students build better models when working in the formalized dynamic learning spaces? (1b)
3. Does working with dynamic causal models enable students to write clearer scientifically correct causal arguments? (1b)
4. Do students find it easy to work with the DynaLearn software?(1c)</td><td colspan="2">1. Difficulties in mastering the modelling language specifically regarding "direct influences" and "proportionalities" [needed to plan intervention to overcome this difficulty}.
2. Difficulties in producing good representation using less formal concept maps. No clear result that supports the advantage of formalized models [modelling is viewed as a promising practice].
3. No indication of improvement in student writing of causal explanation [based on written assignment].
4. After acquaintance, the software was fairly easy to use [needed improvement to ease handling the software].</td></tr>
</table>

<table>
<tr><th>Activity Title II</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>Grounding of terms in models</td><td>LS2</td><td>Photosyn thesis</td><td>Post graduate Certificate in Education Students</td><td>23 - 37</td><td>
<table>
<tr><th></th><th>1 CM</th><th>2 VC</th><th>3 ST</th></tr>
<tr><td>a CU</td><td></td><td></td><td>5</td></tr>
<tr><td>b SR/M</td><td></td><td></td><td>1,2, 3,4</td></tr>
<tr><td>c M/A</td><td></td><td></td><td></td></tr>
<tr><td>d SRL</td><td></td><td></td><td></td></tr>
</table>
</td><td>
<table>
<tr><td colspan="2">Powerpoint Presentation</td></tr>
<tr><td colspan="2">Hands-on modelling using a normal model</td></tr>
<tr><td>X_1 Using grounding
X_2 Using free text comment boxes</td><td>O_1 models written assignments
O_2 models written assignments</td></tr>
</table>
</td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">
1. Does the current repository, DBpedia and grounding facility enable students to ground all terms? (3b)
2. How many terms would currently require the creation of anchor terms in the repository? (3b)
3. What type of terms require the generation of anchor terms in the repository? (3b)
4. How variable and correct are definitions provided by a range of students? (3a)
5. Does grounding (choosing from a predefined list) give students a better conceptual understanding of terms than requiring them to provide their own definition? (3a)
</td><td colspan="2">NO RESULTS – FAILURE OF THE GROUNDING FACILITIES</td></tr>
</table>

<table>
<tr><th>Activity Title IIIa</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>Open-ended modelling in learning space 2</td><td>LS2</td><td>Osmosis</td><td>Post-graduate Certificate In education Students</td><td>18</td><td><table><tr><th></th><th>1 CM</th><th>2 VC</th><th>3 ST</th></tr><tr><td>a CU</td><td></td><td></td><td></td></tr><tr><td>b SR/M</td><td>1,2, 3,4</td><td></td><td></td></tr><tr><td>c M/A</td><td></td><td></td><td></td></tr><tr><td>d SRL</td><td></td><td></td><td></td></tr></table></td><td><table><tr><td>0_1
Pretest
concepts
Causal argumentation</td><td>X
Powerpoint diagrams, lecture, hands-on modelling</td><td>O_2
Posttest
Models Evaluated according to criteria</td></tr></table></td></tr>
<tr><th colspan="4">Evaluation Questions</th><th colspan="3">Results & Conclusions</th></tr>
<tr><td colspan="4">1. How often do naïve modellers make errors in implementation of their model ingredients? (1b)
2. How easily do naïve modellers identify the important components that need to be included in a causal model to represent a scientifically accurate causal argument?(1b)
3. How variable is model complexity between students in terms of number of different ingredients?(1b)
4. Do naïve modellers work at appreciably different rates during a modelling session?(1b)</td><td colspan="3">1. The vast majority pick up the structural aspects of a model. Errors occurred due to including processes in the configuration.
2. Students did not identify or were not able to implement all key domain concepts [personified ontology based feedback is needed to direct students about relevant concepts].
3. The models developed showed high levels of complexity in terms of causal relations [the task is dependent on previous understanding of the domain].
4. Large variability among the students dependent on some cognitive problem solving styles [needed help and feedback support].</td></tr>
</table>

Activity Title IIIb	Learning Spaces	Content Domain	Target Population	Sample Size	Evaluation Spaces	Evaluation Design & Instruments
Influence of teachable agenda mode	LS2	Osmosis diffusion	Post-graduate Certificate in education Students	18 - 37	(see Evaluation Spaces table below)	(see Evaluation Design table below)

Evaluation Spaces:

	1 CH	2 VC	3 ST
a CU	2c 1c	1c	
b SR/M		3c 3c	
c M/A			
d SRL			

Evaluation Design & Instruments:

Experimental				
O_1 Pre test	X_1 Building a model using pre-defined ingre-dients inter-acting with VC	O_1 Model B1 criteria for analy-zing the model for causal depen-dence	X_2 Review the model following the student charac-ter ques-tions and his perfor-mance on a quiz	O_2 Model B2 criteria for ana-lyzing the model for causal depen-dence
Control				
O_0 Pre-test	X_1 Building Model using a list of model ing ingre dients	O_1 Model A_1	X_2 Running simu-lations	O_2 Model A_2 post-test

Evaluation Questions	
1. Does the TA mode give students a better understanding of the system behaviour? (2a) 2. Does the "If, Then" Question style of the TA contribute to the way students build causal arguments in the written text?(2b) 3. Does the TA mode contribute more to students building better, more correct, models than just through pre-defining the ingredients that should be included in the model?(2)	1. The analysis of the written test did not reveal any significant differences between the treatment and control group [maybe because initial knowledge both groups had on the topic. 2. No difference in favour of being scaffolded by "if, then" style questions used by the VC toward building logical causal arguments [single exposure to this learning mode is insufficient.] 3. The models created by both control groups (simulation) and experiment group (VC) do not show any differences students from both groups left out quantities and had errors I building the model. VC scaffolding did not improve the quality of the models.

TAU Evaluation Activities

<table>
<tr><th>Activity Title I</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td></td><td>LS6</td><td>Marine eco-systems</td><td>Under-graduate students of biology</td><td>10 students</td><td>
<table>
<tr><th></th><th>1 CM</th><th>2 VC</th><th>3 ST</th></tr>
<tr><td>a CU</td><td></td><td></td><td></td></tr>
<tr><td>b SR/M</td><td>1.2</td><td></td><td></td></tr>
<tr><td>c M/A</td><td>3</td><td></td><td></td></tr>
<tr><td>d SRL</td><td></td><td></td><td></td></tr>
</table>
</td><td>
<table>
<tr><th>O_1 Concept map</th><th>X Progression of modelling tasks</th><th>O_2 Student final models interview motivation questionnaire</th></tr>
</table>
</td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">Does student involvement in Qualitative Modelling (QM) of ecological systems improve:
1. Understanding of ecological systems?(1b)
2. Modelling capability and scientific reasoning skills? (1b)
3. Motivation to learn science by QM? (1c)

Additional evaluation objectives for this activity were:
1. To provide descriptive information on the integration of QM in regular course.
2. To serve as pilot study for the definition of the evaluation framework and the development of evaluation instruments.</td><td colspan="2">1.& 2. Comparing student concept maps and final models yielded the following findings. Student final models contained less entities (only the relevant ones), quantities that were ignored in the concept maps were properly addressed, differentiated representation of causal relationship. The models were built around specific research questions and hypotheses that were tested through simulations and led in some cases to new questions and insight.
3. Lower scores for increasing conceptual understanding overall. The students showed interest and motivation to learn with the software, also in the future. The main gain perceived was about QM contribution to system thinking and understanding the complexity of the ecosystem.
Less strong, was considered the contribution of QM to learning subject topics.</td></tr>
</table>

Activity Title II	Learning Spaces	Content Domain	Target Population	Sample Size
	LS3, LS4	Marine eco-system	Junior High students	21

Evaluation Spaces

	1 CM	2 VC	3 ST
a CU	1		
b SR/M	2.3		
c M/A	5		
d SRL	4		

Evaluation Design & Instruments

O_1	X_1	O_2	X_3,X_4	$X_{3, 4, 5}$
Pre-test exp.	Trip	Concept map	Modelling	Models and model docu-mentation
Pre-test control group	Trip	Concept map	Web-inquiry	Post-test Challenging question-naire and motivation question-naire

Evaluation Questions	Results & Conclusions
Does LbM with DynaLearn contribute to Junior High School students: 1. Conceptual understanding of a set of key concepts that represent the relevant content-domain (ecological systems)? (1A) 2. Ability to model a complex system and represent it at different levels of complexity using the qualitative reasoning approach embedded in DynaLearn. (1B) 3. Gradual acquisition of scientific and reasoning skills due to the modelling activity? (1B) 4. Capability to apply the knowledge and skills gained for approaching new ecological phenomena? 5. Motivation and attitudes toward learning by modelling?	Students explanations to a set of 20 key concepts in ecology were improved from pre-testing to post-testing (large effect sizes in both groups). However both groups obtained low average scores (to short intervention). The following pre-post changes occurred in the experimental group. Increase in net type representation and increase in adopting an ecological organizing principal In comparison none of the representation in the control group was net-like and less used the ecological organizing principal in these maps. Analyses of models and explanations given to them along the course showed clear advances toward systemic view and understanding of the complexity and causal relationships (chain and loops) that stand behind the system behaviours. Responses to the challenging questionnaire showed a tendency to adopt structure – behaviour – function type of explanations delineation of long chains of causal events. These characteristics were more profound in the experimental group.

<table>
<tr><th>Activity Title III</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td></td><td>LS2</td><td>Sea pollution</td><td>Undergraduate biology students</td><td>14</td><td>
<table>
<tr><th></th><th>1 CM</th><th>2 VC</th><th>3 ST</th></tr>
<tr><td>a CU</td><td></td><td></td><td></td></tr>
<tr><td>b SR/M</td><td></td><td></td><td>1</td></tr>
<tr><td>c M/A</td><td></td><td></td><td></td></tr>
<tr><td>d SRL</td><td></td><td></td><td>2</td></tr>
</table>
</td><td>
<table>
<tr><td>O_1
Blind model exercise and explanations</td><td>X_1
Grounding task</td><td>O_2
Find model and explanations</td></tr>
</table>
</td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">Does the grounding process contribute to:
1. The quality of the models constructed by the students (relevant entities and relationships, correct configuration causal chains).
2, Facilitate self-directed learning of concepts and causal relationships in the modelled system.
(Select appropriate definitions dealing with unknown terms construct knowledge related to an ecological phenomenon)</td><td colspan="2">Before the grounding task about 40% of the students created quality models of the highest scores. After the grounding all but one student's model got the highest score observational data indicated independent successful work in completing the grounding task.</td></tr>
</table>

FUB Evaluation Activities

Activity Title I	Learning Spaces	Content Domain	Target Population	Sample Size
Tree and shade Algal bloom	LS6	Global warming Algal bloom	Deaf students 15-29 years old	30

Evaluation Spaces

	1 CM	2 VC	3 ST
a CU	1		
b SR/M	2,3		
c M/A	4,5		
d SRL			

Evaluation Design & Instruments

O_1 Pre-test written essay	Experimental X_1-X_8 : Lecture Private Modelling	O_2 Post-test questionnaires
O_1 Pre-test Algal bloom written essay questionnaire Motivation and usability	Control X_1-X_2 Lecture ppt. Presentation	O_2 Post-test Algal bloom Written essay questionnaire

Evaluation Questions	Results & Conclusions
1. Do learners present significantly better scores in tests involving concepts on environmental science when using Garp3 conceptual modelling workbench and exploring qualitative models? 2. Does the meta-vocabulary used in qualitative models, from which a conceptual interpretation is built, provide learners a domain independent analytic instrument that enables them to understand more fine grained and thorough analyses of how systems work? 3. Do the students significantly improve their capacity of making inferences (causal reasoning) after using qualitative models as learning tools? Non-trivial conclusions? 4. What is the students' perception of the software, the modelling activities and the use of qualitative models in their learning process? 5. Are the students motivated to work with qualitative models and use Garp3 software?	1. The experimental group presented significant improvement in conceptual understand on the results of post-test after the use of qualitative models; 2. Although a significant difference between the pre and post-test results of the control group indicate a learning effect in the expositive lecture, comparison between the post-tests results in the two groups support the conclusion that the use of qualitative models has produced better results; 3. Significant differences were found between the experimental group's pre and post-tests, in total number of inferences and number of non-trivial conclusions, but no significant differences were found on the number of trivial conclusions; 4. Significant differences were found between post-tests of experimental and control groups, in number of inferences and number of non-trivial conclusions, but no significant differences were found on number of trivial conclusions. 5. The modelling language was accessible to the students and they considered excellent or good the clarity of the representation. About the use of QR language to describe the topics addressed by the models, 20% of the deaf students considered it only regular, probably by the level of complexity of the language. The understanding of the concepts worked out with the support of the models by 80% of the respondents. The students have a positive view on the contribution of qualitative models to their learning process, and would be keen in using these models in the classroom.

<table>
<tr><th>Activity Title II</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td></td><td>LS6</td><td>Trees and shade, Climate change, Economy and global warming</td><td>Public school hearing students age 15-18</td><td>60
(30 x 2)</td>
<td><table>
<tr><th></th><th>1
CM</th><th>2
VC</th><th>3
ST</th></tr>
<tr><td>a
CU</td><td></td><td></td><td></td></tr>
<tr><td>b
SR/M</td><td>2,3</td><td></td><td></td></tr>
<tr><td>c
M/A</td><td>4,5</td><td></td><td></td></tr>
<tr><td>d
SRL</td><td></td><td></td><td></td></tr>
</table></td>
<td><table>
<tr><td>O_1
Concept map</td><td>X
Experimental
X_1 –X_8

Course activities</td><td>O_2
Final model (LS2, LS4)
Motivation measures
Grades on final exam</td></tr>
<tr><td>O_1</td><td>Control
X_1 -X_1</td><td>O_2
Final model (LS2, LS4)
Motivation measures
Grades on final exam</td></tr>
</table></td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">1. Do learners present significantly better scores in tests involving concepts on environmental science when using Garp3 conceptual modelling workbench and exploring qualitative models?
2. Does the meta-vocabulary used in qualitative models, from which a conceptual interpretation is built, provide learners a domain independent analytic instrument that enables them to understand more fine grained and thorough analyses of how systems work?
3. Do the students significantly improve their capacity of making inferences (causal reasoning) after using qualitative models as learning tools?
4. What is the students' perception of the software, the modelling activities and the use of qualitative models in their learning process?
5. Are the students motivated to work with qualitative models and use Garp3 software?</td>
<td colspan="2">1. The experimental group presented significant improvement in conceptual understand on the results of post-test after the use of qualitative models;
2. A significant difference between the post-tests results in the two groups support the conclusion that the use of qualitative models has produced better results than the expositive lecture.
3. In the experimental group, the number of total inferences and of non-trivial conclusions in the post test is highly significantly bigger than in the pre-test, and the number of trivial conclusions was significantly smaller in the post test;
4. When the two groups were compared, the difference in the total number of inferences was not significant; the number of non-trivial conclusions was not significant, but it was exactly on the level of significance (5%); and the number of trivial conclusions was significantly higher in the control group.
5. The analysis of the questionnaires has shown that the students were very satisfied with the use of the didactic material to explain scientific concepts and approved the use of qualitative models and Garp3; the modelling language was accessible to them and they considered good the clarity of the representation. About the use of QR language to describe the topics addressed by the models and to express causality, the students considered it only regular, probably by the level of complexity of the language. The understanding of the concepts worked out with the support of the models for</td></tr>
</table>

	80% of the respondents. The students have a positive view on the contribution of qualitative models to their learning process, and all the students would be keen in using these models in the classroom.

<table>
<tr><th>Activity Title III</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>Teaching with models</td><td>LS1-LS4</td><td>Q.R. and Qualitative modelling</td><td>Public school teachers Integrated science at elementary level and specific domains at secondary level All teachers teach hearing and deaf students</td><td>27</td><td>
<table>
<tr><th></th><th>1
CM</th><th>2
VC</th><th>3
ST</th></tr>
<tr><td>a
CU</td><td>1</td><td></td><td></td></tr>
<tr><td>b
SR/M</td><td>1</td><td></td><td></td></tr>
<tr><td>c
M/A</td><td>1</td><td></td><td></td></tr>
<tr><td>d
SRL</td><td></td><td></td><td></td></tr>
</table>
</td><td>
<table>
<tr><th>X
O_1</th><th>X
O_2</th></tr>
<tr><td colspan="2">Teachers written report collected during the course and in schools; Questionnaires</td></tr>
</table>
</td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">1. Teachers' perception of the contribution of Q.M. to:
• Development of written scientific text
• Vocabulary
• Solving problems
• Producing explanations, formulate hypotheses predict results and produce arguments
• Teachers opinion Q.M.</td><td colspan="2">1. The teachers were able to explore the models presented and recognized the high potential for the development of a number of competences and skills, including the ability to make inferences, analogies and deductions while analysing the behaviour of a system; formulate hypotheses and predict results; analyse and compare possible solutions to the same problem. The teachers also reported substantial improvement on the students' writing skills. We can conclude the didactic material compiled in a DVD was well accepted for the deaf students and considered a bit slow for the hearing students. The teachers considered it a valuable tool for science teaching contributing for learning scientific concepts and the development of reasoning skills.</td></tr>
</table>

Activity Title IV	Learning Spaces	Content Domain	Target Population	Sample Size	Evaluation Spaces	Evaluation Design & Instruments
	LS2	Trees and shade Algae bloom, Erosion, Air pollution, Dengue Fever	Q.R. Experienced female deaf students	4	(see below)	(see below)

Evaluation Spaces

	1 CM	2 VC	3 ST
a CU			
b SR/M			
c M/A	1,2		
d SRL			

Evaluation Design & Instruments

X_1	O_2
Lecture text demonstrations modelling	Questionnaires

Evaluation Questions

This evaluation study aimed at answering the following questions:

1. What is the students' perception of the DynaLearn software, the modelling activities and the use of qualitative models in their learning process?
2. Are the students motivated to work with qualitative models and use DynaLearn software?

Results & Conclusions

- Most students said the conceptual modelling was a new approach, and it made them think about systems in a different way, giving new insight into phenomena and processes on ecology;
- All the students ticked "fully agree" when was said the process of modelling motivated them to learn more about the phenomena, recognising how the conceptual modelling could help them to learn about other topics;
- Most students understand the modelling approach and the goal of the modelling tasks;
- The students rated as "easy" and "hard in part" the conceptual modelling, find the conceptual mapping task and identification and describing of entities and quantities in system;
- Describing system structure in a basic casual model and work with simulations was easy to most of them.

Having had the experience with Garp3, the students quickly got the main features of the DynaLearn interface. Manipulating models was relatively easy for them, and their impression of the software was very positive. All the four students considered DynaLearn more accessible than Garp3, and comment on how the use of qualitative models could help the deaf to learn concepts and to improve their writing skills. One of the students suggested having all the words in the software written in Portuguese, in order to make it easier for understanding the models.

Activity Title V	Learning Spaces	Content Domain	Target Population	Sample Size
	Introduction to LS1-LS4; most activities in LS6	The modelling process	Graduate ecology students at the university level	10 (6 PhD, 3 MA, 1 Researcher PhD)

Evaluation Spaces

	1 CM	2 VC	3 ST
a CU	1		
b SR/M	2		
c M/A	3		
d SRL			

Evaluation Design & Instruments

X_1	O_1
Free modelling activity	Final models Motivation and software usability questionnaires

Evaluation Questions

1. What is the students' perception of the software, the modelling activities and the use of qualitative models in their learning process?
2. Do they foresee the use of qualitative models and DynaLearn as part of their research activities?

Results & Conclusions

1. Some students said the conceptual modelling was a new approach, and the most said it made them think about systems in a different way, giving new insight into phenomena and processes on ecology;
2. All the students ticked "strongly agree" when was said the process of modelling motivated them to learn more about the phenomena, recognising how the conceptual modelling could help them to learn about other topics;
3. Most students understand the modelling approach and the goal of the modelling tasks;
4. The students rated as "easy" and "difficult in part" the conceptual modelling, find the conceptual mapping task and identification and describing of entities and quantities in system;
5. Describing system structure in a basic casual model was easy and work with simulations was difficult in part to most of them.
6. The students rated as easy the qualitative modelling used in classes, and evaluated the experience of working with the DynaLearn as very interesting;
7. They said to build models in different specific Learning Space of DynaLearn is very important for understanding, and the LS4 was the LS who did contribute most with understanding of the concepts represented;
8. The students said completely agreed the modelling enabled them to better understand the complexity of the ecological and environment science, and the software could also be used in other learning topics.

Despite considering the conceptual modelling approach difficult in part, in general, most graduate students that attended the course considered that this approach improved their way of thinking about the environmental systems and their behaviour, as well as about their phenomena. Among them, 66% considered the

	conceptual modelling as a completely new approach, while 44% already knew this approach. However, they all felt that this technique made them think in a different way about the systems. All students found it interesting, in different levels, work with the DynaLearn software, and most of them considers it easy or very easy to operate it, although several students found it difficult. The LS4 was considered the learning space that most contributed to the understanding of the concepts represented by the models, probably because this is the first level at which causality differentiation can be applied and the consequences can be observed in action. The main negative points raised about the use of the DynaLearn software treated on the lack of commands that are found in other software commonly used in Brazil and which allow, in general, the actions of "undo" and "redo" things and "copy" fragments of the models built.

<table>
<tr><th>Activity Title VI</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>Teaching with models</td><td>LS1-LS4</td><td>Pollution, deforestation, bio-magnification, population dynamics</td><td>Secondary school teachers from different disciplines Some with modelling experience</td><td>23</td><td><table><tr><td></td><td>1
CM</td><td>2
VC</td><td>3
ST</td></tr><tr><td>a
CU</td><td></td><td></td><td></td></tr><tr><td>b
SR/M</td><td>2,3</td><td></td><td></td></tr><tr><td>c
M/A</td><td>4,5</td><td></td><td></td></tr><tr><td>d
SRL</td><td></td><td></td><td></td></tr></table></td><td><table><tr><td>O_1

Pre-activity questionnaire

Written pretest (causality)</td><td>X</td><td>O_2

Questionnaire on motivation and usability
Written post-test (causality)</td></tr></table></td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">Software and motivation
1. What are the teachers' perceptions of the software, the modelling activities and the use of qualitative models in their learning process?
2. Are the teachers motivated to work with qualitative models and use DynaLearn software?</td><td colspan="2">1. Most teachers rated as “hard” the modelling used to develop the educational activity, but evaluated the experience of working with DynaLearn software as “interesting”, helping them to understand the ecological problem after explore the topic in DynaLearn;
2. They said to build models in different specific Learning Space of DynaLearn is important for understanding, and the LS2 was the LS who did contribute most with understanding of the concepts represented;
3. The teachers agreed the modelling enabled them to better understand the complexity of the biodiversity loss process, and the software could also be used in other learning topics.
The test about causality detected an increase in the average score in post-test greater than in pre-test, which means that secondary school teachers can recognize more causal relationships after modelling activities. The result supports the conclusion that the use of the software can improve the ability of make and identify causal inferences. Despite they thought hard to use modelling in education activity, the use of DynaLearn was interesting manly to understand environmental issues. This is probably related to diagrammatic approach, the capacity to make predictions and to observe systems behaviour as a whole.</td></tr>
</table>

<table>
<tr><th>Activity Title VII</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td>Conservation biology</td><td>LS1-L4</td><td>Conservation issues
Hydrological erosion</td><td>Secondary school students 15-18 year olds</td><td>21</td><td>
<table>
<tr><th></th><th>1
CM</th><th>2
VC</th><th>3
ST</th></tr>
<tr><td>a
CU</td><td>1,2</td><td></td><td></td></tr>
<tr><td>b
SR/M</td><td>3,4,
5</td><td></td><td></td></tr>
<tr><td>c
M/A</td><td>4,5</td><td></td><td></td></tr>
<tr><td>d
SRL</td><td>6,7</td><td></td><td></td></tr>
</table>
</td><td>
<table>
<tr><th>O_1</th><th>X_1-X_4</th><th>O_2</th></tr>
<tr><td>Pretest causality and Conservation concept
Pre-motivation questionnaire</td><td></td><td>Post-test causality and Conservation concept
Post-motivation questionnaire</td></tr>
</table>
</td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">
1. Do learners present significantly better scores in concept tests when using DynaLearn’s conceptual modelling workbench?
2. Do the students present significantly better scores in concept tests while moving from LS1 – LS2? And from LS2 – LS4?
3. Does the diagrammatic approach (as organized in the DynaLearn settings) actually allow learners to address more complex problems?
4. Does the meta-vocabulary, from which a conceptual interpretation is built, provide learners a domain independent analytic instrument that enables them to construct more fine grained and thorough analyses of how systems work?
5. Do the students present significantly improve their capacity of making inferences (causal reasoning, analogies)?
6. Are the students motivated to work with qualitative models and use DynaLearn software?
7. What is the students’ perception of the software, the modelling activities and the use of qualitative models in their learning process?
</td><td colspan="2">
1. Using the result obtained in the paired t test with bootstrapping we observe a significant difference between pre and post-test in questions on Conservation Biology, and the mean scores increase from pre to post-test. This means that the approach used can collaborate with concept understand of dynamic systems.
2. They said to build models in different specific Learning Space of DynaLearn is very important for understanding, and the LS4 was the LS who did contribute most with understanding of the concepts represented.
3. The students agreed the modelling enabled them to better understand the complexity of the biodiversity loss process, agreed that the software provides a comfortable way of learning and it could also be used in other learning topics.
4. Some students said the conceptual modelling was a new approach, and it made them think about systems in a different way, giving new insight into phenomena of environmental science. Most students agreed that the process of modelling motivated them to learn more about the phenomena, recognizing how the conceptual modelling could help them to learn about other topics.
5. Some students rated as “difficult” and most as “neutral” the qualitative modelling used in classes, and evaluated the experience of working with the DynaLearn as very interesting. Most students understood the modelling approach and the goal of the modelling tasks.
6. The students rated as “easy” and “hard in part” the conceptual modelling, find the conceptual mapping task and identification and describing of entities and quantities in system.
7. Describing system structure in a basic casual model was hard in part and work with simulations was hard in part to most of them.
</td></tr>
</table>

Activity Title VIII	Learning Spaces	Content Domain	Target Population	Sample Size	Evaluation Spaces	Evaluation Design & Instruments
Conservation biology; metapopulation	LS1-LS4	Conservation issues Ecosystems Human Activity Communities Population	Secondary public school Students' aged 15-18	49	(see table below)	(see table below)

Evaluation Spaces:

	1 CM	2 VC	3 ST
a CU	1,2		
b SR/M	3		
c M/A	4,5		
d SRL			

Evaluation Design & Instruments:

O_1	X_1	O_2
Pretests Written essays Diagram-matic products	Experimental groups (modelling) 1. Trained – group A 2. Not trained – group B	Post-test Written essay Modelling products
O^1 Pretests Written essays Diagram-matic products	X_2 Control group C traditional	O_2 Post-test Written essay Modelling products

Evaluation Questions	Results & Conclusions
1. Do learners present significantly better scores in tests involving concepts on environmental science when using DynaLearn's conceptual modelling workbench to build and explore qualitative models? 2. Do the students present significantly better scores in concept tests while moving from LS1 – LS2? And from LS2 – LS4? 3. Do the students present significantly improve their capacity of making inferences (causal reasoning, analogies)? 4. Are the students motivated to work with qualitative models and use DynaLearn software? 5. What is the student's perception of the software, the modelling activities and the use of qualitative models in their learning process?	• The experimental group B presented significant improvement in conceptual understand on the results of post-test after the use of qualitative models; • The experimental groups put together A+B presented significant improvement in conceptual understand on the results of post-test after the use of qualitative models; • The experimental group A considering only students that had trained in DynaLearn presented significant improvement in conceptual understand on the results of post-test after the use of qualitative models and the time to familiarize with software. • The difference on posterior knowledge (post x post) on the concepts addressed during the course, the three groups, experimental A and B and control, was not significant, the significant difference is an important condition to assess the experimental treatment effects on the students' behaviour. • All groups presented significant improvement in conceptual understand of meta-populations considering the questions 14 and 15 in population biology test on the results of post-test after the use of qualitative models; • The difference on previous (pre x pre) and posterior (post x post) knowledge

	on the concepts addressed during the course, the three groups, experimental A and B and control, considering the questions 14 and 15 in population biology test, was not significant, an important condition to assess the experimental treatment effects on the students' behaviour; • All groups presented not significant improvement in conceptual understand of meta-populations on the results of post-test after the use of qualitative models; • The difference on posterior knowledge (post x post) on the concepts addressed during the course, the three groups, experimental A and B and control, was not significant, the significant difference is an important condition to assess the experimental treatment effects on the students' behaviour. • They said to build models in different specific Learning Space of DynaLearn is important for understanding, and the LS2 (group A) and LS1 (group B) was the LS who did contribute most with understanding of the concepts represented; • Most students rated as "easy" the qualitative modelling used in classes, and evaluated the experience of working with the DynaLearn as interesting; • The students agreed the modelling enabled them to better understand the complexity of the biodiversity loss process, agreed the use of the software provides a very comfortable way of learning and the software could also be used in other learning topics; • Some students said the conceptual modelling was a new approach, and the most said it made them think about systems in a different way, giving new insight into phenomena of environmental science; • Most students ticked "agree" when was said the process of modelling motivated them to learn more about the phenomena, recognising how the conceptual modelling could help them to learn about other topics; • Most students understood the modelling approach and the goal of the modelling tasks; • The students rated as "easy/hard in part" the conceptual modelling, as well the conceptual mapping task and identification and describing of entities and quantities in system; • Describing system structure in a basic casual model and work with simulations was easy to most of them.

Activity Title IX	Learning Spaces	Content Domain	Target Population	Sample Size	Evaluation Spaces	Evaluation Design & Instruments
	LS1-LS4	Conservation biodiversity Deforestation Pollution Meta-populations	Secondary public School students	15-control 5-experi-mental 15-18 year olds	(see Evaluation Spaces table below)	(see Evaluation Design table below)

Evaluation Spaces:

	1 CM	2 VC	3 ST
a CU	1,2 4,5		
b SR/M	3		
c M/A	6,7		
d SRL			

Evaluation Design & Instruments:

O_1	X_1	O_2
Pretests Written essays Diagram-matic products	Experimental group (modelling) 1. Trained – group A 2. Not trained – Group B	Post-test Written essay Modelling products
O^1 Pretests Written essays Diagram-matic products	X_2 Control group C traditional	O_2 Post-test Written essay Modelling products

Evaluation Questions	Results & Conclusions
1. Do learners present significantly better scores in concept tests when using DynaLearn's conceptual modelling workbench? 2. Do the students present significantly better scores in concept tests while moving from LS1 – LS2? And from LS2 – LS4? 3. Do the students present significantly improve their capacity of making inferences (causal reasoning, analogies)? 4. Do learners present significantly better scores in tests involving concepts on environmental science when using DynaLearn's conceptual modelling workbench and building and exploring qualitative models? 5. Do the students present significantly better scores in concept tests while moving from LS2 to LS4? 6. Are the students motivated to work with qualitative models and use DynaLearn software? 7. What is the students' perception of the software, the modelling activities and the use of qualitative models in their learning?	<u>In regard to the concept of population</u>: • The experimental (group A) and control groups presented significant improvement in conceptual understand on the results of post-test; • The experimental and control groups showed no significant difference between pre-tests which means that the groups were homogeneous for the subject of the workshop; • The difference in the results of post-tests between groups was not significant and the modelling approach didn't produce better result in that conditions. <u>In regard to the concept of meta-population</u>: • Only experimental group (A) presented significant improvement in conceptual understand on the results of post-test; • The experimental and control groups showed no significant difference between pre-tests which means that the groups were homogeneous for the subject of the workshop; • The difference in the results of post-tests between groups was not significant

	and the modelling approach didn't produce better result considering the design of the experiment. • Most students rated as "easy" the qualitative modelling used in classes, and evaluated the experience of working with the DynaLearn as very interesting; • They said to build models in different specific Learning Space of DynaLearn is very important for understanding, and both LS3 and LS4 was the LS who did contribute most with understanding of the concepts represented: • Most students fully agreed the modelling enabled them to better understand the complexity of the biodiversity loss process, and the use of it provides a comfortable way of learning and could also be used in other learning topics; • The students said the conceptual modelling was a new approach, and the most said it made them think about systems in a different way, giving new insight into phenomena of environmental science; • Most students ticked "totally agree" when was said the process of modelling motivated them to learn more about the phenomena, recognising how the conceptual modelling could help them to learn about other topics; • Most students understood the modelling approach and the goal of the modelling tasks; • The students rated as "very easy" the conceptual modelling, find the conceptual mapping task and identification and describing of entities and quantities in system; • Describing system structure in a basic casual model was very easy or easy/hard in part and work with simulations was very easy to most of them. • The results obtained in this evaluation were similar to the previous and in those conditions, the learning by modelling approach had similar effect as the expository classes. There was an improvement between pre and post-test in both groups, but the improvement in experimental group wasn't greater than the improvement in control group as we expect. • Considering only the metapopulation subject was obtained a little different result: the experimental group had a better result in post-test comparing with pre-test, but the same wasn't observed for control group. Nevertheless, no difference was detected in post-tests between experimental and control group.. • In general, the students had a very good opinion about the course and the modelling activities developed as a motivating task and as stimulate extra class activity. Most of the students looking for new ways, new tools to learning more and better and the activity were a very good experience in this point.

<table>
<tr><th>Activity Title
X</th><th>Learning Spaces</th><th>Content Domain</th><th>Target Population</th><th>Sample Size</th><th>Evaluation Spaces</th><th>Evaluation Design & Instruments</th></tr>
<tr><td></td><td>LS1-LS4</td><td>Global warming</td><td>Secondary school students experienced in DynaLearn</td><td>5-8 students</td><td>
<table>
<tr><th></th><th>1
CM</th><th>2
VC</th><th>3
ST</th></tr>
<tr><td>a
CU</td><td>1</td><td></td><td></td></tr>
<tr><td>b
SR/M</td><td>2</td><td></td><td></td></tr>
<tr><td>c
M/A</td><td>3,4</td><td></td><td></td></tr>
<tr><td>d
SRL</td><td></td><td></td><td></td></tr>
</table>
</td><td>
<table>
<tr><th>X_1 - X_3</th><th>O_1</th></tr>
<tr><td>Modelling activities</td><td>Student modelling products</td></tr>
</table>
</td></tr>
<tr><th colspan="5">Evaluation Questions</th><th colspan="2">Results & Conclusions</th></tr>
<tr><td colspan="5">1. Does the meta-vocabulary, from which a conceptual interpretation is built, provide learners a domain independent analytic instrument that enables them to construct more fine grained and thorough analyses of how systems work?
2. Do the students present significantly improve their capacity of making inferences (causal reasoning, analogies)?
3. Are the students motivated to work with qualitative models and use DynaLearn software?
4. What is the students' perception of the software, the modelling activities and the use of qualitative models in their learning process?</td><td colspan="2">Not yet available</td></tr>
</table>

e-mail: Info@DynaLearn.eu
website: www.DynaLearn.eu