Project no.	**004074**
Project acronym:	**NATURNET-REDIME**
Project title:	**New Education and Decision Support Model for Active Behaviour in Sustainable Development Based on Innovative Web Services and Qualitative Reasoning**
Instrument:	**SPECIFIC TARGETED RESEARCH PROJECT**
Thematic Priority:	**SUSTDEV-2004-3.VIII.2.e**

D7.2
Training Report on using QR for learning about Sustainable Development

Due date of deliverable: *not specified*
Actual submission date: <07/12/2007>

Start date of project: **1st March 2005** Duration: **30 months**

Organisation name of lead contractor for this deliverable:
University of Amsterdam[1]

Revision: final

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission	
RE	Restricted to a group specified by the consortium (including the Commission	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Abstract
This document describes evaluation studies of the products developed in work-package 4 and 6 of the NaturNet-Redime project. Particularly, it focuses on usability and effectiveness of the Garp3 workbench, and of the models developed in the case studies.

Document history

Version	Status	Date	Author
01	Draft – Initial structure and ideas	22/08/2007	Bredeweg
02	Included sections 3 and 4	27/08/2007	Bredeweg
03/04	Included and edited sections 5, 6 and 7	06/12/2007	Bredeweg
final	Inverted section 2 and 3. Included the new section 3, as well as section 1 and 8, and finalised the document.	07/12/2007	Bredeweg

[1] Authors: Bert Bredeweg, Paulo Salles, Dirk Bertels, Joram Rafalowicz, Anders Bouwer, Jochem Liem, Gisele Morison Feltrini, Ana Luiza Rios Caldas, Monica Maria Pereira Resende, Andreas Zitek, and Tim Nuttle.

Contents

1. Introduction

This deliverable discusses five evaluation activities. The studies address software and models developed in work-package 4 and 6, and how stakeholders (can) use these products for performing their tasks. The following activities are described.

- Section 2 describes an experiment in which students in formal education learn about Sustainable Development (SD) by interacting with the Garp3 workbench and models capturing SD concepts.
- Section 3 describes a study that evaluates the 'learnability' and 'memorability' of the icon language used in the Garp3 workbench. 'Easy to learn' and 'easy to remember' are considered key usability requirements for effective user interfaces (Sharp et al., 2007).
- Section 4 describes an evaluation of the Sketch environment (Liem et al., 2006) and how this can be used to support the development of insights by having learners create models. The study follows a small group of participants that work through all the steps of the 'Framework for building Qualitative Reasoning (QR) models' (Bredeweg et al., 2005; 2007) and reports noteworthy issues in this respect.
- Section 5 reports on a comprehensive set of activities concerning the Riacho Fundo case study (Salles, 2007). It addresses evaluations with three stakeholder groups: managers, teachers and domain experts. Section 5 also discussed theoretical background on how to evaluate 'conceptual models'.
- Section 6 addresses a smaller but similar kind of evaluation as described in section 5, and reports on evaluation studies concerning the River Kamp case study (Zitek et al., 2007).
- Section 7 describes a small evaluation of the 'Curriculum for learning about (SD) using QR' as it is implemented in Moodle (Nuttle and Bouwer, 2007).

This document focuses on the evaluation studies and their results, and not so much on the technical and implementation details of the software and the models developed in work-package 4 and 6. For the latter the reader is advised to consult the original documents describing those products.

2. Learning with interactive qualitative simulations

This evaluation focused on learning by interacting with a qualitative simulation using Garp3 and addressed the following hypothesis (cf. Rafalowicz, 2007):

- The Garp3 simulate environment can be used by novices with minimal instructions.
- By doing a guided treatment with the Garp3 simulate environment people can learn about the modelled domain.
- The learning effect will be greater if the model is presented with progressive complexity than when it is presented entirely at once.

Two domains were used: ‘Deforestation’ (DEF) and ‘Fuel and Global Warming’ (FGW), which both related to sustainable development. The context of this study was formal educational. The users were higher educational students with no prior knowledge of the domain and no prior experience with Garp3, but they did have general computer skills. The tasks were opening models, running simulations and inspecting the simulation results. The users had to answer questions about the simulation results. The users worked individually to complete the tasks. They were guided by predefined instructions and assignments. The users got immediate feedback on their assignments. The following equipment was used: Garp3 ran on Windows XP machines. Besides the Garp3 software, a web-browser was used to provide the instructions, the assignments and the pre- and post-test. Only the start-up instructions were on paper. The main knowledge transfer goal was for the users to learn about the cause-effect relations in the two domains.

2.1. Approach and expectations

To test the hypothesis a pre- and pos-test set-up was used with a test group and a control group. In the test group the pre-test was taken before the treatment and the post-test afterwards. The control group did not do a treatment between the tests. The control group is needed to rule out the possibility that the participants ‘learn’ from the pre-test or on their own between the pre- and post-test, thus making the results of the experiment less reliable.

In this experiment the learning effect of conceptual knowledge has to be measured. A good measure for this is the ability to do *prediction* and *post-diction* (or explaining). According to the hypothesis, the participants in the treatment group will learn about the domains by doing a treatment. If the participants learn form the treatment, their scores on the post-test should be significantly higher than on the pre-test. Moreover, the treatment is split in two parts, one for each domain. In the first part, the model on DEF is presented with a progression in complexity for each question set, in the second section the entire model on FGW is presented at once, without any model progression. According to learning theories progression will improve the learning effect (White and Frederiksen, 1990). It is thus expected that the participants will learn more on the first section than the second. The difference in scores between the pre- and the post-test should therefore be greater for the questions on DEF than those on FGW. Finally it is expected that there are no significant differences between test A and test B, for they are designed to be comparable in difficulty.

2.2. Experiment set-up

The experiment was conducted with 28 participants, divided equally over the treatment and control group. The participants in the control group were first year Information

Science students and the participants in the treatment group were first year Artificial Intelligence students, all at the University of Amsterdam. Both the treatment and the control group did a pre- and a post-test, consisting of multiple-choice questions on the domains of DEF and FGW. Between the pre- and the post-test the treatment group worked with Garp3 and the control group attended a lecture on an unrelated subject (Java programming). For the pre- and the post-test two tests were created: test A and test B (Appendix 2). Both tests were designed to be of equal difficulty and content. The tests contained 15 questions each, 9 on the domain of DEF and 6 on the domain of FGW. The distribution of the questions on the two domains was mixed, but comparable for both tests. The tests contained two types of questions, prediction and post-diction questions. Prediction question require forward reasoning: ‘A happens, what will be the effect?’ The post-diction questions require backward reasoning: ‘A is observed, what has caused this?’ The difficulty of a question is determined by its ‘dependency-path’, e.g. ‘A effects B, which effects C’ has a dependency-path of length 3. The questions of the tests were of varying difficulty (Table 2.1). Test A contained 6 forward (4 on DEF and 2 on FGW) and 9 backward reasoning questions (3 on DEF and 3 on FGW) with an average dependency-path length of 2.13. Test B contained 9 forward and 6 backward reasoning questions with an average dependency-path length of 2.33.

		Test A		**Test B**	
		Questions	*D-path*	*Questions*	*D-path*
Forward	*DEF*	4	2.5	6	2
	FGW	2	1.5	3	2
	Total	6	2.17	9	2
Backward	*DEF*	5	2.4	3	3.7
	FGW	4	1.75	3	2
	Total	9	2.1	6	2.8
Total		15	2.13	15	2.3

Table 2.1: Number of forward and backward reasoning questions (Questions) in Test A and B and the average length of the dependency-paths (D-path).

The treatment consisted of inspecting the simulation results of the DEF and FGW models with the Garp3 simulate environment. To guide the participants, they had to answer 12 sets of questions on different subjects related to the two domains, with a total of 44 questions (Appendix 2). The first 6 sets were on DEF and the second 6 sets were on FGW. To answer the questions the participants had to inspect dependencies of the model (shown in Figure 2.1) and value histories of quantities (shown in Figure 2.2). For the DEF domain each question set used a different scenario with increasing complexity, the questions on FGW used only one scenario. Domain experts created the questions.

2.2.1. Procedure: Treatment group

The treatment was conducted in a dedicated computer room especially reserved for the experiment. The participants were placed behind a computer and were handed out instructions and their personal id number. All aspects of the treatment were administered on-line using Garp3 and a web-browser. The first part of the experiment was the pre-test. Participants were randomly assigned. 50 % of the participants received test A as the pre-test and test B as the post-test. The other 50 % received test B as the pre-test and test A as the post-test. The participants had 15 minutes for the pre-test. During the treatment, the participants had approximately 1 hour and 15 minutes to work with the software and answer the treatment questions.

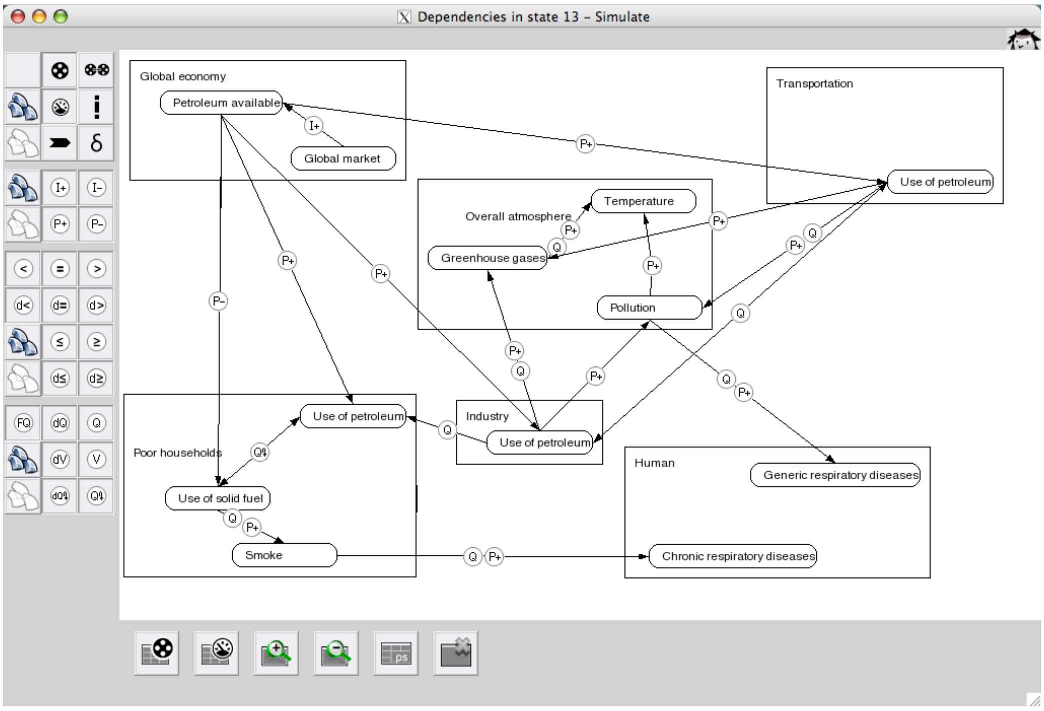


Figure 2.1: Screenshot of Garp3 dependencies of FGW model.

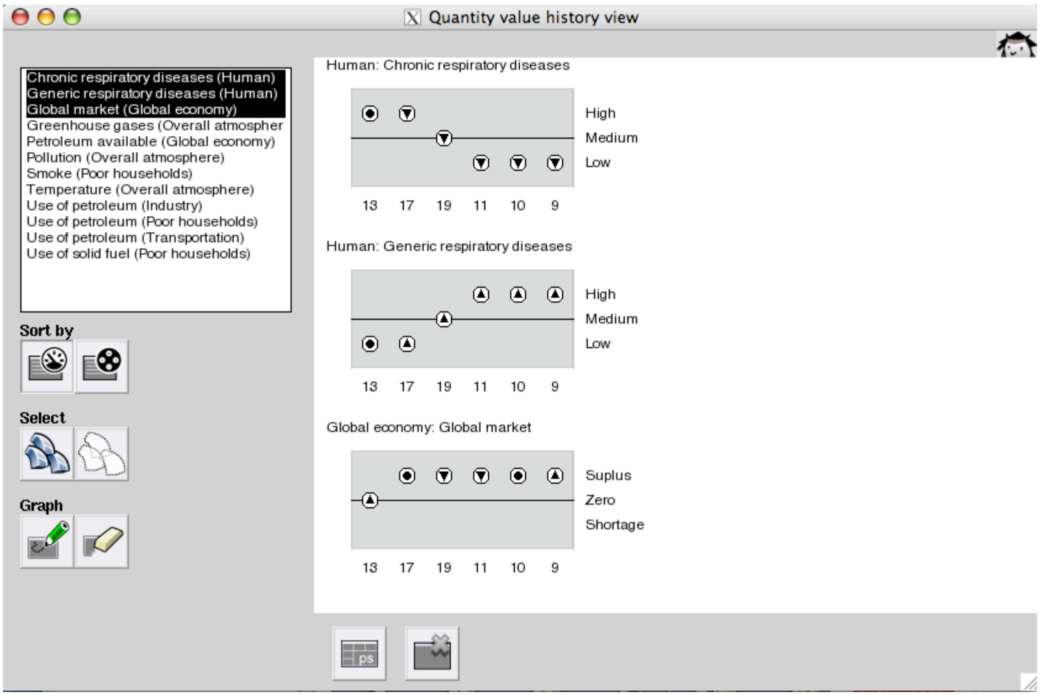


Figure 2.2: Screenshot of Garp3 value history of FGW model.

The web pages contained instructions (Figure 2.3), the treatment questions (Figure 2.4) and the answers with a visual (diagram) explanation (Figure 2.5). The answers and the time spend on each question set was logged for analysis. Figure 2.1 and 2.2 are representative screenshots of what the participants had to inspect in order to answer a treatment question. When the treatment was finished the participants did the post-test, again they had 15 minutes. After the post-test the participants were asked to fill out a questionnaire about their background and an attitude questionnaire on the experiment, both questionnaires were presented and filled out online using an interactive web page.

Figure 2.3: Screenshot of explanation web page during the treatment.

Figure 2.4: Screenshot of question web page during the treatment.

2.2.2. Procedure: Control group

The control experiment was conducted during a Java programming course. In a short introduction the experiment was presented. The participants were asked to try their very best to answer the questions correctly. They did not get any information about the domain nor about the real goal of the experiment. The questionnaires were handed out to the participants on paper, 50 % of the participants got test A as the pre-test and test B as the post-test. The other 50 % got test B as the pre-test and test A as the post-test. This order was assigned randomly. The participants had 15 minutes to complete the first test. After the pre-test the participants attended the regular class for 30 minutes and were then presented with the post-test. After the experiment was finished, the participants were provided with the opportunity to ask questions about the experiment.

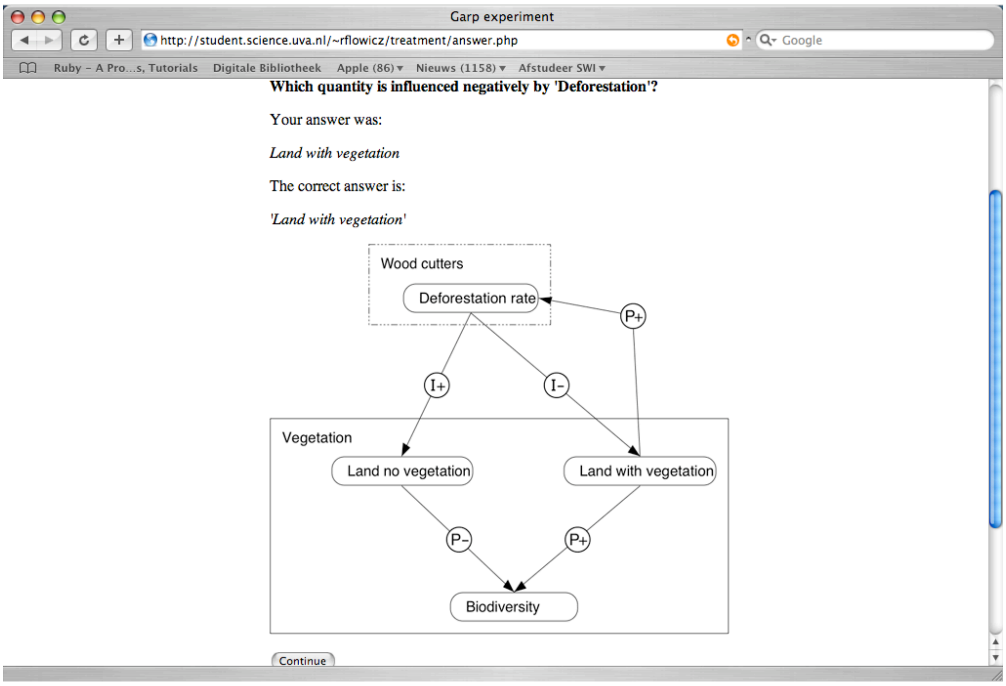


Figure 2.5: Screenshot of an answer web page during the treatment.

2.3. Results

All participants in the treatment and control group finished well within the time limits on the pre- and post-tests. Out of the 14 participants in the treatment, 9 finished all the treatment questions within the time limit, 3 almost finished and 2 finished roughly 50% of the questions, one of those two participants started significantly later than the others (due to circumstances outside scope of the experiment).

For the analysis of the results two-tied t-tests were used; an independent samples t-test for comparison between the control group and the treatment group (inter) and a paired samples t-test for comparison of the pre- and post-tests within a group (intra). To compute the results SPSS was used and excel was used for further analysis. The scores on the tests are in percentiles (values ranging from 0.00 to 1.00). The mean (M) and standard deviation (SD) of the group scores are of interest. For comparison of group scores the t-value (t) and the significance (SIG) are of interest. The results of the t-tests can be interpreted as follows: two sets of values are compared with a t-test, if the significance value is below a threshold, the difference of the values between the two sets is significant; it is not coincidental. For this experiment the threshold is 0.05, roughly meaning a certainty level of 95%.

2.3.1. Results of the pre- and post-tests

The results of the pre- and post-tests are in Table 2.2 and the results for the t-tests can be found in Table 2.3. The control group scored lower on the post-test (M=0.51) than on the pre-test (M=0.53), but this difference was not significant (t=0.418, SIG=0.683). The treatment group scored higher on the post-test (M=0.60) than on the pre-test (M=0.48) and the difference was significant (t=-2.249, SIG=0.043). On the pre-test the control group (M=0.53) scored higher than the treatment group (M=0.48), but this difference is not significant (t=0.843, SIG=0.407). On the post-test the treatment group (M=0.60) scored higher than the control group (M=0.51), but this difference was also not significant (t=-1.206, SIG=0.239). In Table 2.3 the results for the t-tests split by domain are also shown. This shows a significant difference between the treatment and control group on the post-tests for the FGW domain. Also no significant difference is measured between the pre- and post-test for the treatment group on the FGW domain.

Case	N	Mean	SD
Pre-test Control	14	0.53	0.20
Post-test Control	14	0.51	0.23
Pre-test Treatment	14	0.48	0.15
Post-test Treatment	14	0.60	0.20

Table 2.2: Average scores for the four cases.

Inter	ALL	DEF	FGW
Pre			
t	0.843	0.871	0.413
SIG	0.407	0.392	0.683
Post			
t	-1.206	-0.187	-2.526
SIG	0.239	0.853	0.018

Intra	ALL	DEF	FGW
Control			
t	0.418	-0.688	1.307
SIG	0.683	0.503	0.214
Treatment			
t	-2.249	-2.412	-1.389
SIG	0.043	0.031	0.188

Table 2.3: T-test scores for ‘inter’ (LHS) and ‘intra’ (RHS) groups. For all questions and split by domain. For the ‘inter’ tests negative t-values indicate a higher average for the treatment group. For the ‘intra’ tests negative t-values indicate a higher average for the post-test.

2.3.2. Further analysis of the test results

Individual performance: Table 2.4 and 2.5 show the average scores for the individual participants in the control group and the treatment group. In the treatment group t3, t7 and t40 score lower on the post-test than on the pre-test, the other participants improved their scores or stayed at the same level.

Pre-test, N=15			
ID	**Test**	**Mean**	**SD**
c1	A	0.80	0.41
c2	B	0.27	0.46
c3	A	0.53	0.52
c4	B	0.27	0.46
c5	A	0.27	0.46
c6	B	0.80	0.41
c7	A	0.73	0.46
c8	B	0.33	0.49
c9	A	0.53	0.52
c10	B	0.67	0.49
c11	A	0.80	0.41
c12	B	0.47	0.52
c13	A	0.53	0.52
c14	B	0.47	0.52

Post-test, N=15			
ID	**Test**	**Mean**	**SD**
c1	B	0.60	0.51
c2	A	0.13	0.35
c3	B	0.20	0.41
c4	A	0.60	0.51
c5	B	0.27	0.46
c6	A	0.87	0.35
c7	B	0.53	0.52
c8	A	0.80	0.41
c9	B	0.33	0.49
c10	A	0.67	0.49
c11	B	0.40	0.51
c12	A	0.80	0.41
c13	B	0.47	0.52
c14	A	0.40	0.51

Table 2.4: Average scores for the participants in the control group. Test A and B indicate which test the participant got as pre-test and post-test.

Difference between test A and B: The average scores on test A were higher than test B for the pre- and post-test of the control group and the pre-test of the treatment group, only for the post-test of the treatment group were the scores of test B higher than test A (Table 2.6, LHS). However none of these differences were significant (Table 2.6, RHS) supporting the assumption that test A and B are comparable in difficulty.

Pre-test, N=15					*Post-test, N=15*			
ID	**Test**	**Mean**	**SD**		**ID**	**Test**	**Mean**	**SD**
t1	A	0.80	0.41		t1	A	0.80	0.41
t2	A	0.27	0.46		t2	A	0.53	0.52
t3	A	0.40	0.51		t3	A	0.27	0.46
t4	A	0.33	0.49		t4	A	0.33	0.49
t5	A	0.60	0.51		t5	A	0.80	0.41
t6	A	0.53	0.52		t6	A	0.53	0.52
t7	A	0.60	0.51		t7	A	0.53	0.52
t21	B	0.47	0.52		t21	B	0.73	0.46
t22	B	0.60	0.51		t22	B	0.60	0.51
t23	B	0.33	0.49		t23	B	0.87	0.35
t24	B	0.27	0.46		t24	B	0.40	0.51
t25	B	0.40	0.51		t25	B	0.93	0.26
t26	B	0.47	0.52		t26	B	0.60	0.51
t40	B	0.60	0.51		t40	B	0.53	0.52

Table 2.5: Average scores for the participants in the treatment group. Test A and B indicate which test the participant got as pre-test and post-test.

Case	Test	N	Mean	SD
Pre-test Control	A B	7 7	0.60 0.47	0.19 0.20
Post-test Control	A B	7 7	0.61 0.40	0.26 0.14
Pre-test Treatment	A B	7 7	0.50 0.45	0.18 0.13
Post-test Treatment	A B	7 7	0.54 0.67	0.21 0.19

Case	t	SIG
Pre-test Control	1.230	0.242
Post-test Control	1.849	0.089
Pre-test Treatment	0.662	0.521
Post-test Treatment	-1.180	0.261

Table 2.6: Average scores of test A and B for the four cases (LHS), and t-test results for difference between test A and B for the four cases (RHS).

Difference between the two domains: In the pre- and post-test, for both the control and the treatment group, the performance on the FGW questions was lower than on the DEF questions (Table 2.7, LHS). This difference was only significant for the post-test of the control group (Table 2.7, RHS).

Case	Test	N	Mean	SD
Pre-test Control	DEF FGW	9 6	0.57 0.48	0.15 0.18
Post-test Control	DEF FGW	9 6	0.61 0.35	0.11 0.13
Pre-test Treatment	DEF FGW	9 6	0.50 0.44	0.19 0.20
Post-test Treatment	DEF FGW	9 6	0.63 0.57	0.19 0.12

Case	t	SIG
Pre-test Control	1.113	0.286
Post-test Control	4.271	0.001
Pre-test Treatment	0.564	0.583
Post-test Treatment	0.650	0.527

Table 2.7: Average scores of on the DEF and FGW domains for the four cases (LHS), and t-test results for difference between the DEF and FGW domains for the four cases (RHS).

Scores on the treatment questions: The treatment consisted of 12 sets of questions divided over the two domains. The participants had to work with 7 different scenarios, 6 incrementing scenarios on DEF and 1 scenario on FGW. The averages in Table 2.8 (LHS) are computed over the questions that are answered. Not all participants finished all the questions of the treatment. Furthermore every set did not have the same amount of questions and some questions had multiple parts, therefore N differs over the sets. There was a significant difference on the number of correct answers (t=2.912, SIG=0.005) between the questions on DEF (M=0.93) and FGW (M=0.79), set 8 scored especially low.

None of the participants scored notably low on the treatment questions they answered (Table 2.8, RHS), however t3 (N=39) and t40 (N=28) did not get very far, these are also two of the three participants who scored lower on the post-test than on the pre-test.

Scores per question set				
Set	**Domain**	**N**	**Mean**	**SD**
1	DEF	70	0.93	0.26
2	DEF	70	0.89	0.32
3	DEF	56	0.98	0.13
4	DEF	70	0.93	0.26
5	DEF	56	0.89	0.31
6	DEF	83	0.94	0.24
7	FGW	78	0.85	0.36
8	FGW	52	0.52	0.50
9	FGW	120	0.94	0.24
10	FGW	48	0.60	0.49
11	FGW	47	0.83	0.38
12	FGW	39	0.74	0.44

Scores per participants			
ID	**N**	**Mean**	**SD**
t1	58	0.91	0.28
t2	56	0.96	0.19
t3	39	0.90	0.31
t4	61	0.74	0.44
t5	61	0.80	0.40
t6	61	0.85	0.36
t7	61	0.84	0.37
t21	61	0.93	0.25
t22	61	0.89	0.32
t23	61	0.75	0.43
t24	61	0.87	0.34
t25	61	0.85	0.36
t26	59	0.88	0.33
t40	28	0.89	0.31

Table 2.8: Average scores of the question sets in the treatment, were N denotes the total number of questions answered by participants for the question set (LHS), and average individual scores on the treatment questions, were N denotes the number of questions answered by each participant (RHS).

Time spent: The average time spent on question sets in the treatment lies between 50 and 150 seconds for most sets (Figure 2.6) except for set 1 and 8, which took much longer. This was expected for the first set, since it included the introduction, so it contains the start-up time, there is no such reason for set 8. Notice that the domain switch from DEF to FGW while starting with set 7.

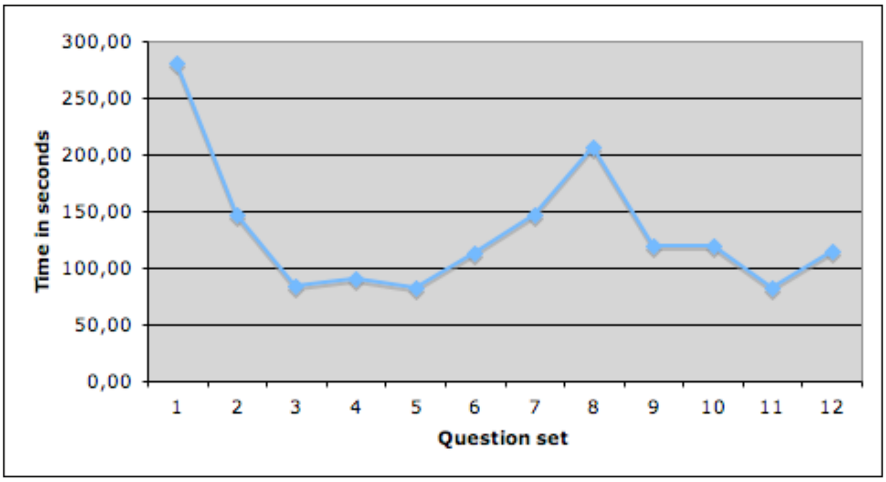


Figure 2.6: Average time spent on the treatment question sets.

Background questionnaire: The questions of the background questionnaire (Appendix 2) were answered on a scale form 1 to 7, 1 being the lowest score and 7 the highest. The participants in the treatment condition were all first year Artificial Intelligence students, 12 male and 2 female, with an average age of 19. From the average scores on the background questionnaire (Table 2.9) it appears that the participants were experienced with computers (5.36), averagely experienced with qualitative reasoning (4.43) and conceptual modelling (3.93) and not very experienced in ecology (2.64). This is what is to be expected of a first year Artificial Intelligence student, except for the somewhat high score on qualitative reasoning. The participants had not yet attended any classes on qualitative reasoning nor received any instruction on this. It is also very unlikely that they had gained experience in qualitative reasoning in any another way. We therefore assume that the participants misunderstood this particular question and that their actual understanding of qualitative reasoning is much lower and not measured accurately by the questionnaire.

Question	**Mean**	**SD**
Computer experience?	5.36	1.15
Ecology expertise?	2.64	1.08
Conceptual Modelling expertise?	3.93	0.92
Qualitative Reasoning expertise?	4.43	1.02

Table 2.9: Average scores of the background questionnaire, N=14. The scores range from none to very high on a scale of 1 to 7.

Attitude questionnaire: The questions of the attitude questionnaire (Appendix 2) were also answered on a scale form 1 to 7. The ease of use and the understandability of the diagrams scored high (Table 2.10), the other issues scored slightly above average.

Question	**Mean**	**SD**
How much was learned?	4.00	1.24
Difficulty of the test questions (1=difficult, 7=easy)?	4.36	1.60
Ease of use of the software interface?	5.06	1.07
How easy did were the diagrams to understand?	5.29	1.07
How enjoyable was the session?	3.71	1.27
How enjoyable was it using the software?	3.93	1.21

Table 2.10: Average scores of the attitude questionnaire, N=14. Unless noted otherwise, scores range from low to high on a 1 to 7 scale.

2.4. Concluding remarks

The results of the experiment support the hypothesis that people can learn conceptual knowledge through observing and inspecting qualitative simulations, even within very limited time. There was no significant difference between the pre- and post-test in the control group, thus it is safe to conclude that any difference observed in the treatment group is a direct effect of the treatment itself. Almost all the participants in the treatment group scored higher on the post-test than on the pre-test. Further supporting this conclusion is the observation that the two participants (t3 and t40), who only got about halve way through the treatment, are also two of the three participants that scored lower on the post-test than on the pre-test. Since there are no significant differences between test A and B in any of the four cases (Table 2.6), the tests can be considered comparable in difficulty.

According to theory on learning, learners learn more if the subject matter is worked through progressively, as done with the DEF domain, than if the subject matter is presented as a single large chunk, as done with the FGW domain. Overall the scores of the questions on the FGW domain were lower than those on the DEF domain, but the only significant difference between those scores was in the post-test in the control group. In fact, as Table 2.3 shows, there is no significant learning effect for the treatment group on the FGW domain (t=-1.389, SIG=0.188), only on the DEF domain (t=2.412, SIG=0.031). The treatment question scores show a similar result, significant lower scores on the FGW questions compared to the scores on the DEF questions. This all supports the theory that model progression improves learning.

An alternative explanation for the difference in learning could be that the subject matter on FGW was more complex. This is however less likely. First, the questions in the pre- and post-test seem less difficult for the FGW, because of a shorter average causal path (Table 2.1). Moreover, both models contain 12 entities, but the DEF model has in total 16 dependencies while the FGW model has 14 dependencies. The longest causal path in the DEF model has a length of 5, while the longest path in the FGW model has a length of 4. These facts suggest that the FGW model is not more complex than the DEF model, on the contrary.

The time spike for question set 8 of the treatment can be explained by the nature of the questions in this set, especially the second question. While most questions required the participants to inspect some values and relations, the second question of set 8 asked the participants to compare the behaviour of multiple quantities. This most likely took more time than the other questions, because the participants had to inspect and compare multiple value histories.

3. Evaluation of the Garp3 icon language

Usability is an important aspect of user interfaces, and includes issues such as: effective to use, efficient to use, safe to use, have good utility, easy to learn, and easy to remember how to use (cf. Sharp et al., 2007). Particular the issues of an interface being 'easy to learn' (also referred to as 'learnability') and 'easy to remember' (also referred to as 'memorability') are considered important, because they provide the basis for other issues. The Garp3 icon language was therefore evaluated on these two aspects.

3.1. Garp3 icon design

As the vocabulary used in Qualitative Reasoning (QR) is in principle unknown to non-computer scientists it is almost impossible to develop a set of icons that is immediately understood by such users. By definition, the vocabulary and the corresponding icons have to be *learned* by the users. Therefore, more important than immediately 'understanding' the meaning of an icon, is the ability for users to learn the meaning and remember the meaning for future use of the software. To support this process the icons in the Garp3 workbench were organised in a structure that highlights their meaning as much as possible. Table 3.1 shows one of the key icons of this vocabulary, the 'blue earth circle'. This icon refers to the *model* that is processed using the workbench at a certain moment. There are three types of manipulations that a user can perform: *filing* a model (save, open, etc.), *building* ingredients to compose a model (create entities, quantities, etc), and *simulating* a model (running a specific scenario, inspecting the generated causal model, etc).

Table 3.1: Model, model ingredients, and model simulation

		
File: Model	Build: Ingredient	Simulate: Model

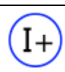
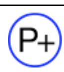
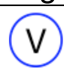
Based on these initial icons for 'model' and 'model ingredient' further icons were developed to signify the different ways in which models and ingredients can be manipulated. Some of the typical icons are shown in Table 3.2.

Table 3.2: Model and model ingredient action buttons

					
Save ingredient in model	Delete ingredient from model	Create new ingredient	Show ingredient properties	Copy ingredient	Delete ingredient

Similar to the icons for operating the software, a set of icons was developed to signify the model ingredients. A subset of those is shown in Table 3.3 and Figure 3.1 illustrates how these icons are used to create model fragment contents.

Table 3.3: A small selection of the Garp3 icons for model ingredients

						
Model fragment	Entity	Quantity	Influence *(positive)*	Proportionality *(positive)*	Value correspondence	Agent

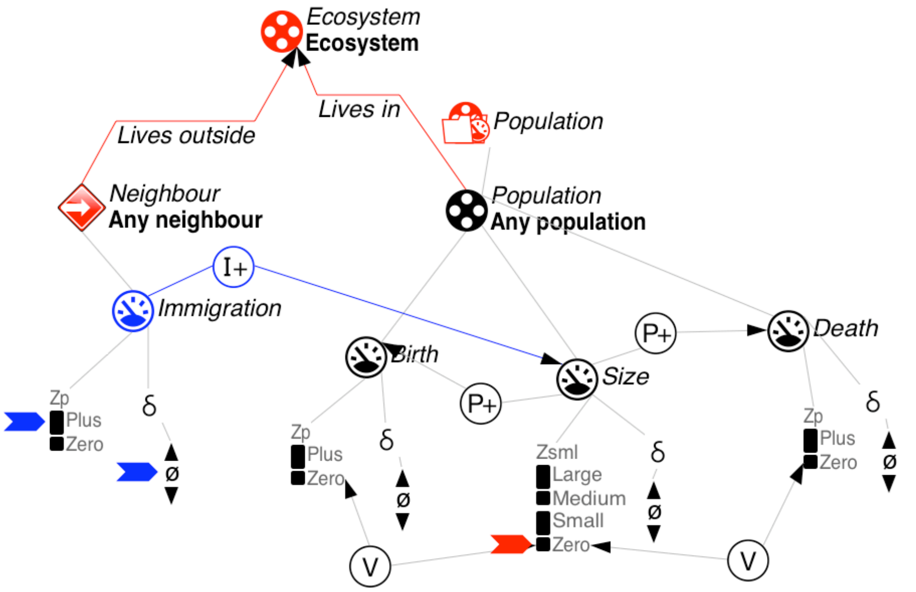


Figure 3.1: Model ingredient icons visualising model fragment contents, in this case the notion of ‘Colonisation’ (cf. Salles and Bredeweg, 2003).

Summarising, to increase the usability of the Garp3 workbench a graphical language was developed from which icons were derived to signify the meaning of the underlying vocabulary and how to operate the software (for details see Bredeweg et al., 2006a; 2006b; Bertels, 2007). The study performed to evaluate the result of this approach is described below.

3.2. Setup and materials

The Garp3 icon language was evaluated in a series of studies performed during the QR user group meetings in Amsterdam (11-14, October, 2005) and Sophia (Central Laboratory of General Ecology in Sofia, Bulgaria 27-31 March, 2006), and during the QR summer school of 21st International Workshop on Qualitative Reasoning in Aberystwyth (27-29 June, 2007).

A questionnaire was developed (Appendix 1) consisting of 50 icon-items representing the approximately 300 different kinds of icons used in the software. Table 3.4 shows one of the items. Each item had the form of a multiple-choice question with which the user had to identify the meaning of an icon by selected one of the 4 available options (in Table 3.4 the correct answer is shown in italic).

Table 3.4: Example question used for the Garp3 icon language testing

Icon	*What action do you associate with this icon?*	*Your answer:*
	1. Select a scenario to simulate 2. Select a path 3. *Full simulation ‘current’ scenario* 4. Add new item	

Three hypothesis were formulated:

- New users have no knowledge of the QR vocabulary (hence: initial users will score low on the questionnaire).
- Users will easily learn the meaning of the icons when working with the software (hence: after working with the software users will score significantly higher on the questionnaire compared to new users)

- Users will remember the meaning of the icons even when they are not actively using the software (hence: after not working with the software for a certain amount of time, users will still score significantly higher on the questionnaire compared to new users, and score similarly to users who have been using the software)

The above described questionnaire was administered three times, following the schema shown in Table 3.5, although only a subset of the participants did the 2nd post-test (T3). All participants had little to no prior experience with QR. Before being introduced and having worked with the software, the participants had to identity the meaning of the icons (T1). After working with the software for approximately one day, they filled out T2, examining whether they had mastered the icons used in the software. Notice that there was no explicit explanation of the icon language during this 'treatment' phase. Working with the software meant doing modelling assignments and receiving feedback on the correctness of their results. Finally, after not having worked with the software for approximately 6 months a subset of the participants filled out T3, examining to what extend they remembered the correct meaning of the icons.

Table 3.5: Evaluation schema

Pre-test (T1)	*Using software*	*Post-test (T2)*	*Not using software*	*Post-test (T3)*
20 min	6-8 hours	20 min	6 months	20 min

3.3. Main results

The results (correct answers on the questionnaire) are shown in Table 3.6 and Figure 3.2. Test T1 and T2 was filled out by 18 participants and T3 by 8. The questionnaire had 50 items. The average number of correct answers on T1 was 26, on T2 it was 38, and on T3 it was 42.

The Wilxon signed rank test shows that there was a statistically significant change in the scores on T2 and T3, compared to T1 ($p<0.001$). There was no significant difference between T2 and the T3 (the follow-up).

Table 3.6: Total correct scores on questionnaires per participant.

Participant	Pre-score (T1)	Post-score 1 (T2)	Post-score 2 (T3)
1	23	44	44
2	24	45	45
3	14	32	40
4	29	42	48
5	37	43	46
6	20	25	42
7	33	40	45
8	25	39	33
9	36	46	-
10	12	24	-
11	25	44	-
12	40	44	-
13	32	44	-
14	21	29	-
15	31	39	-
16	22	43	-
17	27	33	-
18	24	40	-

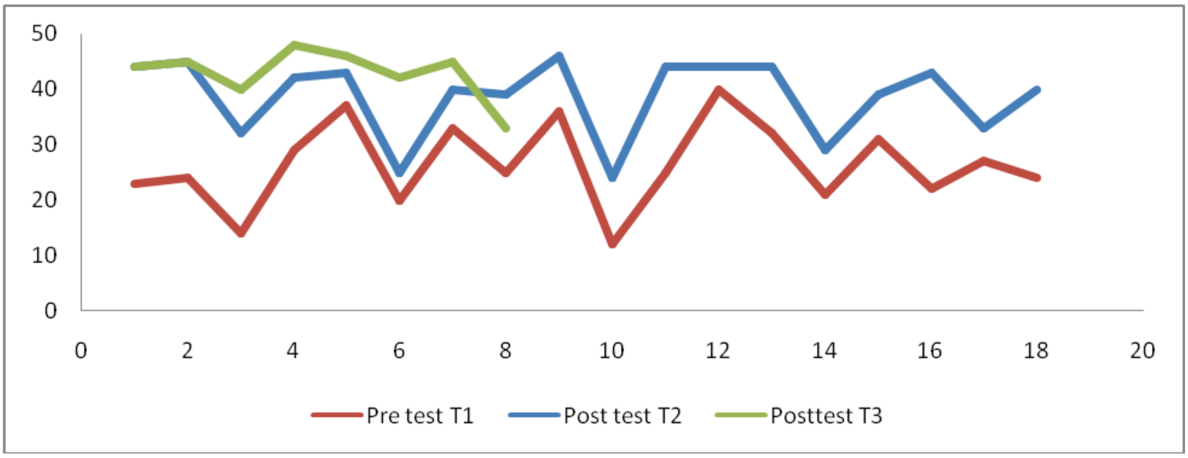


Figure 3.2: Total correct scores on the questionnaires per participant.

The average scores on T1 and T2 for each of the icons are shown in Figures 3.3, 3.4, 3.5 and 3.6. Notice that for almost all icons there is an increase in the number of correct answers, although there is variation in the amount of increase. The icons shown in Figure 3.7 illustrate sets of icons that are interesting for further discussion. Icon (a) refers to a group of icons that already have a high score on T1 and also have a high score on T2. Apparently, the meaning of these icons was immediately clear to the participants, and remained clear when using the software. Icon (b) refers to a group of icons for which the score on T2 drops compared to T1, which suggest that the meaning became somewhat unclear for the participants. The reason for this could be a combination of ‘confusion’ and ‘lack of use’. The icons for which the scores decrease are part of sets of icons that look rather similar and have slightly varying behaviour in a similar contest. To know the difference between these icons users need to understand in detail how the system operates. Moreover, these icons are for advanced usage of the software, typically associated with ‘expert behaviour’. Thus when first confronted with these icons users had ‘some idea’ of their meaning (T1), but instead of developing this insight further they confused the closely related icons, due to insufficient exposure and use of the icons involved (T2). Notice, that the occurrence of this phenomena is rather rare for the workbench (e.g. only 2 out of the 50 icons in the questionnaire had it). Moreover, it is expected that the meaning of these icons will quickly become apparent when users have to actually use them to perform their tasks.

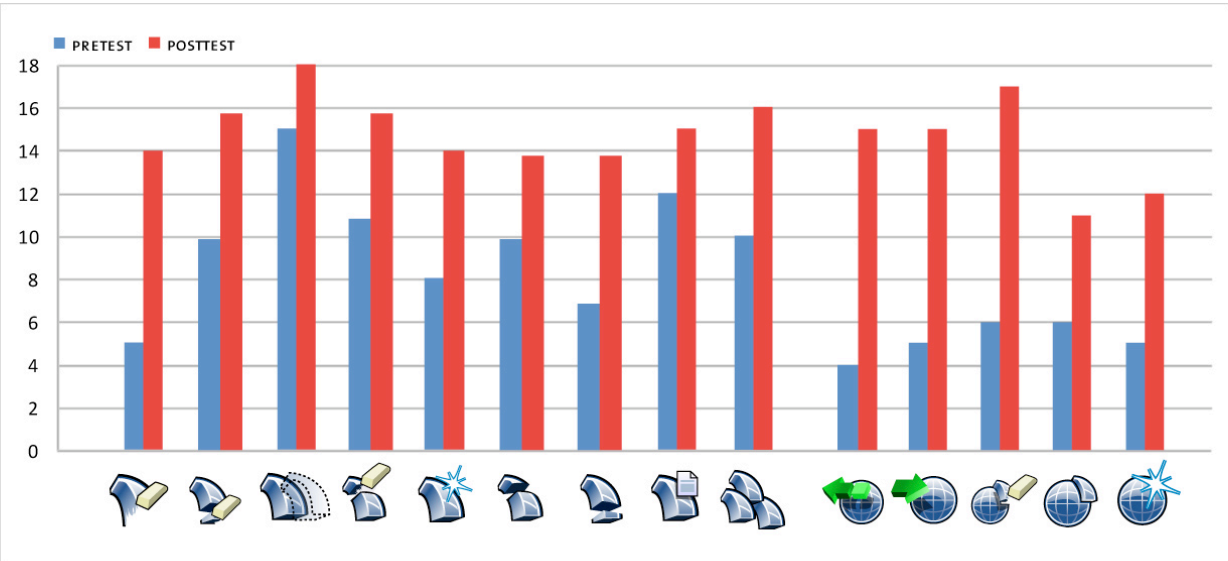


Figure 3.3: Scores on the icon questionnaire: Build Environment (part 1) and File manipulation.

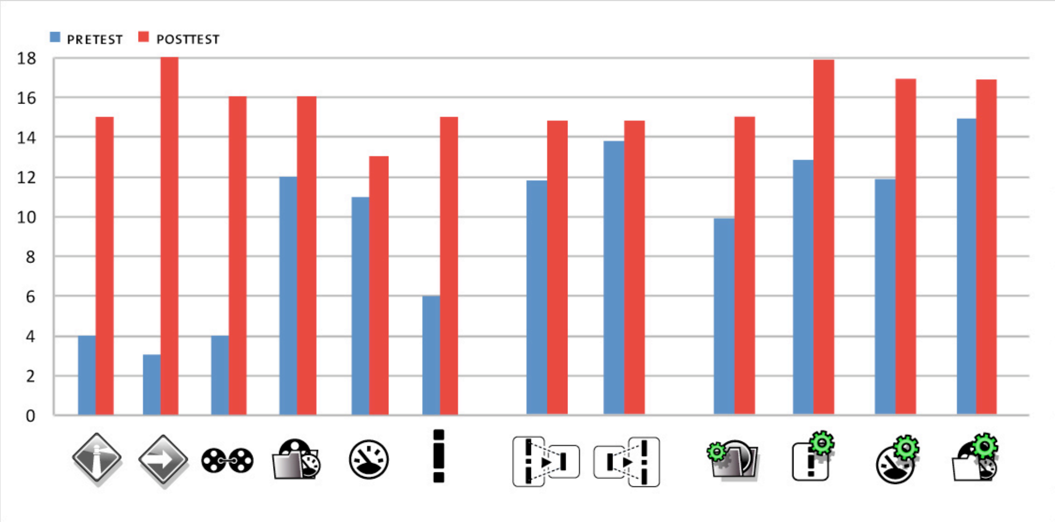


Figure 3.4: Scores on the icon questionnaire: Build Environment (part 2).

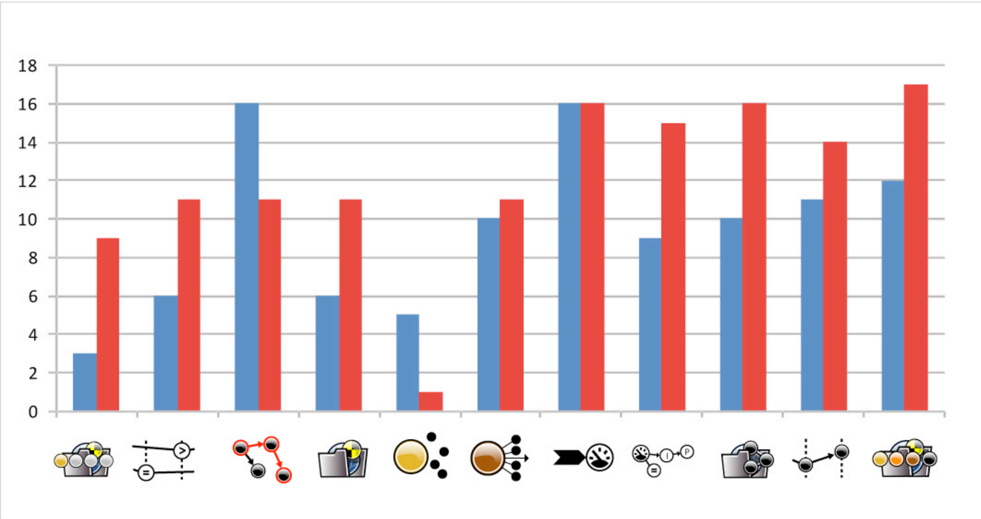


Figure 3.5: Scores on the icon questionnaire: Simulate Environment.

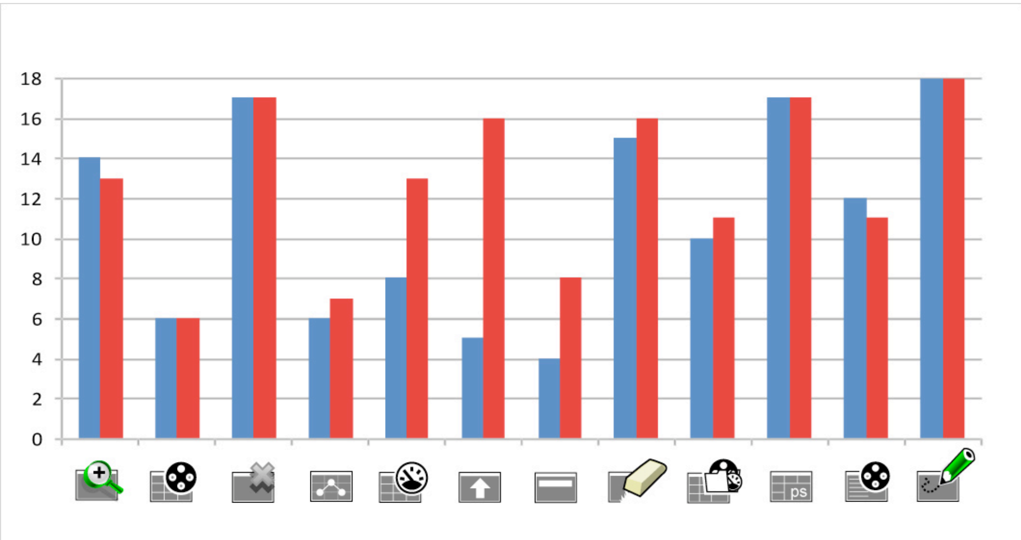


Figure 3.6: Scores on the icon questionnaire: Window and Content layout manipulation.

(a): Draw (b): Select a path (c): Save current model to a different file

Figure 3.7: Icons with some specific features

Icon (c) refers to a third group of icons. This group shows a rather low score on T1 and a much higher score on T2 (e.g. the five right most icons in Figure 3.3). This seems to suggest that these icons are ‘easier’ to learn than other icons. The reason for this may be that the functionality implemented by these icons refers to in general well-known aspects of user interfaces such as ‘open file’, ‘save file’, etc. However, instead of using commonly known icons for these cases in the Garp3 workbench, dedicated icons have been created. Initially these icons are thus not familiar to users (hence the low score on T1), but after encountering the icons in the software, the participants easily associated the well-known behaviour to these new icons (hence the high score on T2). To accommodate new users, it would be an option to use more common icons for these well-known functions. On the other hand, ensuring that all icons consistently fit into an overall ‘graphical language’ is beneficial for learnability and memorability. For the Garp3 workbench the latter was considered more important and therefore taken as approach. Finally, notice that the average number of correct scores on T1 is already significantly higher than random. In a way, this contradicts the hypothesis that new users ‘have no knowledge of the QR vocabulary’, and supports the idea that the icons correctly communicate their purpose (even at first encounter), and are thus well designed.

3.4. Concluding remarks

The Garp3 workbench was developed to offer easy access to high-end qualitative simulation software, providing non-AI/QR experts the possibility to use QR technology without having to understand low-level implementation details of such automated reasoners. A series of evaluation studies have been performed to investigate whether this goal was realised. These evaluation studies support the hypothesis that the workbench interface is easy to learn and easy to remember, addressing key usability goals concerning software development.

4. Discovering knowledge by building models

Building a qualitative model is a complex task. To support this task a structured approach to building qualitative models was developed (Bredeweg et al., 2005; 2007). The Garp3 workbench supports this approach in the form of the sketch environment. To examine how this approach supports the modelling effort an exploratory case study has been carried out which is described in this section. This study focused on actual usage of the structured approach to qualitative modelling using the sketch environment in an educational setting (cf. Rafalowicz, 2007). The goal was:

- to determine how working with the sketch environment helped the modelling process,
- to determine how working with the sketch environment supported learning,
- to map the issues which were encountered by the users of the sketch environment, and
- to create recommendations for guidance of structured approach in an educational setting.

4.1. Method

A case study was chosen as a means to gain deeper insights in applying the structured approach using Garp3. A qualitative in-depth exploratory case study was conducted in an educational setting. Three participants were followed in their model-building efforts following the structured approach. Observations of the model building process, weekly open-ended questionnaires, and analysis of intermediate model representations were combined in this study to gain insight into the process of working with Garp3 and the sketch environment. The participants gave informed consent to being part of this study.

Domain: Each participant was presented with one of three modelling problems. These three problems were related to sustainable development. Each of the problems was a target in the Millennium Development Goals (MDG) project of the United Nations (http://www.unmillenniumproject.org). The MDG project is a combined effort to improve the quality of life in third world countries by the year 2015. The MDG project contains 8 goals to be achieved by 2015, each goal has a number of targets with corresponding indicators. The modelling problems for this course were each based on one of the three targets for goal 7: Ensure Environmental Sustainability (see Table 4.1). The participants had to build models that would give better insight into how the targets could be achieved. This was achieved by explicating the structure and causal relations of the involved systems in a qualitative model.

Users: The users were Artificial Intelligence MSc students at the University of Amsterdam. They had some experience with model building using Garp3 and qualitative reasoning theory. The users had computer experience, but little prior domain knowledge.

Environment: This case study involved three students enrolled in the course ‘Qualitative Reasoning’ at the University of Amsterdam. The course lasted 16 weeks and was divided into two parts. The first part contained theory and practical exercises on qualitative reasoning and Garp3, in the second part the students had to do a practical project related to Garp3. Three of the students in this course decided to build a qualitative model as their project, and were therefore chosen as the participants of this study. Two experts supervised the students during the 8 weeks of the modelling. The students presented and discussed their progress at weekly meetings. During these meetings the experts commented on the models and also provided explanations on the domain and modelling techniques. The students were guided in their modelling effort by weekly assignments, corresponding to the steps of the structured approach mentioned above. The students worked individually on their models. There was no predetermined work environment.

Tasks: The users had to use the sketch environment to design their model and implement a functioning model with the build and simulate environment.

Equipment: The users used Garp3 and worked on their own computers. For this task the users were given literature on the domain (Lee and Ghanime, 2004), (United Nations Development Group, 2001), (Smeets and Weterings, 1999) and modelling the domain (Salles et al., 2005). The users were also given an article on the structured approach (Bredeweg et al., 2007).

Goals: The task goals of this usage were to create a working qualitative model and fully document it. The knowledge transfer goals were to learn more about qualitative reasoning and about the domain.

Target 9	Integrate the principles of sustainable development into country policies and programs and reverse the loss of environmental resources.
Indicators	25. Proportion of land area covered by forest (FAO) 26. Ratio of area protected to maintain biological diversity to surface area (UNEP-WCMC) 27. Energy use (kg oil equivalent) per $1 GDP (PPP) (IEA, World Bank) 28. Carbon dioxide emissions per capita (UNFCCC, UNSD) and consumption of ozone-depleting CFCs (ODP tons) (UNEP-Ozone Secretariat) 29. Proportion of population using solid fuels (WHO)
Target 10	Halve, by 2015, the proportion of people without sustainable access to safe drinking water and basic sanitation.
Indicators	30. Proportion of population with sustainable access to an improved water source, urban and rural (UNICEF-WHO) 31. Proportion of population with access to improved sanitation, urban and rural (UNICEF-WHO)
Target 11	Have achieved by 2020 a significant improvement in the lives of at least 100 million slum dwellers.
Indicators	32. Proportion of households with access to secure tenure (UN-HABITAT)

Table 4.1: Millennium Development Goal 7: Ensure Environmental Sustainability

4.1.1. Observation of modelling process

The weekly meetings were recorded on video for further analysis. The videos were analyzed on for a temporal division. The temporal division is an analysis of proceedings of the meetings. The meetings are divided temporally according to the following categories:

- *Participant:* when a Participant is presenting his model, this is usually only in the beginning of their allotted time.
- *Comment:* most likely given by one of the teachers, it is a comment on what the participant has created, this can lead to further discussion or explanation.
- *Discussion:* When an issue arises and it is not clear how this should be resolved or someone is not convinced a certain approach is correct it can be discussed.

- *Explanation:* When more explanation is required, one of the teachers could elaborate on certain subjects. During the explanation the teacher could ask questions to the participant for more interaction.

All meetings were examined extensively and were categorised the category systems described above.

4.1.2. Open-ended questionnaire

The participants were asked to answer a few open-ended questions each week on what they did and what issues they encountered for the purpose of this study (Table 4.2).

Name
Date
Representation
What did you do?
What were the problems you encountered?
What is the status of the representation?
Remarks
Are you satisfied with the representation?

Table 4.2: Open-ended questionnaire for weekly report by participants.

4.1.3. Intermediate model representations

To analyse the structured approach the model progress was mapped for the three participants. The structured approach contains a number of assumptions with regards to the re-use, refinements and formalisation of ingredients[2] (see figure 4.8). According to the structured approach the ingredients should be re-used and refined in following steps rather than creating new ingredients in each step. The structured approach is a process of refinement and iteration. For analysis all the intermediate representations were examined on content for each week. The representations were compared to the representations of the previous week and other current representations. This analysis focussed on where the ingredients came from, where they re-used or refined from another representation or were they new. For the implementation of the model only the final model was examined for this analysis. This analysis should give an insight in the model building process and the structured approach.

To analyse the formalisation of the intermediate representation into the actual model the final intermediate representations were compared to the final model. Each of the formalisation steps (see in Figure 4.8 the 'formalises into' arcs) were measured for each participant and for all the participants combined. The ratio between the re-used, refined and new ingredients in the final model is used to compare the different approaches and the different formalisation steps. The goal of this analysis is to see if there is a difference in ratios between the three participants and between formalisation steps. If a certain formalisation step has a lower ratio than other steps it could mean that the participants did not create a correct intermediate representation, did not formalise correctly or that the sketch environment does not support the formalisation correctly.

4.2. The modelling effort

Each week the participants got assignments on what to do. These assignments corresponded with the structured approach. The next sections describe what the participants had to do each week. The participants had to present and discuss their

[2] Ingredient is the term used to refer to elements in the Garp3 Build, Simulate and Sketch environments.

progress each week at the meetings. At the end of each meeting the participants got their next assignments.

- *Week 1*: During the first meeting three modelling problems were presented to the participants, who could choose which one they would like to model. All three problems were targets of the sustainable development project. After choosing which target they had to model, the participants had the rest of the week to study the target documentation, fill in the ‘Abstract, model goals, intended audience and general remarks’ and create the concept map. This corresponds with the first step of the structured approach: orientation and initial specifications.
- *Week 2*: In the second week the participants had to create the structural model, the processes and the actions and external influences. This corresponds with the second step and part of the third step of the structured approach: System selection and structural model and Global behaviour.
- *Week 3*: In week 4 the participants had to create the causal model and the behaviour graph. This is the rest of the third step of the structured approach: Global behaviour.
- *Week 4-6:* The participants had three weeks to build the actual model. This corresponds with step 5 of the structured approach: Implementation.
- *Week 7-8:* The last two weeks were for writing the report and model documentation. This corresponds with step 6 of the structured approach: Model documentation. The final meeting took place in week 7, in this meeting the participants had to give their final presentation.

4.3. Results

4.3.1. Intermediate representations

The intermediate representations made each week were examined for each participant. To see how the iterative process of the structured approach worked an overview was created for each participant. The results of this analysis can be seen in Figures 4.1, 4.2 and 4.3. The three participants are referred to as P1, P2 and P3.

The blocks in these figures represent the different sketch representations (e.g. concept map, structural model) and build representation (e.g. quantities and model fragments). In each block the different ingredients for each representation are denoted and the amount of ingredients present in that representation (e.g. 28 concepts and 9 relations in concept map in Figure 4.1). The vertical lines separate each iteration and the dashed vertical line separates the sketch representation from the build representation. P2 and P3 worked on their sketch representation one week more than P1, and therefore have one iteration extra. The lines between ingredients denote re-use or refinement of sketch ingredients and formalisation of sketch into build ingredients. The number above the line represents the number of re-used ingredients, while the number below the line represent the number of refined ingredients (e.g. 5 concepts are re-used and 5 concepts are refined as entities in the structural model in Figure 4.1). For the model fragments block, the instantiations denote the amount of model fragments that are specific variants of another model fragment.

Figure 4.1 shows less relations than Figures 4.2 and 4.3, this indicates less refinement and re-use steps. Table 4.3 summarizes Figures 4.1, 4.2 and 4.3. The total amount of ingredients denoted in the figures are given and also split into sketch and build sets. The total number of re-uses and refinements are counted and the percentage of the total amount of ingredients calculated. The table shows that P1 often created new ingredients during the modelling effort; 58.5% of all created ingredients were new, as opposed to 24.3% and 43.3% by P2 and P3 respectively. All the participants' final models were of comparable size in terms of number of ingredients.

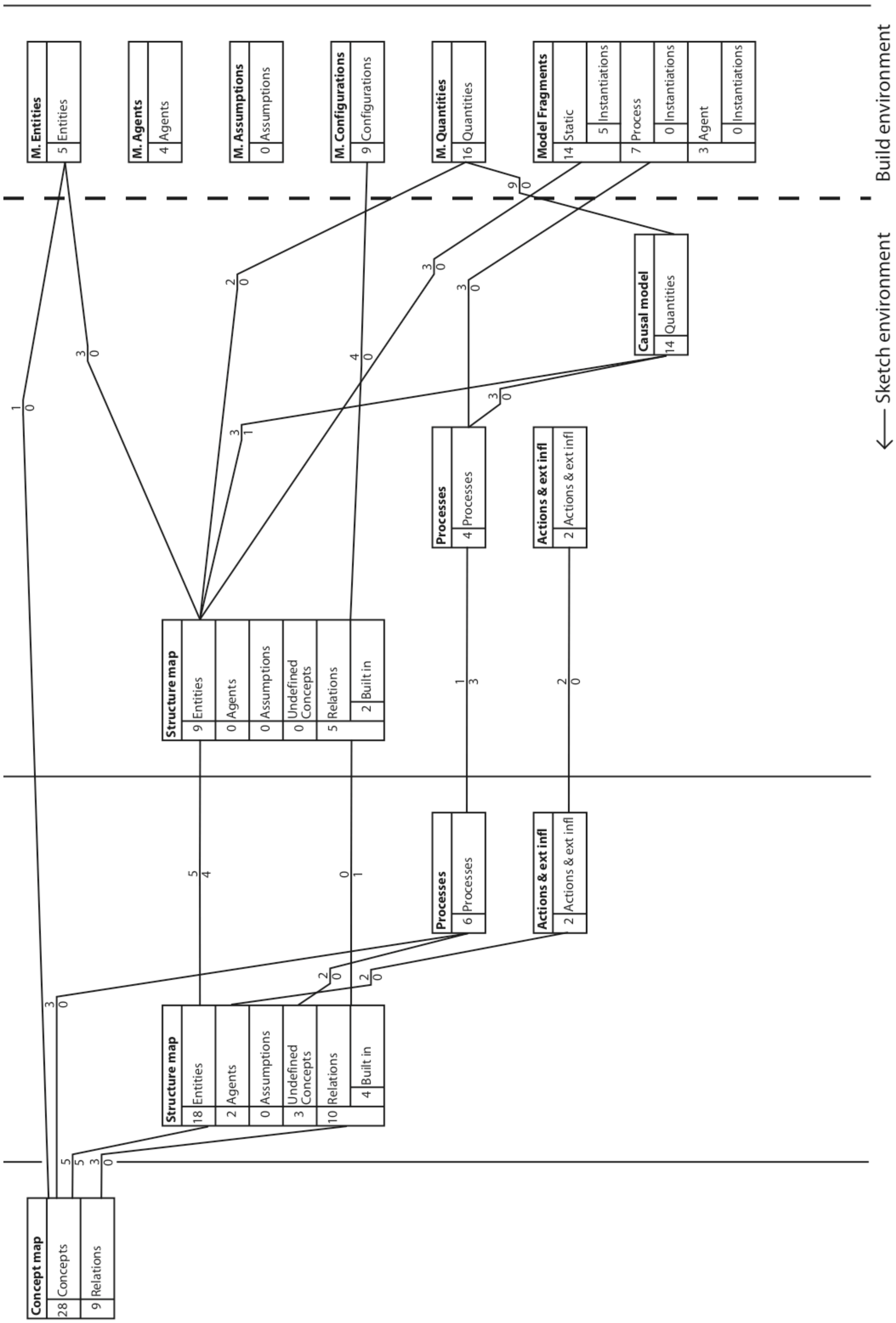


Figure 4.1: Overview of intermediate results for P1.

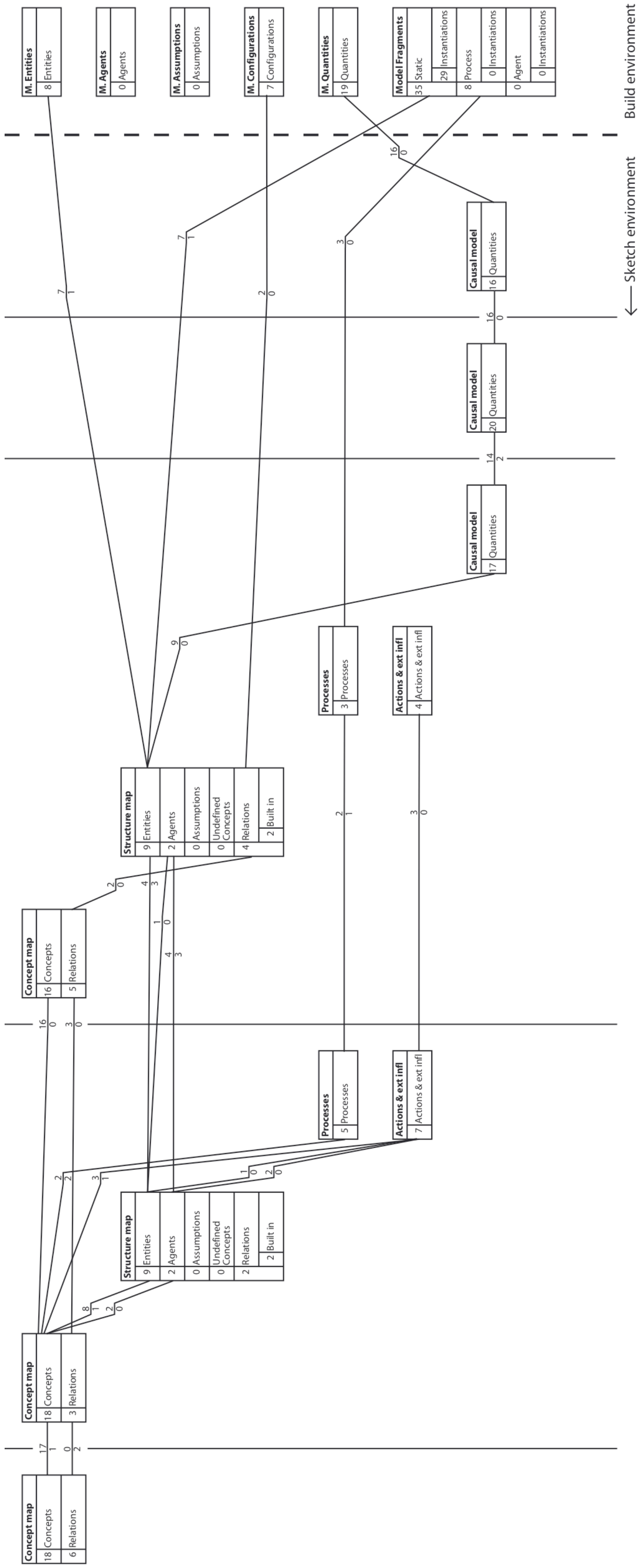


Figure 4.2: Overview of intermediate results for P2.

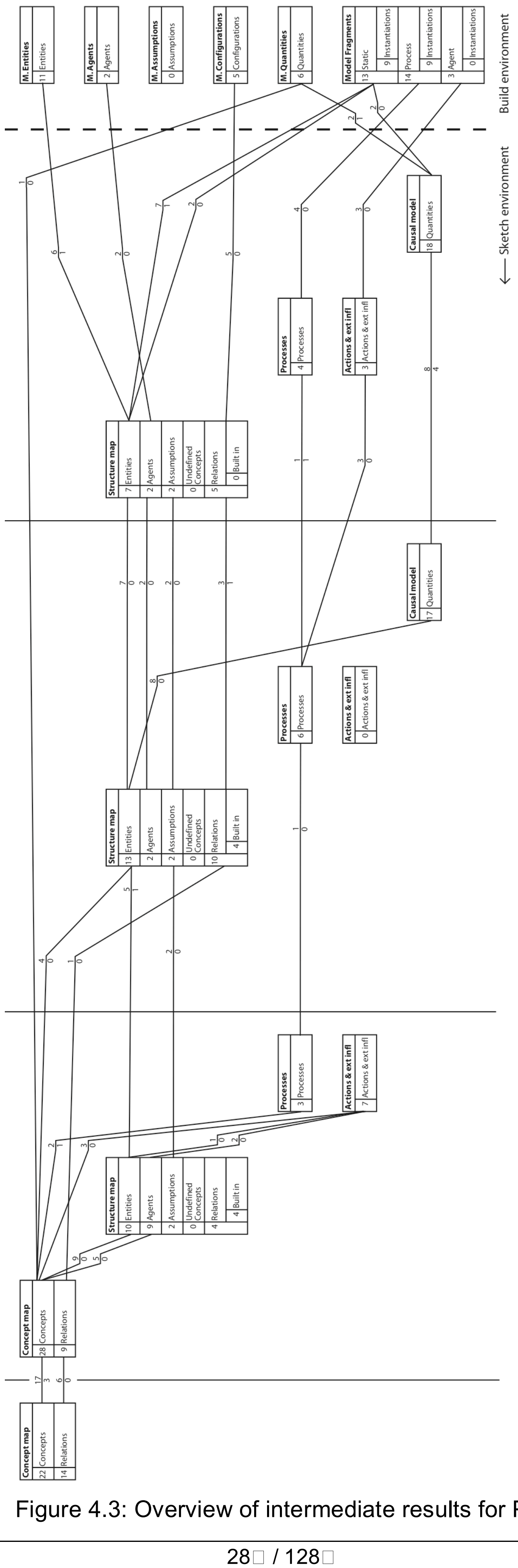


Figure 4.3: Overview of intermediate results for P3.

Participants / # Ingredients	P1	P2	P3
Total	164	214	245
Sketch	106	162	191
Build	58	52	54
Re-use	54	144	126
Re-use (%)	33%	67.2%	51.4%
Refinements	14	18	13
Refinements (%)	8.5%	8.5%	5.3%
Overall new	96	52	106
Overall new (%)	58.5%	24.3%	43.3%

Table 4.3: Overview of ingredients as processed by the participants.

4.3.2. Temporal division

The videos of the meetings were analysed for the types of activities and their duration. In total there where seven meetings, the first meeting was used for introductions and the final meeting was used for the final presentations. Five meetings were used for presentations and discussion. These five meetings were used for this temporal analysis. Each meeting lasted for approximately 105 minutes, the total time of the meetings was 550 minutes. In the first two meetings the ‘other’ students in the course used about half of the meeting time to present their work, so the total time available in the meetings for the participants was approximately 450 minutes. Table 4.4 presents an overview of the time spent explaining different subjects. Table 4.5 shows the time taken by the participants for explaining their progress. Table 4.6 shows the time taken for comments during the meetings. Table 4.7 shows the time taken for discussions during the meetings. All the temporal divisions are divided by subject. The most time was spent on the concept map, the causal model, the processes and modelling. The total time from the four categories is 427 minutes (some time was lost due to set-up of equipment and presentations).

Explanation subject	Total time (minutes)
Behaviour graph	00:30
Causal model	02:30
Concept map	03:00
Course	01:30
Domain	11:00
Model fragments	03:00
Modelling	18:30
Processes	24:00
Structural model	10:00
Total	74:00

Table 4.4: Total time and subjects for explanations during weekly meetings.

Subject	P1	P2	P3	Total
AMgIaGr	0:01:30	0:01:00	0:02:00	0:04:30
Behaviour graph	0:02:30			0:02:30
Causal model	0:09:00	0:00:30	0:01:30	0:11:00
Concept map	0:09:30	0:15:00	0:05:30	0:30:00
Entities			0:00:30	0:00:30
Model fragments	0:05:00	0:10:00	0:02:00	0:17:00
Processes	0:07:00	0:02:00		0:09:00
Scenario		0:01:30		0:01:30
Scope		0:01:30		0:01:30
Simulation	0:13:00	0:01:00	0:08:00	0:22:00
Structural model	0:04:00	0:03:30	0:01:30	0:09:00
Total	0:51:30	0:36:00	0:21:00	1:48:30

Table 4.5: Overview of participants speaking during meetings by subject for all three participants. AMgIaGr stands for Abstract, Model goals, Intended audience and General remarks.

Subject	P1	P2	P3	Total
A and EI		0:04:30		0:04:30
Behaviour graph			0:07:00	0:07:00
Causal model	0:11:00	0:06:30	0:07:30	0:25:00
Concept map	0:04:30	0:05:30	0:07:00	0:17:00
Model fragments		0:07:30		0:07:30
Model help	0:32:00	0:21:00	0:27:30	1:00:30
Processes		0:03:30		0:03:30
Structural model	0:06:30	0:11:30	0:03:30	0:21:30
Total	0:54:00	0:60:00	0:52:30	2:46:30

Table 4.6: Overview of comments during weekly meetings by subject for all three participants.

Subject	P1	P2	P3	Total
Approach		00:01:00		0:01:00
Causal model	00:04:00	00:17:00	00:03:00	0:24:00
Concept map		00:02:30	00:05:30	0:08:00
Domain		00:07:30	00:02:00	0:09:30
Modelling	00:04:00	00:14:30	00:07:00	0:25:30
Processes	00:04:00	00:02:00		0:06:00
Scope	00:04:30			0:04:30
Simulation			00:03:00	0:03:00
Structural model	00:05:30		00:01:30	0:07:00
Total	0:22:00	0:44:30	0:22:00	1:18:30

Table 4.7: Overview of discussions during weekly meetings by subject for all three participants.

4.3.3. Formalisation

The formalisation process is the process from the intermediate representations to a final model. To analyse this, the final (intermediate) Sketch representations were compared to the final model representations (in the Garp3 Build environment). Each of the formalisations from Figure 4.8 (see: ‘formalises into’) was measured quantitatively. For each formalisation the amount of ingredients re-used, refined and new were measured.

Figures 4.4, 4.5 and 4.6 contain the ratios between re-used, refined and new ingredients for the three participants. Figure 4.7 shows these ratios for the three participants combined. P1 has the lowest ratios and P3 the highest. Overall it shows that entities and quantities have a high re-use rate and that model fragments and behaviour have a low re-use rate. Configurations have a medium re-use rate. Agents, attributes, assumptions and scenarios have too few occurrences to give a clear view.

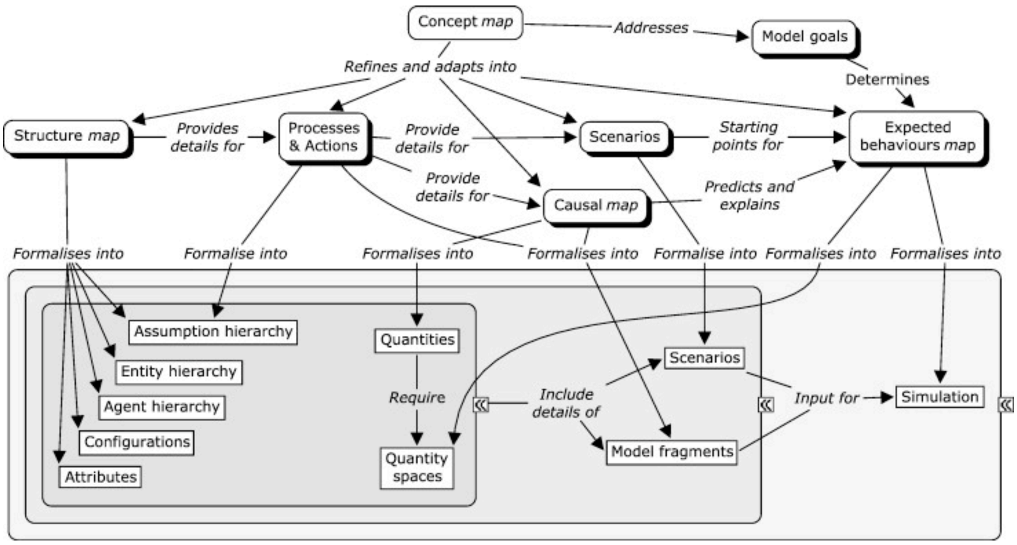


Figure 4.8: Overview of the intermediate representations used in the structured approach. Taken from [Bredeweg et al., 2007]

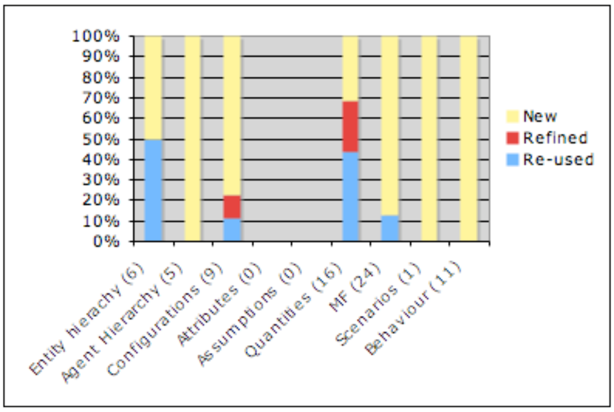


Figure 4.4: Formalisation by P1

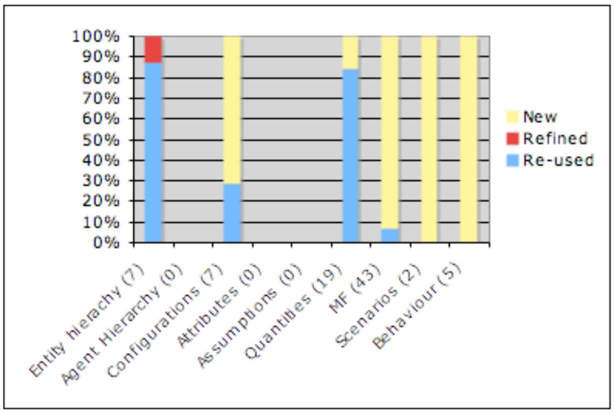


Figure 4.5: Formalisation by P2

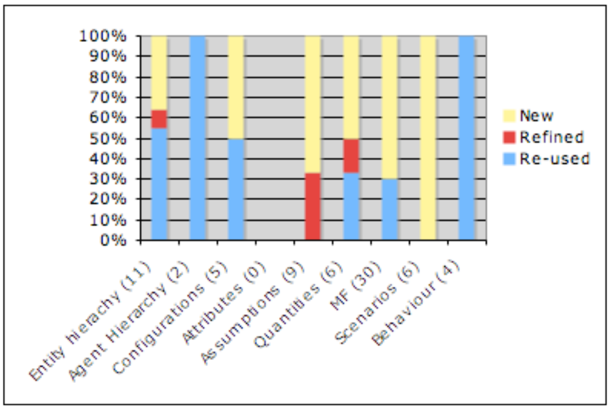


Figure 4.6: Formalisation by P3

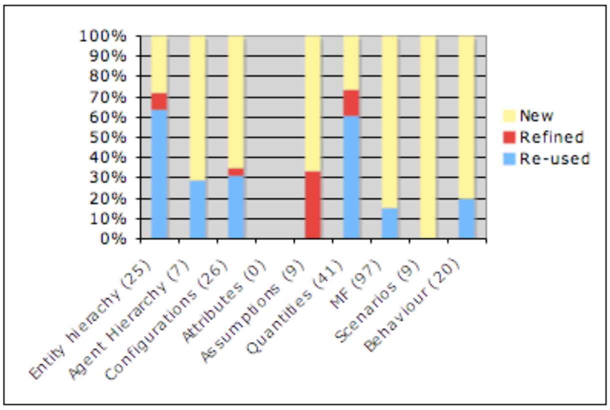


Figure 4.7: Formalisation for all participants.

4.4. Concluding remarks

The *concept map* did not prove to be a problem. The three participants did create different concept maps. The *structure model* provided some problems. Mainly because the participants did not clearly understand what needed to be represented in the structure model and what it's use is. Table 4.4 shows that the structure model needed the second most explanation of the sketch tools during the meetings. Figures 4.1, 4.2 and 4.3 show that the structure model was refined and adapted the most together with the processes and actions and external influences. This all suggests that the structural model may need more explanation than the other models in the sketch environment. The *processes* proved to be the most problematic of the sketch tools. It needed the most explanations during the meetings, all participants had to refine and adapt their processes. Since *actions and external influences* have a lot in common with the processes, the explanations on processes also applied on the actions and external influences. There were no further notable observations. During the meetings a lot of time was spend on the *causal models*, especially comments and discussion. This had to do with the modelling of the processes. The difficulties the participants had with the processes were reflected in the causal map. The participants found it cumbersome to create the *behaviour graph* in the sketch environment. The expected behaviour of P1 and P2 did not correspond to final model behaviour generated by their model. The expected behaviour graph of P3 corresponded completely with the behaviour of the model created. During the implementation of the models the participants had a lot of *modelling* problems. The problems the participants encountered were mostly 'model problems' and 'model errors' (that is, getting different simulation outputs then expected).

This study has shown that is possible to gain a deep conceptual understanding of a domain by using the structured approach to qualitative model building. Even though this understanding has not been tested explicitly, the domain experts that supported the students agreed that the participants had shown the acquisition of their knowledge of the domain through their work. The participants were graded on their final report and their participation. All three participants passed the course.

The participants in this experiment showed a different approach to model building, even within the structured approach. P1 re-used less of the model ingredients than P2 and P3. P1 also did not refine earlier sketch representations in later steps, whereas P2 and P3 did. This does not mean that P1 had a wrong approach or that the structured approach was not useful. The structured approach serves two purposes, namely to support the model building effort and to externalise intermediate steps. The structured approach supports the model building effort by 'forcing' different viewpoints on the modeller. In each step the modeller has to focus on a particular aspect of the model, the modeller must look at the problem from a certain angle. This helps the modeller in the model building process and through the steps the modelling problem is tackled. This applies to all the participants: by working through the steps different problems could be discussed and explained in a structured manner.

The second purpose, the externalisation of intermediate representations, is very useful to get insight in the modelling effort. When multiple modellers are collaborating in a model building effort intermediate representations can be very helpful in comparing and discussing the effort. Although the participants did not collaborate on the same model in this study, they did work with overlapping vocabulary. Furthermore the representations are also insightful for the experts who guide the modellers. When earlier representations are changed it reflects that the modeller has adapted his view of the model. In an educational situation, like the one in this study, it can help the participants and the experts to communicate and discuss the progress. As P1 did not refine earlier representations to reflect changes in approach, the changes had to be explained verbally more often. This caused some confusion during the meetings.

In conclusion it can be said that the structured approach and the sketch helped the participants in their modelling effort by providing focus on the important viewpoints. It also helped the communication and discussion of the effort through the intermediate representations. Although it is not essential to re-use and refine a lot of the intermediate ingredients for the individual, it can hinder communication.

5. Evaluation of Riacho Fundo model with stakeholders

5.1. Introduction

5.1.1. Evaluation of QR models

This section presents the results of evaluation activities held in Brasília with the purpose of validating the Riacho Fundo case study model (Salles, 2007). Validation and verification are very important aspects of model evaluation (Konikow and Bredehoeft, 1992; Oreskes *et al*., 1994; Rykiel, 1996; Sargent, 2004). Validation means that the scientific and conceptual contents of the model are acceptable for its intended use. Verification is a demonstration that the model is correctly implemented. Given that verification was done during the model building effort, this document does not discuss verification, only focusing on validation.

Following Rykiel (1996), before validation is undertaken it is important to:

- Define the purpose of the model
- Define the performance criteria
- Specify the model context

The purpose of the Riacho Fundo model is to support learning sustainability concepts and decision-making initiatives regarding water resources. The model is expected to improve understanding about how land use and occupation affect water quantity and quality, so as to stakeholders be able to foresee the consequences of the current situation as well as to assess the impacts of their decisions.

Criteria to validate the model include its capability to express relevant dynamic aspects of natural ecosystems and human activities and in order to explicitly represent cause-effect relations between land use and water resources. Accordingly, the model should support simulations that represent qualitatively distinct system behaviours, given a qualitative description of the system structure, as well as present concepts and simulation results in order to adequately convey the relevant scientific concepts.

The context in which the case study models were developed in NaturNet-Redime (NNR) project, and are meant to be used, are very diverse. For the Brazilian case study in the Riacho Fundo basin, the model is to be used by experts in water resources, schoolteachers, and students and water resources managers. In this context, the evaluation process should answer the following questions:

- Is the model 'in accordance' to the theories and available understanding of the functioning of similar systems?
- Is the model accessible for the user (vocabulary, visual representation, common sense ideas)?
- Does the model produce acceptable qualitative representations of the relevant phenomena?
- Are the concepts represented in the model justified by scientific knowledge? If not, can they be justified by educated commonsense?
- Can the model be used as a learning tool, in order to improve the learning of concepts expressed in the model?
- Finally, does the user recognize the potential of the model for representing real world problems and maybe for suggesting related solutions?

The work reported in this section follows a set of evaluation procedures as proposed in the literature following Rykiel (1996) and Sargent (2004). According to Rykiel (1996), the validation process can be decomposed in three components: conceptual, operational, and data validation. These components can be described as follows:

- *Conceptual* validity depends on providing a scientifically acceptable explanation of the cause-effect relationships included in the model. This can be achieved if it is possible to demonstrate that the theories and assumptions underlying the conceptual model are correct, or at least justifiable, and that the representation of the multiple aspects of the system structure are reasonable for the intended use of the model.
- *Operational* validity (whole model validation) demonstrates whether the model output meets the performance standards required for the model purpose. If the simulated system behaviour corresponds to observed behaviour of the real system, then the model is an adequate representation of the system. Statistical tests of comparison between simulated and real data are widely used to evaluate model behaviour.
- Rykiel (1996) points out that data used to develop and to evaluate the model should also be validated. In fact, *data validity* is a sensitive issue, as the relative inaccuracy and imprecision of ecological data also places limits on model testability. Given that we are not using numerical data, this component is not an issue in our work.

5.1.2. The Riacho Fundo model

Riacho Fundo is a small basin (225,48 km^2) located in Brasília, central Brazil. Since the new capital was built in the late 50's, it has been the most impacted area of the Paranoá Lake water basin. Most of the impacts are related to changes in land use, essentially consisting of transforming natural areas into rural and urban areas, leading to the current figures of circa 200.000 inhabitants. Due to the urbanization process, springs, streams, and natural vegetation are disappearing, and biodiversity is being reduced. Changes in the habitat also put severe pressure on many species, including the Riacho Fundo's largest mammal, the capybara (*Hydrochoerus hycrochaeris*). Details about the physical and biological environment of the Riacho Fundo basin can be found in the (Salles and Caldas, 2006).

According to stakeholders at the Riacho Fundo basin, the most relevant problems in that area are the following (cf. Salles, 2001):

- Uncontrolled land occupation
- Deforestation and destruction of natural habitats
- Problems with basic sanitation (including garbage and sewage deposition in the drainage network, open land and water bodies)
- Unsustainable practices by farmers and by the industrial sector
- Deficit in community participation, in part due to lack of knowledge about local degradation processes, environmental concern and mobilization

Accordingly, the Riacho Fundo model explored sustainability aspects from three viewpoints, namely the Urban, Semi-urban and Rural one (Salles 2007; Salles et al., 2007). The topics addressed in each of these perspectives are presented in the Table 5.1.

Table 5.1: Perspectives on the Riacho Fundo

	Perspectives		
Land use	Urban	Semi-urban	Rural
Main problems	Drainage system; flooded areas; transported garbage	Urbanization; infiltration; erosion	Erosion; biodiversity
Economic features	Services: garages; residues of petroleum products and garbage Economic damage caused by floods;	Industry: textile industry; food industry	Agriculture: cattle; crops
Soil	Impermeability of the soil	Soil particle aggregation	Soil fertility Manure production
Water resources	Urban drainage system: uncontrolled flow of water; drained water	Sediments and water bodies: effects of erosion and underground water on springs and rivers Water quality	Sediments and water bodies: effects of erosion and underground water on springs and streams Water quality
Biological entities	Mosquitos, Pathogens	Semi-urban vegetation	Rural vegetation; Vertebrates; Capybara
Human	Human well-being: garbage and water related diseases		
Agents	Rainfall	Urbanization	
Sustainability	Control of diseases; Control of residues	Water quality; Control of residues	Soil fertilization; Control of residues; Soil organic fertilization

5.1.3. Workshops and evaluation activities

Workshops with stakeholders in the Riacho Fundo basin were both dissemination and evaluation events. In all the meetings, initially a power point presentation explained the project and the main results obtained, particularly with respect to Redime case studies and the qualitative models produced. At the end, we asked who would like to participate in the evaluation activity. The volunteers were then, in a second meeting, confronted with a dedicated representation of the Riacho Fundo model and were asked to answer a questionnaire.

The following groups were participated in the evaluation: two groups of experts (academic and technicians from the Water and Sewage Company of Brasilia - CAESB), three groups of teachers, and one group of water managers who are involved in a Commission which is involved in the creation of the Rio Paranoá Water Basin Committee. This Committee, following Brazilian legislation for water resources, is a sort of parliament where representatives of the government, civil society and productive sectors discuss management of water related problems. This group was selected because the Riacho Fundo basin is comprised in the Rio Paranoá basin. Nine dissemination / evaluation events were held, being two of them with experts, six of them with teachers, and one of them with water resources managers. 31 questionnaires were obtained during these meetings: 1 from an academic expert, and 16 from the CAESB experts; 11 from teachers, and 3 from managers. The description of these meetings is presented below.

5.1.4. Questionnaires

The questionnaires addressed a number of aspects related to the modelling language, the simulator (Garp3), and the contents of the model, in its three perspectives (Rural, Semi-urban and Urban). They consisted of three types of questions:

- A statement followed by a scale of agreement, e.g. {*fully agree, agree, …, fully disagree*}
- A table with the name of the model ingredient or the type of the activity of a potential user of the model and the options, e.g. {*very easy, easy, …, very difficult*}, for each ingredient / type of user
- Open questions for the evaluator to present his / her opinion to specific problems, e.g. '*In your opinion what are the uses of the model?*'

A number of figures were included to show model fragments, scenarios, behaviour graphs, causal models and value history diagrams. In all the questions the evaluator could add comments or suggestions to their answers. In what follows, section 5.2 presents details of the workshops and of the evaluators. Section 5.3 presents the results obtained with the CAESB experts. Section 5.4 presents the results obtained with the water resources managers. Section 5.5 presents the results obtained with the teachers. Section 5.6 and 5.7 present the discussion and the conclusions.

5.2. Description of workshops and of the evaluators

In this section we describe the workshops and characterize the evaluators. In all the dissemination / evaluation events, the first activity was a power point presentation of the NNR project and the its most important results, mainly those related to the Redime case studies and qualitative reasoning models. After this introduction, the Riacho Fundo model was presented and discussed, both using slides and running simulations in Garp3. Considering that different groups of evaluators were involved, four types of questionnaires were produced. Finally, the appropriate questionnaire was given to those who volunteered to evaluate the model.

5.2.1. Experts in water resources

Two groups of experts were involved in the evaluation process. The first group consisted of two experts with academic degrees who were interviewed. One of them answered a complete questionnaire about the model. Both were experienced in planning and water management. One expert has a PhD degree and is a lecturer at the University of Tocantins. The other expert has a MSc degree and is a top level technician of the national Ministry of Environment and Water Resources, and was heavily involved in the preparation of the Brazilian National Plan for Water Resources. This meeting was held at the University of Brasília on 30/7/2007.

The second group consisted of 25 experts who work for the Brasilia Water and Sewage Company (CAESB). The session was held in premises of the Company in Brasília, on 26/9/2007. From this group, 16 questionnaires were obtained. The 16 technicians who answered the questionnaire can de characterized as follows: five operational analysts, one environmental analyst, two civil engineers, two biologists, and two operational technicians. Four of them had MSc degrees. Their age variation is shown in Table 5.2.

Table 5.2: Technicians from CAESB characteristics

Age	18-25 years	26-35 years	36-45 years	>46 years
Number (%)	3 (18,8%)	6 (37,5%)	5 (31,3%)	2 (12,5%)

Their professional experience was related to water quality (19%), GIS-based water quality and land use studies (13%), water bodies protection (13%), hydrology (6%), multi-criteria decision-making and reuse of water (6%), sustainability indicators (6%) and other activities. Four of them (25%) had no experience with water resources management.

The group had a diversified professional experience. Three worked with different applications of GIS (studies on land use and occupation, water quality and quantity, support for hydrological constructions, water resources management); two referred to work experience in governmental institutions related to the environment and management of water resources; two of them dealt with the protection of water bodies; and the remaining ones worked with multi-criteria analysis on decision-making and reuse of water; hydrological studies, management of water basins, qualitative analysis of phytoplankton, and water quality. Four respondents did not mention their professional experience.

The most comprehensive questionnaire was presented to the academic experts, three topics being addressed, namely the scientific knowledge involved in the model, concepts related to sustainability, and the value of the model as a representation of real-world problems. The questionnaire consists of 53 questions with the following distribution: general aspects of the model (2 questions), economic activities (5 questions), the Urban Riacho Fundo perspective (16 questions), the Semi-urban perspective (13 questions), the Rural perspective (14 questions), and final remarks (3 questions). For the CAESB experts, the questionnaire was a simplified version of the one presented to the academic experts, exploring the same types of expert knowledge, consisting of 34 questions. Here, the details of the questionnaires used for the CAESB experts are described. The answers given by the academic experts to their questionnaire, as well as the interview, are included in the discussion.

5.2.2. Water resources managers

A meeting was held with water managers in the governmental Administrative Center of the Riacho Fundo, on 22/08/2007. Initially, 8 people attended the meeting for the presentation of NNR project results. The group was quite heterogeneous, with respect to their professional background, including one student, civil servants, farmers, teachers, and people from NGOs. Eventually 3 of them volunteered to answer the questionnaire. The first participant, the leader of the group, was 49 years old, and works as a teacher and a farmer. As a farmer he was specialized in irrigated agriculture. He also was involved in a number of projects in the Riacho Fundo basin, both as a coordinator and as a member of the projects. He also had experience as a member of the local government, having worked as the Riacho Fundo administrator (‘mayor’) for 3 years. The second participant was a teacher, 39 years old. The third participant was 37 years old and currently the president of an NGO, which was actively involved in projects in the Riacho Fundo basin.

The questionnaire used for the water resources managers was the simplest one, consisting of 16 questions, two topics being addressed, namely the conceptual clarity, the use of the model as a tool to improve the understanding of the system, and to support argumentation and conflict resolution. The questions are presented below, along with the results of the questionnaires.

5.2.3. Secondary school teachers

The teachers were all working in state schools and had more than 10 years of professional experience. They were first presented with the NNR material in a power point exposition, and then they were presented with the evaluation activity and the questionnaires. To have an adequate environment for the evaluation activity, most of the teachers were brought to the Informatics Laboratory of Biological Sciences Institute at the University of Brasília (UnB). The teachers were separated in three groups and the following events were held:

- First group of teachers
    - First session – presentation of NNR project at the Centro Educacional 6 in Taguatinga, a satellite city located 30 km from Brasilia, on 25/07/2007. A group of 43 teachers attended this meeting, and 31 from them answered a simple questionnaire with their personal information and general opinion about the presentation. The evaluation activity was held at UnB on 08/08/2007, and was attended by only 1 teacher who answered the complete questionnaire
    - Second session – presentation of NNR project at UnB, on 06/08/2007, for 11 teachers. Among them, 7 answered the same simple questionnaire with their personal information and general opinion about the presentation. The evaluation activity with this group was held at UnB on 15/08/2007, and 5 teachers answered to the complete questionnaire.
- Second group of teachers
    - Presentation of NNR project and evaluation activity at the Centro Educacional Ave Branca in Taguatinga, on 29/08/2007. 8 teachers attended this meeting, and 2 answered the evaluation questionnaire.
- Third group of teachers
    - Presentation of the project and evaluation activity held in the Informatics Laboratory of the Biological Sciences Institute at UnB on 03/10/2007. 4 teachers attended the meeting and answered the questionnaire the same day. The teachers of this group are MSc students in the Post-graduate program in Science Education at UnB.

Altogether 11 teachers answered the questionnaire. Their profile is shown in Table 5.3.

Table 5.3: Profile of schoolteachers

Teachers	Subject	Teaching experience (years)	Graduate studies (MSc)	Age		
				26-35 years	36-45 years	>46 years
1	Biology	25	finished			X
2	Biology	no answer	no		X	
3	Biology	20	ongoing		X	
4	Physics	no answer	ongoing	X		
5	Biology	20	no		X	
6	Biology	15	no		X	
7	Chemistry	10	finished	X		
8	Physics	21	ongoing		X	
9	Mathematics	14	no		X	
10	History	17	no			X
11	Chemistry	17	ongoing		X	

Teachers were asked to give their opinion about educational aspects of the model, answering a questionnaire with 63 questions. These questions included the quality of the communication mediated by the model and the degree of difficulty of the modelling language for the students (4 questions); concepts about the Riacho Fundo basin from the Urban perspective (17 questions), the Semi-urban perspective (8 questions), and the Rural perspective (8 questions). Special attention was given to the development of mental competences and abilities, two central aspects currently addressed in the education system in Brazilian secondary schools (20 questions); the material conditions, and the human resources that prove to be necessary for the use of the software and the models in secondary schools (5 questions).

5.3. Riacho Fundo model: The expert opinion

5.3.1. Entities and configurations

The structure of a system consists of *entities* (shown in Figure 5.1) and *configurations*. In qualitative models they form the basis for the expression of behaviour. Experts were asked to evaluate entities and configurations for the whole model, and for each of the perspectives.

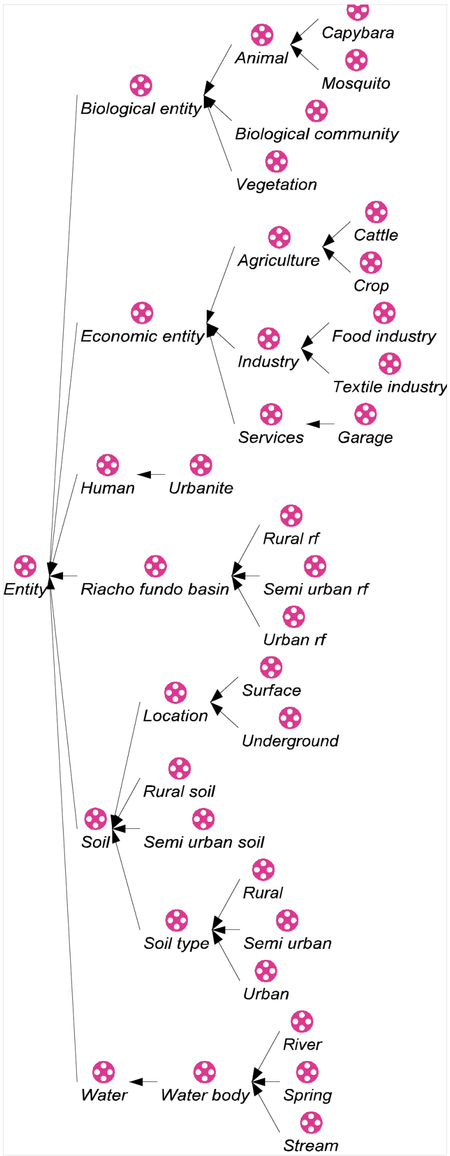


Figure 5.1: Entity hierarchy tree of the Riacho Fundo model.

The list of entities selected to represent the social, economic and environmental system in the Riacho Fundo basin was evaluated in terms of two parameters: clarity and correctness. As shown in the table below, 25% of the respondents answered they are clear and correct, and 75,0% considered them partially clear and partially correct:

"(Ex-G-01) Are the entities selected to represent the social, economic and environmental system in the Riacho Fundo basin clear and correct?"

Options	Yes	Partially	No	No answer
Number (%)	4 (25,0%)	12 (75,0%)	0	0

The reason why the respondents considered the list to be *'partially' correct* was that it did not include a number of items that they believed were necessary to represent the Riacho Fundo system – as can be inferred from their comments, below:

"Where are the humans that live in rural areas?"; *"What about the underground water?"; "Climate factors are missing."*; *"There is a need to represent water supply to public and private uses."*; *"Missing Organisms living in the water."*; *"There is an overlap of the entities 'animal', 'biological community' and 'vegetation'."*

However, it is not possible to include all the factors in the model. It is well-known that modelling always requires the selection of the most relevant parts of the system being modelled, according to the purposes of the model. The important point then is that none of the evaluators considered the representation to be unclear and/or incorrect.

A similar pattern was observed in answers to the question about the relevance of the selected entities to represent sustainability in the basin, as shown in the following table:

"(Ex-G-02) Are the selected entities enough to model the most relevant aspects of the sustainability in the Riacho Fundo basin?"

Options	Yes	Partially	No	No answer
Number (%)	6 (37,5%)	10 (62,5%)	0	0

Again, at least part of the reason why 62,5% of the respondents have chosen the parameter *'partially'* relevant can be attributed to the view that some items were missing in the representation, as can be inferred from their comments: *"Include changes in land use and occupation along time."*; *"It should include commerce among the services"*; *"Missing the cumulative effect of human actions."*

It is important to note that none of the evaluators considered the selection of entities not to be enough to represent sustainability.

Answers to the questions exploring the representation of entities and configurations (shown in Figures 5.2, 5.3, and 5.4) used on different viewpoints are shown in the following table:

Are the entities and configurations enough and relevant for a broad and clear representation of the [Urban/Semi-urban/ Rural] Riacho Fundo system?				
	Yes	Partially	No	No answer
Urban RF	5 (31,3%)	6 (37,5%)	1 (6,3%)	4 (25,0%)
Semi-urban RF	4 (25,0%)	10 (62,5%)	0	2 (12,5%)
Rural RF	4 (25,0%)	11 (68,8%)	0	1 (6,3%)
TOTAL	13 (27,1%)	27 (56,3%)	1 (2,1%)	7 (14,6%)

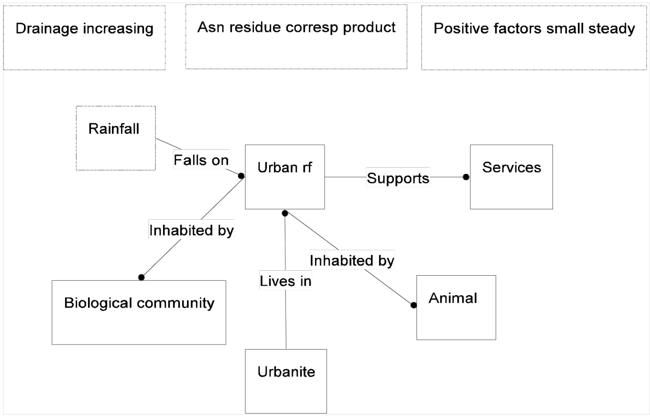


Figure 5.2: Entities and configuration – Urban RF

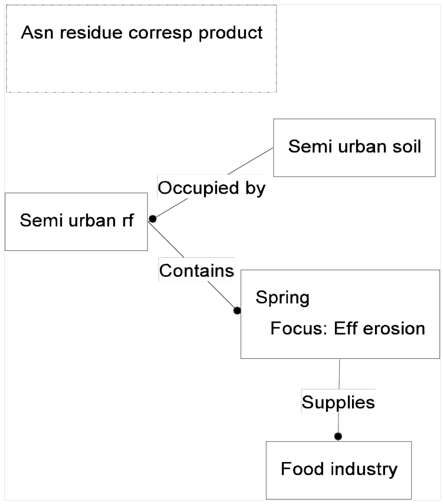


Figure 5.3: Entities and configuration – Semi-urban RF

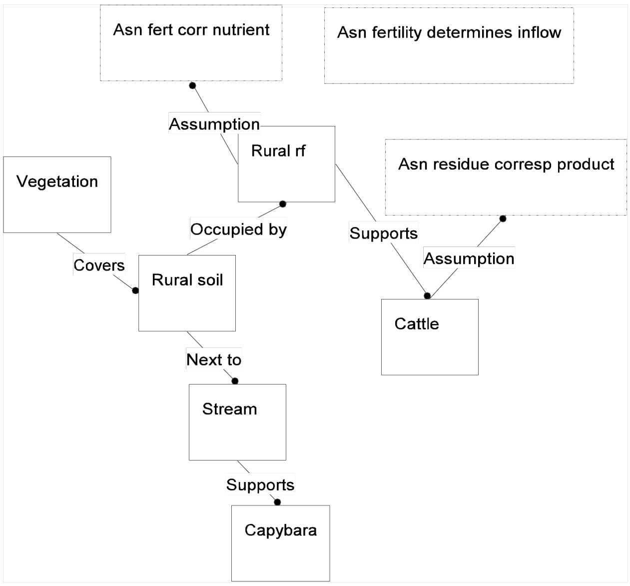


Figure 5.4: Entities and configuration – Rural RF

From the comments of the experts, it is clear that the most important limitation is that the list of entities is incomplete. Although one expert said that the answer was ‘yes’ *“considering the question the model was supposed to answer”*, all the remaining answers followed the previous pattern, as can be inferred from the comments of the respondents: *“Where are the humans that live in rural areas?”; “What about the underground water?”; “Climate factors are missing.”; “There is a need to represent water supply to public and private uses”; “Organisms living in the water are missing.”*

5.3.2. Quantities and their qualitative values

Quantities are expressions of relevant properties of the entities, and their qualitative values represent relevant qualitative states of the entities. Together, quantities and qualitative values are crucial to provide focus to the model and to create the vocabulary for reasoning about the system behaviour. The experts were asked about these issues in two questions, applied to each of the perspectives from which the Riacho Fundo basin was inspected – namely the urban, the semi-urban and the rural perspectives.

The answers are summarized in the following tables:

Do the selected quantities capture the most interesting properties of the entities used to represent the urban/ semi-urban/ rural Riacho Fundo system?				
	Yes	Partially	No	No answer
Urban RF	5 (31,3%)	7 (43,8%)	0	4 (18,8%)
Semi-urban RF	8 (50,0%)	5 (31,3%)	0	3 (18,8%)
Rural RF	10 (62,5%)	5 (31,3%)	0	1 (6,3%)
TOTAL	23 (47,9%)	17 (35,4%)	0	8 (16,7%)

Taking into consideration the comments, it can be pointed out that the answer is 'partially' due to the fact that the respondents pointed to the need of additional factors, as can be inferred from the answer of one of the respondents who believe that the model should include underground water problems in the urban perspective. Another expert suggested the inclusion of the quantity *quality of water* in the semi-urban part of the model.

Answering the question taking into consideration the rural perspective, some of the experts suggested to include other quantities. For example, types of crops, types of vegetation and water quality to the entity 'Spring'. One of the respondents commented *"maybe capybara is not the best indicator of habitat quality to represent the community of animals"*.

Do the qualitative values and quantity spaces of the quantities capture the most interesting qualitative states of the entities in the urban/semi-urban/rural Riacho Fundo system?				
	Yes	Partially	No	No answer
Urban RF	7 (43,8%)	3 (18,8%)	0	6 (37,5%)
Semi-urban RF	8 (50,0%)	3 (18,8%)	0	5 (31,3%)
Rural RF	9 (56,3%)	4 (25,0%)	0	3 (18,8%)
TOTAL	24 (50,0%)	10 (20,8%)	0	14 (29,2%)

With respect to the urban perspective, one of the experts pointed to the need for including underground water problems. Another suggested to include the quantity *quality of water* in the semi-urban part of the model. A third expert agreed with the selection of possible values, saying they were useful for a *qualitative* model of the rural perspective. There was also a suggestion that *"the quantity manure should include the value maximum"* to the rural perspective.

5.3.3. Simulations

Experts were asked to evaluate three simulations exploring the Urban (Figure 5.5), Semi-urban (Figure 5.6), and Rural perspective (Figure 5.7). In each case, the most complex scenario was chosen and the causal model and the values of the most relevant quantities in specific behaviour paths produced in the simulation were presented for evaluation. The answers are commented in what follows.

Based on the causal models shown in the Figures 5.5-5.7, the experts answered to the question: "Can the causal relations represented by (I's e P's) in the causal model of urban / semi-urban / rural RF be justified by scientific knowledge available?" As shown in the table below, the Semi-urban RF perspective causal model achieved the best score in terms of having causal influences justified by scientific knowledge available, and the Urban RF causal model was the one that received the lowest score in this item.

An overall of 35,4% of the expert opinions indicates that all the causal relations expressed in the three perspectives can be justified on the grounds of available scientific knowledge. Few experts add comments to this question. Referring to the Rural RF causal model, one of the respondents mentioned that *"some relations could only be explained in terms of repeated occurrence"* and not on scientific knowledge, and another mentioned that *"sediments should also affect water quality"*.

"Can the causal relations represented by (I's e P's) in the causal model of urban / semi-urban / rural RF be justified by scientific knowledge available?"				
	Yes, all of them	Some of them	No, none of them	No answer
Urban RF	4 (25,0%)	9 (56,3%)	0	3 (18,8%)
Semi-urban RF	8 (50,0%)	7 (43,8%)	0	1 (6,3%)
Rural RF	5 (31,5%)	8 (50,0%)	0	3 (18,8%)
TOTAL	17 (35,4%)	24 (50,0%)	0	7 (14,6%)

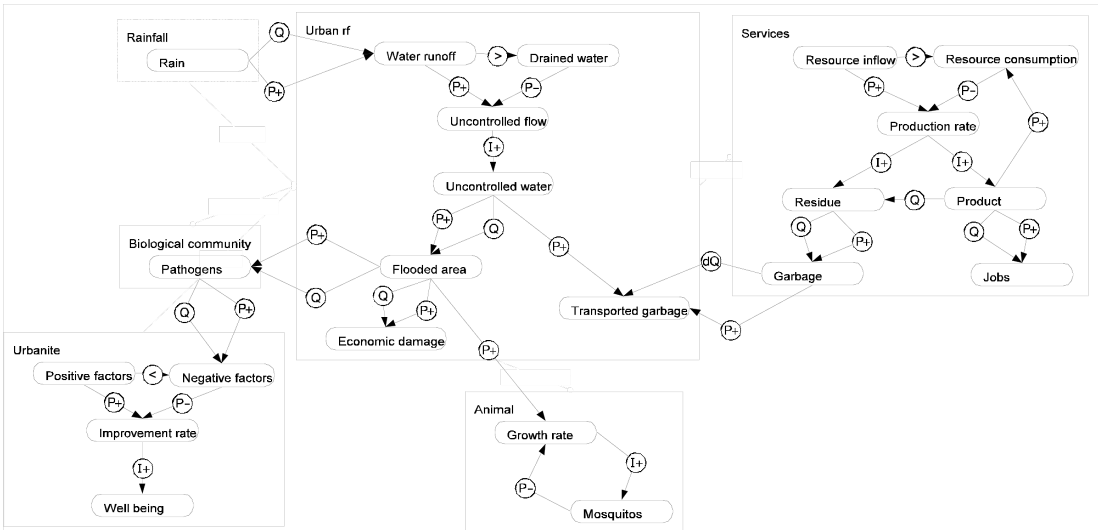


Figure 5.5: Causal model obtained in a simulation with a Urban RF scenario

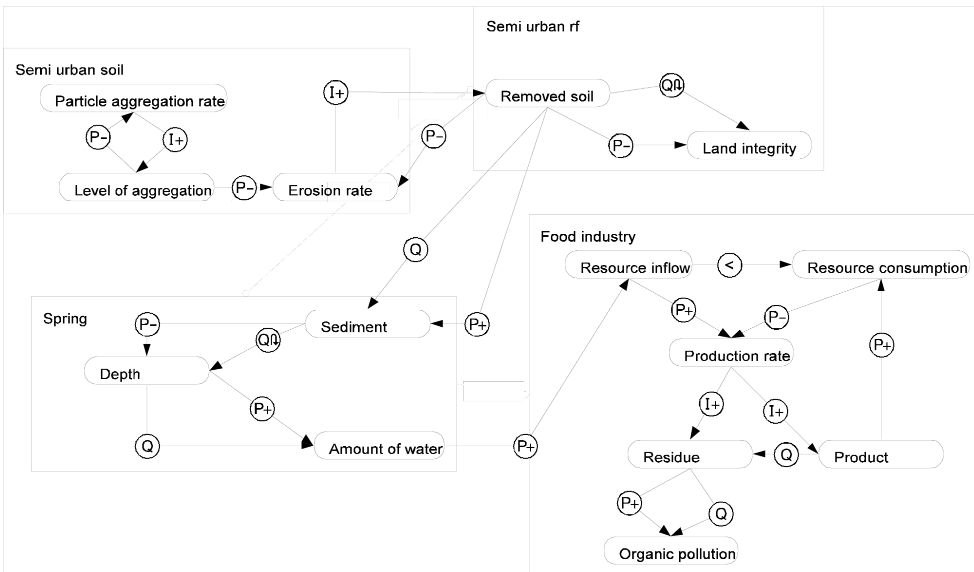


Figure 5.6: Causal model obtained in a simulation with a Semi-urban RF scenario

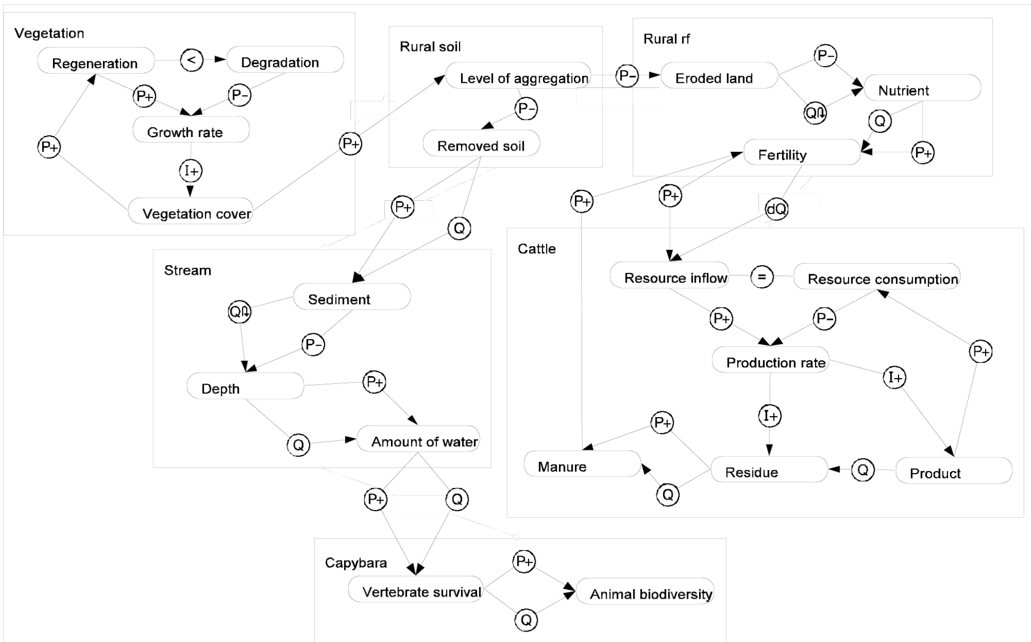


Figure 5.7: Causal model obtained in a simulation with a Rural RF scenario

In the next question the experts were asked on whether the commonsense knowledge could support the causal relations that cannot be justified by scientific knowledge. The answers are summarized in the following table:

Can causal relations (I's e P's), which have no scientific basis, be justified by commonsense knowledge?"				
	Yes, all of them	Some of them	None of them	Did not answer
Urban RF	6 (37,5%)	6 (37,5%)	0	4 (25,0%)
Semi-urban RF	5 (31,3%)	5 (31,3%)	1 (6,3%)	5 (31,3%)
Rural RF	5 (31,3%)	6 (37,5%)	0	5 (31,3%)
TOTAL	16 (33,3%)	17 (35,4%)	1 (2,1%)	14 (29,2%)

The table shows that apparently most of the causal influences that cannot be justified by scientific knowledge can indeed be supported by commonsense knowledge. Among those who added comments to their evaluation report to the Rural RF perspective, one of the respondents mentioned that *"good commonsense can create parameters for understanding phenomena"*, while another one added that *"logic deduction, along with commonsense, may explain certain causal relations used in the model"*.

The experts were presented with a series of value history diagrams in which the behaviour of the most relevant quantities in each simulation of the three perspectives was shown (Figure 5.8a-c). Some questions addressed the plausibility of the results of the simulation, and a final one in this set addressed the possibility of explaining the system behaviour in terms of the causal flow represented in the causal model as well as the values assumed by the quantities. The questions and the answers are given in what follows.

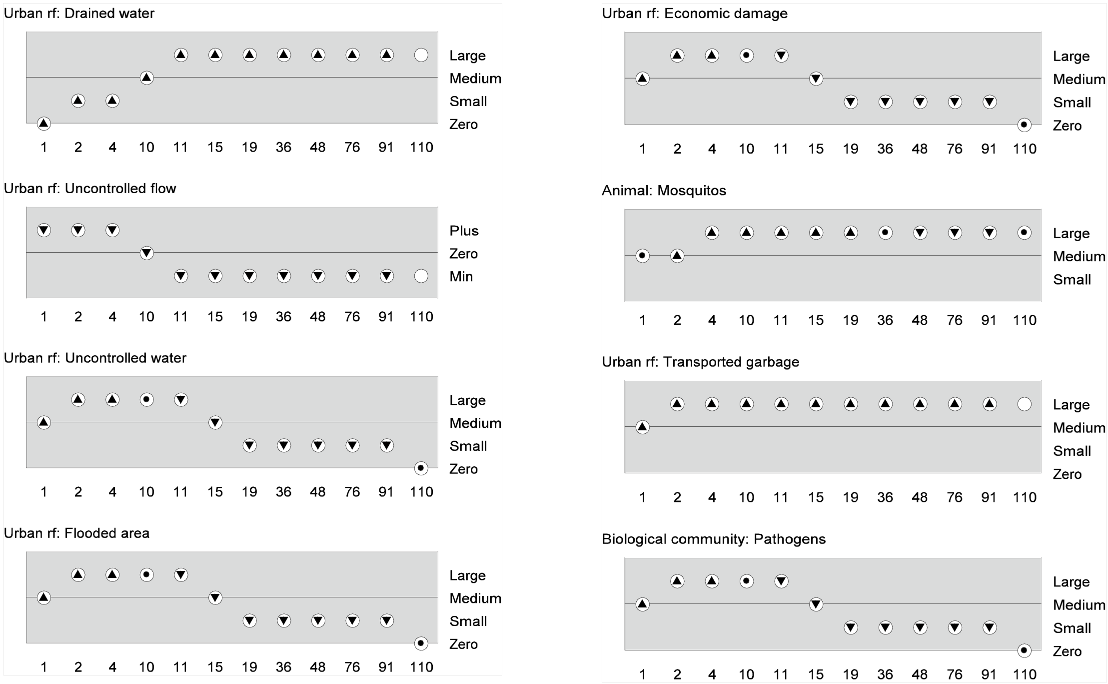


Figure 5.8a,b: Value history diagrams of a simulation with a Urban RF scenario.

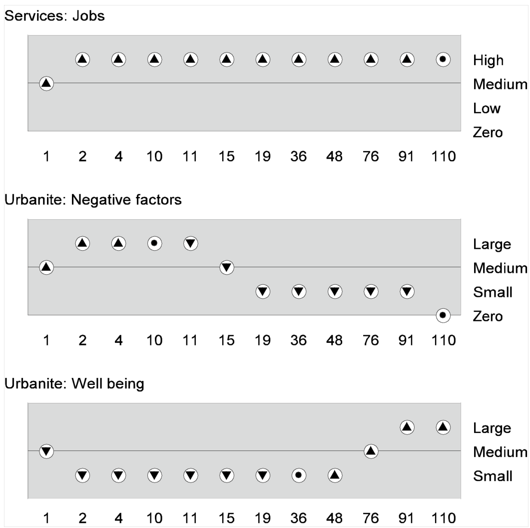


Figure 5.8c: Value history diagrams of a simulation with a Urban RF scenario

A behaviour path taken from the most complex simulation of the Urban RF perspective was presented to the evaluators, the questions and results being shown in the tables that follow:

(a) “(Ex-U-07) Does the behaviour path [1, 2, 4, 10, 11, 15, 18, 39, 50, 77, 9, 110] correspond to a behaviour shown by the system under natural conditions?”

Options	Yes	Maybe	No	Did not answer
Number (%)	2 (12,5%)	7 (43,8%)	1 (6,3%)	6 (37,5%)

(b) “(Ex-U-08) Are the values assumed by the quantities in each behaviour path correct?

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	3 (18,8%)	7 (43,8%)	0	6 (37,5%)

(c) “(Ex-U-09) Do these values represent values obtained in studies with the real system?

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	3 (18,8%)	7 (43,8%)	0	6 (37,5%)

(d) “(Ex-U-10) Is it possible to explain the system behaviour based on the causal flow represented in the causal model as well as in the values assumed by the quantities?”

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	7 (43,8%)	4 (25,0%)	0	5 (31,3%)

None of the experts added comments to these four questions.

The tables above about the simulation with a scenario of the Urban RF perspective show that 56,3% of the respondents considered the behaviour produced as a model output to be a behaviour that happens or may happen under natural conditions in the urban RF. In the answer to the question on whether the values of the quantities were correct, 62,6% considered the values ‘completely’ or ‘partially’ correct, and the same proportion considered these values to be ‘completely’ or ‘partially’ validated by studies with the real system. Finally, answering the question on whether it is possible to explain the system behaviour based on a combination of the causal model and the values of the quantities in the selected behaviour path, the answer ‘yes, completely’ and ‘partially’ were given by 68,8% of the respondents.

The same procedure was taken with respect to a behaviour path in a simulation with a scenario of the Semi-urban RF perspective. The value histories are shown in Figure 5.9a,b). The questions and results are shown in the following tables:

(a) “(Ex-S-07) Does the behaviour path [3, 5, 13, 14] correspond to a behaviour shown by the Semi-urban system under natural conditions?”

Options	Yes	Maybe	No	Did not answer
Number (%)	7 (43,8%)	4 (25,0%)	1	4 (25,0%)

(b) “(Ex-S-08) Are the values assumed by the quantities in each behaviour path correct?

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	4 (25,0%)	4 (25,0%)	0	8 (50,0%)

(c) “(Ex-S-09) Do these values represent values obtained in studies with the real system?

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	2 (12,5%)	6 (37,5%)	0	8 (50,0%)

None of the experts added comments to these three questions.

(d) “(Ex-S-10) Is it possible to explain the Semi-urban system behaviour based on the causal flow represented in the causal model and in the values assumed by the quantities?”

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	7 (43,7%)	4 (25,0%)	0	5 (31,3%)

The tables with the results concerning the simulation of the Semi-urban perspective as presented above show that 68,8% of the experts considered the behaviour produced as a model output to be a behaviour that happens or may happen under natural conditions in the semi-urbanized RF. Answering the question on whether the values of the quantities were correct, 50% considered the values ‘completely’ or ‘partially’ correct, and 75% considered these values to be completely or partially validated by studies with the real system. Finally, answering the question on whether it is possible to explain the system behaviour based on a combination of the causal model and the values of the quantities in the selected behaviour path, the answers ‘completely’ and ‘partially’ were given by 68,8% of the respondents. One of the experts provided the following comment: *“there are other variables that could interfere in the causal flow represented in the model”*.

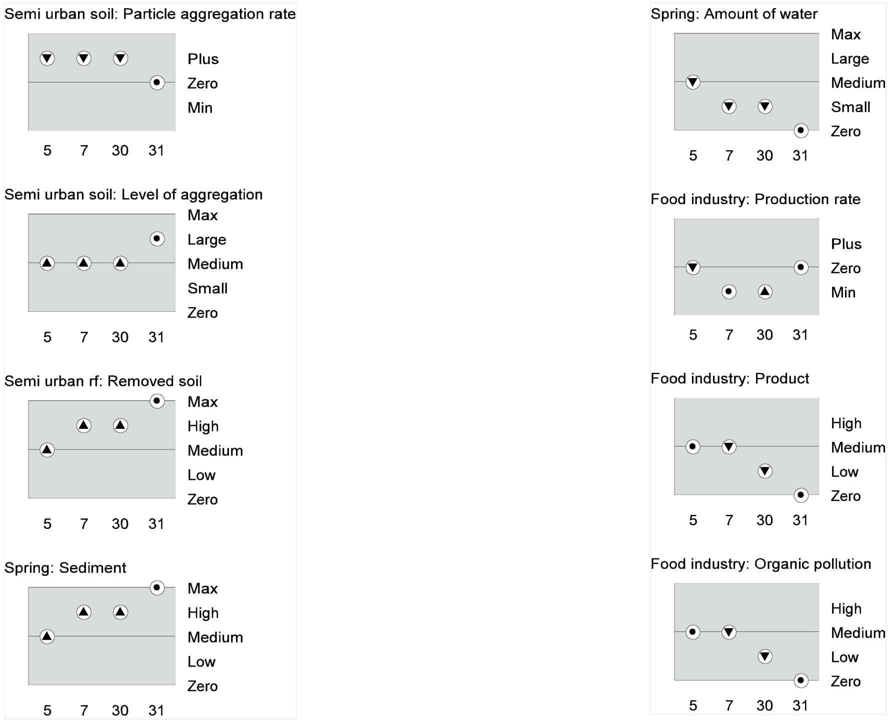


Figure 5.9a,b: Value history diagrams of a simulation with a Semi-urban RF scenario

The same procedure was taken with a behaviour path taken from the simulation with a complex scenario from the Rural RF perspective. The value histories are shown in Figure 10a-e. The questions and results are shown in the following tables:

(a) “(Ex-R-06) Do the behaviour paths [1] and [2, 6, 10, 25] correspond to behaviours shown by the system under natural conditions?”

Options	Yes	Maybe	No	Did not answer
Number (%)	8 (50,0%)	6 (37,5%)	1 (6,3%)	1 (6,3%)

(b) “(Ex-R-07) Does the behaviour path [3, 5, 13, 14] correspond to a behaviour shown by the system under natural conditions?”

Options	Yes	Maybe	No	Did not answer
Number (%)	5 (31,3%)	4 (25,0%)	4 (25,0%)	3 (18,3%)

Although 25,0% answered negatively to the possibility of observing this behaviour path under natural, no comments were added to the first two items.
(c) “(Ex-R-08) Are the values assumed by the quantities in each behaviour path correct?

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	6 (37,5%)	4 (25,0%)	0	6 (37,5%)

One of the experts added the following remark to the third item: *“given that we are supposing things in the model, nothing is completely correct or wrong”*.

(d) “(Ex-R-09) Do these values represent values obtained in studies with the real system?

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	2 (12,5%)	10 (62,5%)	0	4 (25,0%)

One of the experts made the comment that some of the results do not correspond to what happens in reality because *“the real system does not behave in such a simple way”*.

(e) “(Ex-R-10) Is it possible to explain the system behaviour based on the causal flow represented in the causal model and in the values assumed by the quantities?”

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	8 (50,0%)	4 (25,0%)	0	4 (25,0%)

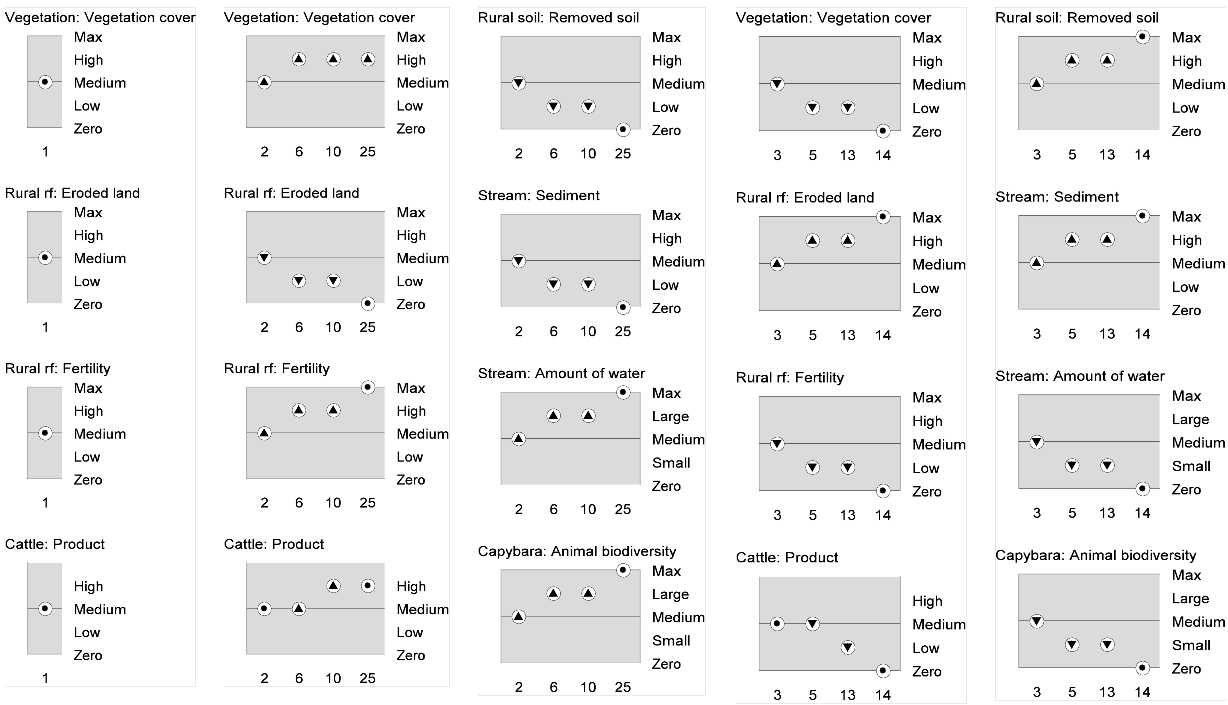


Figure 5.10a-e: Value history diagrams of a simulation with a Rural RF scenario

The behaviour produced as an output from a simulation with the Rural perspective model presented the best results of the whole set of questions: 87,5% of the experts considered that the behaviour produced by the model happens or may happen under natural conditions. Answering the question on whether the values of the quantities were correct, 62,5% of the experts considered the values completely’ or ‘partially’ correct, and

75,0% considered these values to be completely or partially validated by studies with the real system. Finally, answering the question on whether it is possible to explain the system behaviour based on a combination of the causal model and values of the quantities in the selected behaviour path, the option ‘completely’ and ‘partially’ was given by 75,0% of the respondents. One of the experts added the following comment: *“there are other variables that could interfere in the causal flow represented in the model”*.

5.3.4. Final remarks of the experts

The final part of the questionnaire for experts called for additional comments about the model in terms of the following parameters: ‘possible uses’, ‘possible users’ and ‘other remarks, if needed’.

(Ex-G-03) You are kindly asked to give your view on the possibilities of using qualitative models about sustainability in the Riacho Fundo for the following activities: teaching and training, research, natural resources management. The most representative answers are presented below.

Teaching and training
The opinion of the experts on the use of qualitative models in teaching activities is very positive, as can be inferred from their comments:

(a) “This is a highly relevant tool.”
(b) “(…) [it is] useful for the global assessment of a situation (involving the interaction of events), as an alternative to using compartmentalised knowledge.”
(c) “It can be used to develop the capacity of identifying cause and effect relations within a system.” (…) and “[t]o show the relations between different factors that are acting in a particular area.”
(d) “It could be used for teaching basic notions of Physics, Chemistry and biological sciences at the University.”
(e) “If simple models are used, learning can be faster.”

Some debate still exist on who can benefit from these models:

(f) “I still think this is a complex model to be used with the general public.”
(g) “Being qualitative, this type of model can be easier for teaching and communicating with a wide public. It can be a good tool for teaching and making predictions regarding the situation of Riacho Fundo.”

It is also suggested that the model can be useful for training activities:

(e) “I believe it is necessary to have knowledge about natural systems for interpreting the results generated by this kind of model, and this way I think it can be used for training managers.”
(i) “[it is] useful for teaching and training and for the development of new models.”

Research

(a) “Qualitative models represent an advancement, a change in the way of doing research.”
(b) “Qualitative models provide a global overview, making it possible to aggregate a number of non-measurable variables, and to delimit relevant factors.”
(c) “It represents the improvement of modelling approaches.”
(d) “In research, qualitative models will be important in the development of hypotheses to be tested.”
(e) “Useful for the development of theoretical concepts in mathematical models.”

(f) "This modelling approach is interesting because it can provide support for the development of quantitative models."
(g) "The improvement of the model and the representation of its quantities open a number of opportunities in different areas of research."
(h) "It is possible to improve the models by adding relevant factors that are required to develop specific researches" and *"to show with real facts the causes of a particular effect."*

Natural resources management

(a) "Excellent opportunities to use the model."
(b) "The model should be carefully used in the beginning. We should be careful with generalizations and answers to some problems that apparently can be used to all situations but in fact are not universally applicable. It is important to always point out that models have limitations and that, when dealing with natural systems, each situation has to be analyzed and we should include as many variables as we can in order to capture the reality."
(c) "It is useful to follow the development and the interactions between natural processes in the water basin."
(d) "It is useful for vulnerability analysis and assessment of water bodies, water bodies that receive effluents, and for water basin management."
(e) "It is useful to support decision-making about activities to be developed in the water basin."
(f) "It is useful for generating scenarios for regional planning."
(g) "It can be useful as many variables cannot be quantified, and with the model, inferences can be made about the system and the model works."
(h) "It is useful to predict solutions for future problems."
(i) "It would be interesting to combine with a quantitative model (at least in some relevant aspects)";
(j) "It can be applied in predictions and to reduce the burden of building mathematical models."
(k) "It is good for understanding the cause and effect relations involving different factors, and therefore allowing for changing one of them to get consequences in the whole system."

(Ex-G-04) In your opinion, what kind of users could benefit from this model?

(a) "Teachers, researchers, consultants."
(b)" Managers, students."
(c)"Technical staff involved with water resources management."
(d) "Teachers (secondary school and University) and professionals related to natural resources management."
(e) "Everyone related to studies on the improvement of quality of life."
(f) "At least for trained technicians who have finished secondary school."
(g) "Professionals, technicians, students, inspectors."
(h) "Environmental managers."
(j) "Architects, geographers, agronomists and those in charged of making decisions."
(k) "Users who have finished secondary school and have some knowledge in informatics could create simple models."
(l) "I believe models still represent an 'educational barrier' for the users. Although I cannot foresee an ideal model that could become a tool for the interpretation and the manipulation causal relations that could be used by the general public allowing them to become familiar with sustainability issues, I believe and praise the initiative of creating qualitative models for interpreting reality."

(Ex-G-05) Would you like to make any other comment?

(a) "This is a tool for planning."

(b) "This tool may be useful in the development of more complex models. In a way, it shortens the way to the development of complex models."

(c) "Training is required to use this tool."

5.4. Riacho Fundo model: The managers' opinion

It was difficult to undertake the evaluation event with the managers in the Riacho Fundo basin. After a number of unsuccessful attempts, 8 people with different background met at the Riacho Fundo local government building. The interaction with them was very positive, and they expressed interest both in the NNR project and in the use of qualitative models to support water resources management. Unfortunately, only 3 of them answered the questionnaire. However, these are valuable opinions to the project, and this is why they are included here.

5.4.1. Communication mediated by qualitative models in management

Communication is central for managers to interact and discuss the complex problems they have to decide upon. As we are investigating the potential for using qualitative models in water management, it is important to figure out the impression that the managers have about models and the modelling language.

After the presentation of the Riacho Fundo model, a table with the list of modelling elements was presented to managers with the following question: How difficult was it for *you* to understand each modelling element used in qualitative models? The answers are presented below:

Modelling element	Degree of Difficulty				
	Very easy	Easy	Medium	Difficult	Very difficult
Entities		2			
Entities and configurations		2			
Quantities (variables)		1	1		
Qualitative values	1	1			
Quantity spaces		2			
Model fragments		2			
Processes	1	1			
Direct influences (I+/I–)		2			
Qualitative proportionalities (P+/P–)	1	1			
Correspondences	1	1			
Situations described (qualitative states)		2			
Scenarios	1	1			
Simulations		1	1		
Total	5	19	2		

Only 2 respondents answered this part of the questionnaire. Their answers show that they found it easy to understand the modelling language.

Another question asked their opinion about how difficult it could be for people with different background, who probably would be involved in water management in a Water

Basin Committee, to understand the modelling language used in the qualitative models. Their answers are presented below:

Profiles	Degree of Difficulty				
	Very easy	Easy	Medium	Difficult	Very difficult
Farmers			1	1	
Businessmen	1	1			
Workers in the commerce		1	1		
Civil servants working for the local government	1	1			
Governmental technicians working with water resources	1	1			
Primary school students		1		1	
Secondary schools students	1		1		
Undergraduate students	1	1			
NGO members	1		1		
Housewives			1		1
Water basin committees members	1				
Total	7	6	5	2	1

As expected, according to the managers, people with less years of formal education, such as the farmers, housewives and those who have the primary school students are expected to have more difficulties with the modelling language. In general, for the others things should be easier.

5.4.2. The managers' view on the causal models

The managers were also exposed to causal models obtained in simulations with scenarios of Urban, Semi-urban and Rural perspectives of the Riacho Fundo model (see Figures 5.5, 5.6, 5.7) and questions explored clarity and correctness, and the possibility of creating explanations and making predictions based on the causal model. The results are shown in the table below:

Is the [Urban / Semi-urban / Rural] causal model conceptually clear and correct?				
	Yes	More or less	No	Did not answer
Urban RF	2	1	0	0
Semi-urban RF	2	1	0	0
Rural RF	2	1	0	0
TOTAL	6	3	0	0

The remarks about the causal model of the Urban perspective were in general positive (*"Yes, the diagram is very important for everybody's understanding."*; *"Yes, it is very clear as it is the reality of the Riacho Fundo."*). One of the managers answered *"More or less: it requires oral explanations."* One manager considered the Semi-urban causal model a good representation for the reality in Riacho Fundo: *"Yes, this is exactly what happens in our basin."* About the Rural perspective causal model a manager answered: *"More or less the diagram should be smaller."* Another one made a very interesting remark: *"I learned with the model that soil nutrients are reduced because of erosion."*

All the respondents were affirmative about creating explanations about the effects of the drainage system in urban areas of the basin, and two of them were affirmative about explaining the relation between erosion and industrial production: "*increasing erosion, the problems also increase and contribute to the decrease in the industrial production."* None of the respondents made comments about the Rural perspective causal model. The answers are given in the table below:

<table>
<tr><td colspan="5">Taking into account the relations of causality (I's and P's), could you create an explanation about…
[the improvement on human well being after the engineered drainage system was implemented? (Urban)]
[the relation between erosion and industrial production? (Semi-urban)]
[the relation between deforestation, animal biodiversity and agricultural production? (Rural)]</td></tr>
<tr><td></td><td>Yes</td><td>More or less</td><td>No</td><td>Did not answer</td></tr>
<tr><td>Urban RF</td><td>3</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>Semi-urban RF</td><td>2</td><td>0</td><td>0</td><td>1</td></tr>
<tr><td>Rural RF</td><td>0</td><td>0</td><td>0</td><td>3</td></tr>
<tr><td>TOTAL</td><td>5</td><td>0</td><td>0</td><td>4</td></tr>
</table>

All the respondents answered that they would be able to make predictions about mosquitos and pathogens after implementing the drainage system, based on the Urban perspective causal model (see table below). Two of them explained their answers: *"Yes, because in the Riacho Fundo few areas are covered by the drainage system, we have too many mosquitos."; "Yes, diseases transmitted by mosquitos decrease."* About the Semi-urban causal model, a manager said: *"Yes, the industrial production requires too much water"*, and about the Rural perspective, a manager justified his answer saying: *"Yes, water is missing for crop production."* Unfortunately the only person who answered 'no' to this question did not explained why (s)he could not make predictions based on the Semi-urban causal model.

<table>
<tr><td colspan="5">Taking into account the relations of causality (I's and P's), could you make predictions about …
[what happens to mosquitos and pathogens after the engineered drainage system was implemented? (Urban)]
[erosion, and the loss of springs and decrease in the industrial production? (Semi-urban)]
[the relation between deforestation, animal biodiversity and agricultural production? (Rural)]</td></tr>
<tr><td></td><td>Yes</td><td>More or less</td><td>No</td><td>Did not answer</td></tr>
<tr><td>Urban RF</td><td>3</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>Semi-urban RF</td><td>1</td><td>0</td><td>1</td><td>1</td></tr>
<tr><td>Rural RF</td><td>1</td><td>0</td><td>0</td><td>2</td></tr>
<tr><td>TOTAL</td><td>5</td><td>0</td><td>1</td><td>3</td></tr>
</table>

5.4.3. Qualitative models and the managerial activities

The next two questions explored the managerial experience, focusing on whether the managers had experience discussing the issues represented in the model, and on whether they would use a model similar to the one the had just seen. The answers were as follows:

Did you ever explore problems related to… [the drainage system in your activities on water resources in the Riacho Fundo? (Urban)] [erosion and industrial production in your activities on water resources in the Riacho Fundo? (Semi-urban)] [deforestation, animal biodiversity and agricultural production in your activities on water resources in the Riacho Fundo? (Rural)]				
	Yes, completely	Partially	No, almost nothing	Did not answer
Urban RF	1	1	1	0
Semi-urban RF	1	0	1	1
Rural RF	1	0	0	2
TOTAL	3	1	2	3

The answers to this question were well scattered among the options. Only one manager completely explored similar problems to those represented in the Urban causal model (*"Yes, many times. I was involved in a number of projects aiming at re-vegetation of riparian forests in the Riacho Fundo."*) and another one in the Rural causal model (*"Yes, many times, in projects of cleaning and re-vegetation of degraded areas."*).

Would you use this qualitative model in your discussions about … [flooding in the Riacho Fundo? (Urban)] [erosion and the loss of water resources in the Riacho Fundo? (Semi-urban)] [the relation between deforestation, animal biodiversity and agricultural production in the Riacho Fundo? (Rural)]				
	Yes	Maybe	No	Did not answer
Urban RF	2	1	0	0
Semi-urban RF	3	0	0	0
Rural RF	1	2	0	0
TOTAL	6	3	0	0

Finally, the managers were asked if they would use the qualitative model in their discussions about the problems represented in the causal model. The three respondents said 'yes' in discussions about erosion and the loss of water resources in the Riacho Fundo: *"Yes, it is very didactic."*; *"Yes, it is very good and easy to understand."*. Two out of three answered 'yes' in discussions about flooding in the Riacho Fundo: *"Yes, it is very useful to make things clear."*; *"Yes, mainly as an educational tool."*. The third manager put a condition for using it: *"Maybe, along with explanatory material."* About using the Rural perspective model, one manager said: *"Maybe, in the rural areas didactic activities should be simpler and practical."*

5.4.4. The potential for using qualitative models

The managers were asked to give their opinion about how qualitative models could support management of water resources. The options were ranked as follows: {fully agree, agree, maybe, disagree, fully disagree}. Below we present the answers, following respectively this sequence of options. The last number refers to the number of people who did not answer.

(a) Understanding how nature works = {0:2:0:0:0:1}
(b) Exploring the consequences of human actions = {1:1:0:0:0:1}
(c) Solving problems within the Water Basin Committee = {0:1:0:0:0:2}
(d) Building up arguments to support my ideas = {1:1:0:0:0:1}
(e) Planning water uses in the Riacho Fundo water basin = {0:2:0:0:0:1}

The conclusion is that, in spite of the reduced number of managers answering the questionnaire, the results support the conclusion that the participants had a very positive view about the use of qualitative models in water resources management.

5.5. Riacho Fundo model: The teachers' opinion

A very comprehensive questionnaire was presented to the teachers. A number of issues for educational actions were addressed, but the main focus was on the conceptual aspects of the modelling language, on the use of models and software in the classroom and on the potential of qualitative models to develop cognitive competences and abilities. Eleven teachers answered the questionnaire, in three evaluation events. The data collected with them is presented in the following section.

5.5.1. Communication mediated by qualitative models in educational contexts

Qualitative models, as those developed for the NNR project, are considered *conceptual models* as they contribute to improve the understanding of the structure and the behaviour of a system. In this evaluation effort, we asked the teachers to assess the conceptual aspects of the model, as this is a crucial part of their work. The first question asked for a general opinion about the use of qualitative models. Next, six questions explored the representation of concepts in model fragments, as these modelling elements should provide a summary of the main concepts. Due to lack of time, only model fragments from the Urban perspective were evaluated.

Before moving forward, teachers were asked about the degree of difficulty of communication using the modelling language. The answers are presented in two tables, as shown below. The first table contains their opinion about how difficult it is for the teacher to understand each modelling element:

Modelling element	Degree of difficulty				
	Very easy	Easy	Medium	Difficult	Very difficult
Entities	2	5	2	1	
Entities and configurations	1	3	7		
Quantities (variables)	2	5	4		
Qualitative values	2	5	3		
Quantity spaces		4	6		
Model fragments	1	5	5		
Processes		4	6		
Direct influences (I+/I–)		1	10		
Qualitative proportionalities (P+/P–)	1	5	5		
Correspondences		6	5		
Described situations (qualitative states)	2	3	6		
Scenarios	3	3	5		
Simulations	2	2	5		
Total	16	51	69	1	

It is interesting to note that none of the modelling elements was considered 'very difficult' and only one was 'difficult' to be understood by the teachers. The majority of the elements was considered 'easy' or 'medium', as confirmed by the following comments: *"In the beginning, when I first had contact with Garp3, it was difficult to understand the* *difference between scenarios, model fragments and processes."; "My difficulties were due to lack of familiarity with the meaning of the icons."*; *"As soon as we explore models already built and try to build our own models, we become familiar with the modelling elements."*

But, what about the students? The teachers' answers to the question "How difficult is it for *your student* to understand each modelling element used in qualitative models?" are shown in the following table:

Modelling element	Degree of difficulty				
	Very easy	Easy	Medium	Difficult	Very difficult
Entities	1	1	7	1	
Entities and configurations		1	5	3	
Quantities (variables)	2	3	4	1	
Qualitative values	2	3	4	2	
Quantity spaces	1	1	6	2	
Model fragments		6	2	3	
Processes			7	3	
Direct influences (I+/I–)		1	3	6	
Qualitative proportionalities (P+/P–)		1	3	7	
Correspondences		1	7	3	
Described situations (qualitative states)		4	3	3	
Scenarios		3	5	2	
Simulations		2	6	2	
Total	6	27	62	38	

The table shows that the overall opinion is that the modelling elements are considered to be 'medium / difficult' (75% of the answers). Again, none of the elements was considered 'very difficult'. Some teachers are optimist, as can be inferred from the following comments: *"Handling the models will become easier for the students when they play with the models and with the symbols used for expressing the concepts."*; *"If the concepts are worked out beforehand, the use of the models will be easier."* One of the teachers pointed out difficulties: *"We could build models using a simplest language than this one. Garp3 should be translated into Portuguese."* Another teacher noted that there exist differences between students: *"It is difficult to say whether these elements are difficult or easy for the students. We have students that are able to cope with scientific concepts, while others barely cope with them."*

In the final part of the questionnaire, the teachers were asked to express their opinion about the success of scientific education mediated by qualitative models, under the assumption that the 'language used to represent concepts in qualitative models' is already known to the students.

(Te-G.04) Suppose that the students are already familiar with the modelling language used to represent concepts in qualitative models. In your opinion what are the chances of success of scientific education mediated by qualitative models?

Option	Very small	Small	Medium	Big	Very big
Number (%)	0	0	3 (27,2%)	6 (54,6%)	2 (18,2%)

The possibility of success is considered ‘very big’ or big’ for 72,8% of the teachers, and none of them select the options ‘small’ or ‘very small’. Additional comments were the following: *“The language may become the biggest barrier for using qualitative models in secondary schools.”*; *“The language used in the models should be the same as the one used in classroom, otherwise the students will be less motivated.”*

5.5.2. Representing concepts

Model fragments (MF) are ways of expressing complex concepts, by combining different modelling elements. After all, a qualitative model is composed by model fragments, selected by the software according to the structural description of the system. Similarly, scenarios are made of a number of modelling elements that describe a situation in which the system can be found, which in turn may evolve into typical behaviours that the system can assume. Together, these modelling elements provide a summary of most of the knowledge encoded in a qualitative model. Six model fragments and one scenario were explored by the teachers in order to assess their potential for representing concepts, as described in this section.

Initially, the teachers were asked to answer the following question:

(Te-G.01) What is your opinion about the possibility to express concepts by means of qualitative models?

Option	Very small	Small	Medium	Big	Very big
Number (%)	0	0	2 (18,2%)	7 (63,7%)	2 (18,2%)

Some of the comments are useful to understand their opinion: *“Science education usually explores ideas based on models.”*; *“The possibility of expressing concepts with qualitative models is very big, provided they are well explored by the teacher.”*; *“The visual approach make the models easier to understand.”*

The following model fragments were evaluated: ‘Precipitation’ (Figure 5.11); ‘Control of drained water by drainage system’ (Figure 5.12); ‘Transported garbage’ (Figure 5.13); ‘Mosquito population growth’; ‘Water related diseases’; ‘Human well being improvement’. Questions with similar structure were used to evaluate the model fragments, as illustrated by the following question: “Is the model fragment that describes precipitation conceptually clear and correct?”

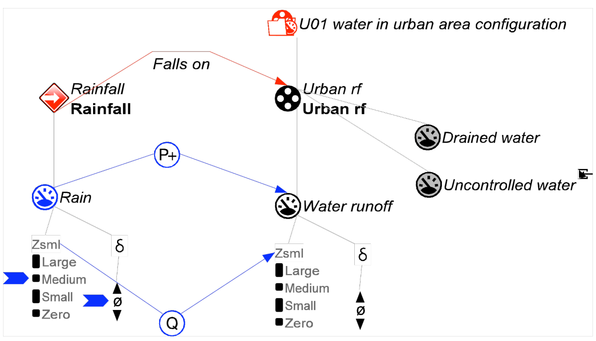


Figure 5.11: Model Fragment Agent ‘Precipitation’

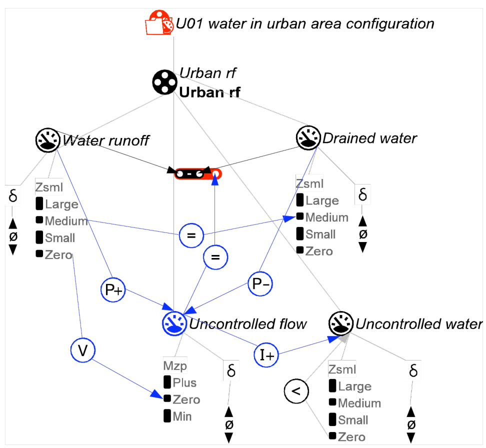


Figure 5.12: Model Fragment ‘Control of drained water by drainage system’

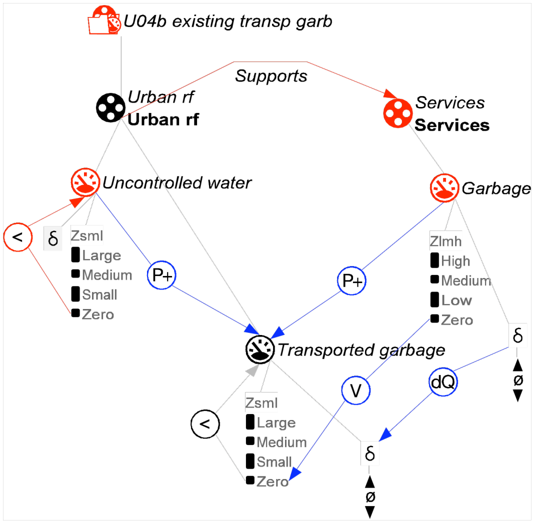


Figure 5.13: Model Fragment 'Transported garbage'

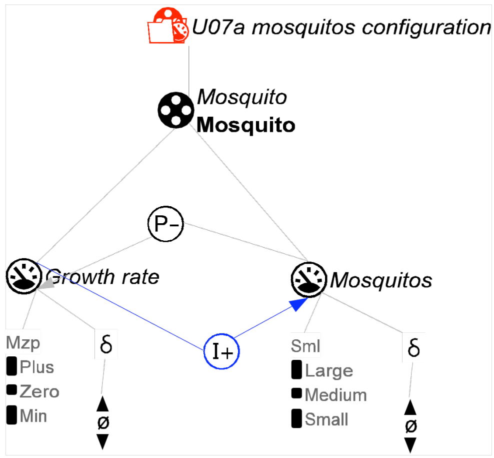


Figure 5.14: Model Fragment 'Mosquito population growth'

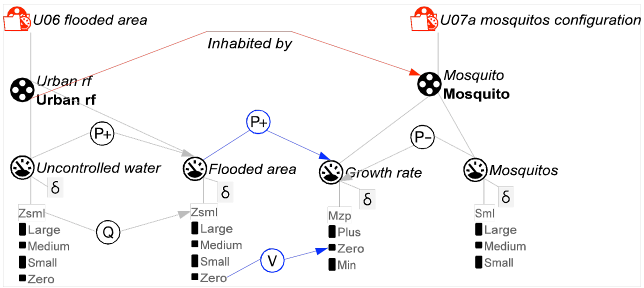


Figure 15. Model Fragment 'Flooded areas and mosquitos'

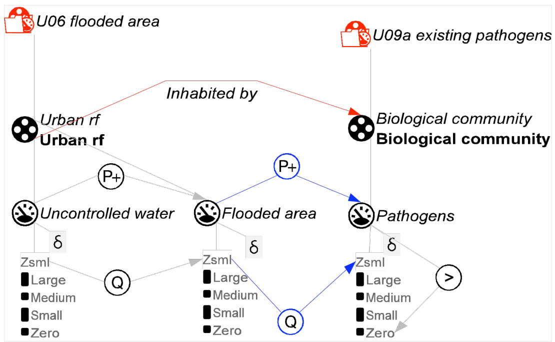


Figure 16. Model Fragment 'Water related diseases'

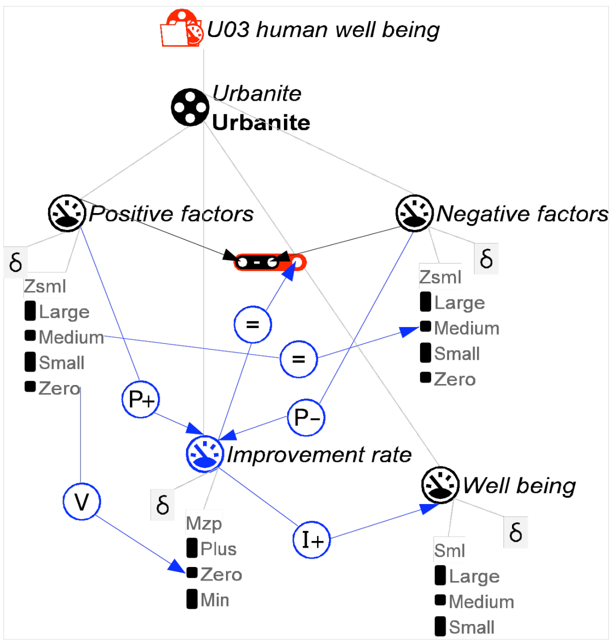


Figure 5.17: Model Fragment 'Well being improvement'

The answers are presented in the table below:

Is the model fragment that describes […] conceptually clear and correct?				
	Yes	More or less	No	No answer
MF ‘Precipitation’	6 (54,6%)	5 (45,5%)	0	0
MF ‘Control of the water runoff by drainage system’	5 (45,5%)	6 (54,6%)	0	0
MF ‘Uncontrolled water transported garbage’	5 (45,5%)	5 (45,5%)	0	1 (9,1%)
MF ‘Uncontrolled water and mosquito population growth’	5 (45,5%)	4 (36,4%)	1 (9,1%)	1 (9,1%)
MF ’Water related diseases’	7 (63,7%)	3 (27,3%)	0	1 (9,1%)
MF ‘Positive and negative factors that determine human well being’	3 (27,3%)	7 (63,7%)	0	1 (9,1%)
Total	31 (46,5%)	30 (45,0%)	1 (1,5%)	4 (6,0%)

The overall impression of the teachers was positive as 46,5% of the answers considered the model fragments conceptually clear and correct, and 45,0% considered them ‘more or less’ clear and correct. The best score was achieved by the MF ‘Water related diseases’ (63,7% of ‘yes’) and the worse score was given to the MF ‘Positive and negative factors that determine human well being’ (63,7% of ‘more or less’). Only one teacher considered a model fragment (MF) unclear and incorrect, namely the so-called MF ‘Uncontrolled water and mosquito population growth’.

About the MF ‘Precipitation’ some teachers presented the following comments: *“It is possible to draw some conclusions by inspecting this model fragment.”* “*The relations expressed in the model fragment correspond to what happens in reality.”*

Although the MF ‘Control of the water runoff by drainage system’ received more intermediate marks than the other MFs. One of the teachers considered it conceptually clear and correct adding the following comment: “ (…) because it describes how the water reaching the soil is divided further defining the proportion of water that constitutes the drained flux.”

Some comments may account for the choice of the option ‘partially yes’ and ‘no’, as in the evaluation of the MF ‘Uncontrolled water and mosquito population growth’: *“The relation between ‘growth rate’ and ‘mosquitos’ (I+ and P–) is not very clear.”* One of the teachers suggested changes in this model fragment: *“The model fragment should specify ‘population growth rate’, because the expression ‘growth rate’ may be understood as the growth of insect size.”*

Finally, some teachers expressed their worries about students dealing with the model fragments. For example, answering the question on whether the representation of the MF ‘Water related diseases’ was correct and clear, the option ‘more or less’ was chosen, with the following comment: *“More or less; the relations are correct, but I imagine the secondary school students will not be able to see them given the high number of relations expressed in this model fragment.”*.

About the MF ‘Positive and negative factors that determine human well being’, the option ‘more or less’ was found followed by the following comment: *“I found the concept of ‘Improvement rate’ not very clear. I think it should be replaced by another one, clearer than this one. Also, I believe that, in order to use the model in the classroom, it is* *necessary to be specific about which positive or negative factor is being focused. For the students the current terminology is far too general."*

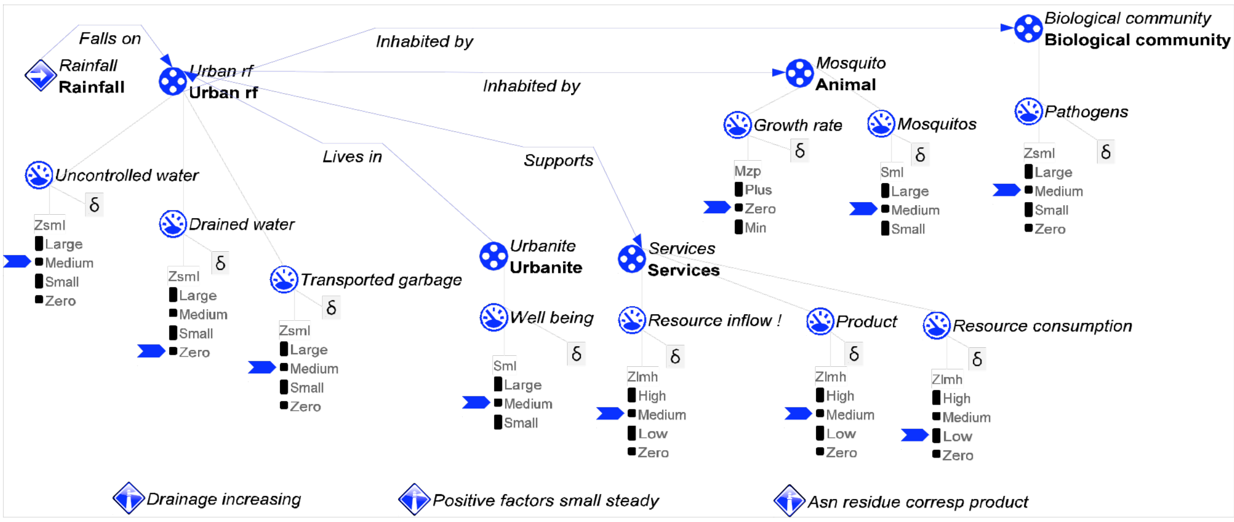


Figure 5.18. A scenario of the Urban RF perspective

A scenario (see Figure 5.18) for the more complex simulation of Urban RF perspective was conceptually evaluated by the teachers as follows:

(Te-U-07) Is the scenario above conceptually clear and correct?

Options	Yes	More or less	No	Did not answer
Number (%)	5 (45,5%)	4 (36,3%)	1 (9,1%)	1 (9,1%)

Some teachers justified their answers as follows: *"Yes, although it is a big model, it is clear and objective."*; *"More or less. We could arrive to different conclusions."*; *"More or less. The relations are not very clear."* One teacher has chosen the option 'no', but from his/her comment it is clear that (s)he did not understand the difference between scenario and model fragment: *"No. I believe the qualitative proportionalities are missing."*

5.5.3. Simulations and causal models

Teachers were asked to evaluate a simulation exploring the Urban RF perspective. The scenario was evaluated in the previous section. The causal model and the values of the most relevant quantities in specific behaviour paths were presented for evaluation (Figures 5.3, 5.6 and 5.9). The answers are commented as follows.

The teachers were presented with three causal models, obtained in simulations of the Urban, Semi-urban and Rural Riacho Fundo perspectives (see Figures 5.5, 5.6, 5.7). The majority of the teachers agreed that the causal models were clear and correct, as shown in the table below.

Is the [Urban / Semi-urban / Rural] causal model conceptually clear and correct?				
	Yes	More or less	No	Did not answer
Urban RF	9 (81,9%)	2 (18,1%)	0	0
Semi-urban RF	10 (91,0%)	1 (9,1%)	0	0
Rural RF	7 (63,7%)	3 (27,3%)	0	1 (9,1%)
TOTAL	26 (78,8%)	6 (18,2%)	0	1 (3,0%)

Some of the comments of the teachers are presented below.

Urban RF perspective: *"Yes, the model gives us an overview of cause-effect relations and shows how we could add more information to the model."*; *"Having a previous good training on the concepts represented by I's, P's and Q's, it is easy to read the diagram."*; *"Yes. I don't agree with some of the P– included in the model (between 'resource consumption' and 'production rate'; 'mosquitos' and 'growth rate'), but I know that this relation may be in accordance to what the author wants to show."; "More or less. I have some doubts, for example, about negative factors being stronger than the positive ones, and about the I+ as an influence on 'well being'."*

Semi-urban RF: *"More or less. It still confuses me the use of P+, P – and I+ and other modelling elements."*; *"Yes. Soil aggregation rate starts changes in the system that affects the spring and the food industry. If soil aggregation rate is positive, the system has a decrease in erosion rate and land integrity will increase. This quantity is inversely proportional to the amount of removed soil and thus sediments on the water decrease, increasing the amount of water in the spring, further increasing resources offer and influencing the food production, which in turn increases the residues production and the organic pollution. Alternatively, there will be an increase in the consumption of resources, causing the food production rate to decrease, as well as residues and organic pollution."*; *"Yes, it shows all the possibilities."*; *"Yes, observing the quantities behaviour it is possible to draw this conclusion."*

Rural RF: *"More or less. As the sediment increases, depth decreases. How can amount of water increase?"*; *"Yes. If vegetation regeneration is smaller than its degradation, then vegetation growth is also small and this affects vegetation cover, which will influence soil features that will affect the stream, agricultural production and animal biodiversity."*; *"Yes, the relations are well defined."*; *"More or less. Vegetation regeneration is smaller than degradation. Therefore growth rate is negative."*

Next, the teachers were asked if they could use the causal model to create explanations and to make predictions about the Urban, Semi-urban and Rural Riacho Fundo systems. As shown in the table below, all the answers ranged from 'more or less' to 'yes', and the majority admitted they could use the causal models to support explanations and predictions.

Taking into account the relations of causality (I's and P's), could you create an explanation about… [the improvement on human well being after the engineered drainage system was implemented? (Urban)] [the relation between erosion and industrial production? (Semi-urban)] [the relation between deforestation, animal biodiversity and agricultural production? (Rural)]				
	Yes	More or less	No	Did not answer
Urban RF	7 (63,7%)	4 (36,3%)	0	0
Semi-urban RF	9 (82,9%)	2 (18,2%)	0	0
Rural RF	8 (72,8%)	2 (18,2%)	0	1 (9,1%)
TOTAL	24 (72,7%)	8 (24,3%)	0	1 (3,0%)

Some of the comments of the teachers are presented below.

Urban RF: *"Yes. Using the model, I can follow the path from 'drained water' until human 'well being'."*; *"Having a drainage system, there are no flooded areas, pathogens are controlled and these things contribute to the well being."*; *"Partially. If water is drained, the 'uncontrolled flux', which influences the 'uncontrolled water', will decrease, reducing 'flooded areas' and 'transported garbage'. This will be a negative influence on mosquito population growth."*

Semi-urban RF: *"Partially. I am still unable to understand the interactions between industrial production, erosion, pollution etc."*; *"Partially. If the quantity soil removed increases, land integrity destabilize, negatively influencing the industrial production."*; *"Yes. Industrial production will lead to increasing urbanization growth, and this will affect the environment (erosion)."*

Rural RF: *"Partially. If vegetation degradation increases, there is erosion, reducing nutrients and fertility, and this will affect agricultural production and animal biodiversity."*; *"Yes. After the causal model analysis, it is possible to clearly see the relations among the quantities, and this will contribute to the formulation of concrete proposals on sustainable development for the Riacho Fundo basin."*; *"Yes. Deforestation decreases and animal biodiversity and agricultural production increases."*

Taking into account the relations of causality (I's and P's), could you make predictions about … [what happens to mosquitos and pathogens after the engineered drainage system was implemented? (Urban)] [erosion, the loss of springs and decrease in the industrial production? (Semi-urban)] [the relation between deforestation, animal biodiversity and agricultural production? (Rural)]				
	Yes	More or less	No	Did not answer
Urban RF	8 (72,8%)	3 (18,2%)	0	0
Semi-urban RF	10 (91,0%)	1 (9,1%)	0	0
Rural RF	7 (63,7%)	3 (27,3%)	0	1 (9,1%)
TOTAL	25 (75,8%)	7 (21,2%)	0	1 (3,0%)

Some of the comments of the teachers are provided below:

Urban RF: *"Yes, (…) because increasing the urban drainage capacity it is possible to reduce negative factors and increase the well being."*; *"Yes. After the simulations, we can inspect the behaviour of the natural system."*; *"Partially. If water is drained, uncontrolled flow decreases, and so do flooded areas and transported garbage, which is a negative influence on mosquitos population growth."*

Semi-urban RF: *"Yes. Industrial production may increase or decrease according to the erosion rate."*; *"Yes. If there is erosion, it will affect industrial production."*; *"Yes. With the industrial production, the soil degrades causing erosion."*

Rural RF: *"Yes. If deforestation is large, biodiversity decreases. If deforestation decreases, biodiversity increases. Agricultural production, in turn, may increase or decrease according to resource inflow. When this quantity increases due to increase in fertility, production rate increases. When resource inflow decreases because fertility has decreased, production rate also decreases."*; *"Partially. Land without vegetation cover = removed soil = sediments in the rivers = less water for animal survival."*; *"Yes. If the* *farmer does not provide nutrients to the soil, it will not be possible to increase agricultural production."*

5.5.4. System behaviour

Taking the value history diagrams for relevant quantities in the simulation with a scenario from the Urban RF perspective (see Figure 8 a-c), the conceptual representation of the Rural RF system behaviour was evaluated by the teachers in four questions, as follows. Note that, although the sequence of states is slightly different when comparing the numbers in Figures 8 a-c ([1, 2, 4, 10, 11, 15, 19, 36, 48, 76, 91, 110]) and those in the questions below, the behaviour path is exactly the same in both circumstances. The difference is due to the English version of the model (from where the figures were taken) and the Portuguese version (used to prepare the questionnaire).

(Te-U-12) Does the behaviour path with the states [1, 2, 4, 10, 11, 15, 18, 39, 50, 77, 94, 110] correspond to a trajectory shown by the system under natural conditions?

Options	Yes	Maybe	No	Did not answer
Number (%)	4 (36,4%)	3 (27,3%)	3 (27,3%)	1 (9,1%)

This question addresses the only behaviour path that received more negative evaluations in the whole set of data collected in the evaluation events described here. As shown in the table, 54,6% of the respondents said that the trajectory either was not found at all in nature or could be found in natural conditions. Some of the comments are the following: *"Partially. The values assumed by one of the variables are not correct."*; *"No. If flooded area stabilized in the final state, why 'transported garbage' was still large?"*; *"Partially. 'Transported garbage' should decrease when uncontrolled water decreases."* In fact, 'transported garbage' is not influenced by 'drained water', and therefore does not react to the implementation of infrastructure. This is an issue to be improved in future versions of the Riacho Fundo model.

(Te-U-13) Are the values shown by the quantities (presented in the diagram) in the behaviour path [1, 2, 4, 10, 11, 15, 18, 39, 50, 77, 94, 110] correct?

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	4 (36,4%)	5 (45,5%)	0	2 (18,2%)

Only two teachers made comments about this question: *"Yes. It is possible to predict what will happen along the whole process."*; *"Partially. The model shown is correct, except the values of 'transported garbage."*

(Te-U-14) Based on the causality flow represented in the causal model and in the values of the quantities, is it possible to explain the Urban RF system behaviour?

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	8 (72,8%)	3 (27,2%)	0	0

None of the teachers added comments to this question.

(Te-U-15) Based on the causality flow represented in the causal model and in the values of the quantities, is it possible to make predictions about the Urban RF system behaviour?

Options	Yes, completely	Partially	No, not at all	Did not answer
Number (%)	9 (81,9%)	2 (18,1%)	0	0

Comments:
"Yes, of course, we can draw a conclusion about the behaviour of the system."

The last two questions confirm that the causal model combined with the quantity values may be the basis for explanations and predictions. It is interesting to note that the teachers found an improvement in generating explanations and making predictions when the causal model is combined with the values of the quantities. In the previous section, the proportions were 63,7% for answer 'yes' and 36,3% for answer 'more or less' about the possibility of generating explanations based on the Urban causal model alone; in the present section, for the combination of causal model and quantity values, the option 'yes' was chosen in 72,8% of the answers and the option 'more less' in 27,2% of the answers.

Similarly, the possibility of making predictions based on the causal model alone received 72,8% of 'yes' and 18,2% of 'more or less'; and the combination of the causal model with the quantity values received 81,9% of 'yes' and 18,1% of 'more or less'.

5.5.5. Educational applications of qualitative models

Three questions explored the content of the qualitative models and the secondary school curriculum, for each of the Urban, Semi-urban and Rural perspectives, in connection with the experience of the teacher with these contents and the possibility of using the model to support teaching activities on these topics. The results are presented in this section.

Are the concepts addressed in this model about… [urban drainage included in the curriculum contents of secondary schools? (Urban)] [the relation between erosion and industrial production included in the curriculum contents of secondary schools? (Semi-urban)] [the relation between deforestation, animal biodiversity and agricultural production included in the curriculum contents of secondary schools? (Rural)]				
	Yes, completely	Partially	No, almost nothing	Did not answer
Urban RF	4 (36,4%)	3 (27,3%)	1 (9,1%)	3 (27,3%)
Semi-urban RF	4 (36,4%)	2 (18,2%)	3 (27,3%)	2 (18,2%)
Rural RF	6 (54,6%)	1 (9,1%)	1 (9,1%)	3 (27,3%)
TOTAL	14 (42,4%)	6 (18,2%)	5 (12,2%)	8 (24,2%)

The answers were quite heterogeneous, as expected, because the individuals included in the sample were teachers of different disciplines. Some comments illustrate this aspect.

Urban: *"Partially. Given that the students have only one semester of ecology, it [the educational application] depends on the school or the teacher."*; *"Partially. The subject receives a superficial approach."*; *"Partially. I don't know the curriculum of other disciplines, maybe in Geography and Psychology."*; *"No. This subject can be approached in an interdisciplinary project."*

Semi-urban: *"Yes: human populations quality of life, and structural theme of the PCN+ (the Brazilian National Curriculum Parameters)."*; *"Partially. Its importance depends on the teachers. They may be explored within the diversified part of the curriculum."*

Rural: *"Yes, these concepts are explored, for instance, in ecology."*

Did you ever explore this content in the classroom? [urban drainage (Urban)] [the relation between erosion and industrial production (Semi-urban)] [the relation between deforestation, animal biodiversity and agricultural production (Rural)]				
	Yes, completely	Partially	No, almost nothing	Did not answer
Urban RF	3 (27,3%)	1 (9,1%)	5 (45,5%)	2 (18,2%)
Semi-urban RF	0	2 (18,2%)	6 (54,6%)	3 (27,3%)
Rural RF	3 (27,3%)	3 (27,3%)	2 (18,2%)	3 (27,3%)
TOTAL	6 (18,2%)	6 (18,2%)	13 (39,4%)	8 (24,2%)

Some comments of the teachers on whether the relevant content is included in their teaching activities are presented below.

Urban: *"No. I do not work with these issues."*; *"No. Almost nothing from these issues is part of the Chemistry curriculum."*

Rural: *"Yes, in biology classes."*; *"Partially. If the teacher is interested in exploring these contents more deeply, it is possible."*

Would you use this qualitative model in your classes about this content? [urban drainage (Urban)] [the relation between erosion and industrial production (Semi-urban)] [the relation between deforestation, animal biodiversity and agricultural production (Rural)]				
	Yes, completely	Partially	No, almost nothing	Did not answer
Urban RF	5 (45,5%)	4 (36,4%)	1 (9,1%)	0
Semi-urban RF	7 (63,7%)	0	2 (18,2%)	2 (18,2%)
Rural RF	6 (54,6%)	2 (18,2%)	1 (9,1%)	2 (18,2%)
TOTAL	18 (54,6%)	6 (18,2%)	4 (12,1%)	4 (12,1%)

Some comments of the teachers about the viability of using the models in school are presented below.

Urban: *"Yes. I think it helps the development of creativity and makes it easy to find correlations between the concepts."*; *"Yes, because it provides support in teaching and learning. The student will reflect on the subject and make questions about the concepts further formulating conjectures and predictions about the behaviour of the system."*; *"Maybe, if it is made simpler than it is now."*; *"Yes. By using this model the student may use his/her knowledge in a constructivist [Piagetian] way broadening his/her knowledge and having more significant results."*; *"Maybe. It is necessary to adapt the model in accordance with the academic level of the student."*; *"I do not deal with these contents."*

Semi-urban: *"Yes, as soon as I can understand it better."*; *"Perfectly."*; *"Yes. In order to make it easier for the students to understand the role of human population in nature."*; *"Yes, but I would try to make it a bit simpler before using it."*; *"No, the contents are not addressed in my classes."*; *"Not this model, because the contents are not addressed in my classes."*

Rural: *“Yes, because it shows the whole process and its consequences.”*; *“Maybe, because I believe that too many classes are required to explore the whole model.”*; *“Maybe, to work in an interdisciplinary project.”*

5.5.6. Cognitive competences, abilities and qualitative models

The University of Brasilia runs one of the most advanced programs of selection of students for enrolment in undergraduate courses. The candidates are selected annually under an evaluation process that is based on parameters addressing not only the formal knowledge acquired during the previous school years, but also the cognitive ‘competence’ in using different types of language (Portuguese, mathematical language etc.), in dealing with problem-solving activities, in building up consistent argumentation, in making decisions and in presenting or assessing proposals of intervention in the reality based on technical and ethical criteria. Each of these competences may be mobilized by a number of (technical/ procedural) ‘abilities’, which can be also tested. In this section we present the opinions of the teachers about using qualitative models to support mobilization of abilities and development of competences by means of the contents illustrated by the models of the Urban, Semi-urban and Rural perspectives regarding the Riacho Fundo sustainability.

(Te-CA-01) The use of a limited set of modelling elements (entities, influences, proportionalities, qualitative values) for building up different qualitative models (urban drainage, erosion, deforestation) shows multiple meanings of the modelling language.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	4 (36,3%)	6 (54,6%)	1 (9,1%)	0	0	0

Comments: *“I fully agree. This modelling approach could explore any curricular component.”*; *“I agree, because it would be possible to adapt the subjects in the modelling language.”*

(Te-CA-02) The use of qualitative models contributes to the development of the student’s ability to identify central and peripheral information, presented in different contexts.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	4 (36,3%)	6 (54,6%)	1 (9,1%)	0	0	0

Comments: *“I fully agree. For example, when exploring the consequences of deforestation (Rural Riacho Fundo), information about vegetation cover is of central importance, as it is directly related to deforestation. Soil fertility and water availability are peripheral information concerning deforestation.”*; *“I fully agree, because the student is not restricted to quantity values.”* Some of the teachers mentioned additional requirements: *“With the adequate support from the teacher.”*; *“I agree, since we could make clearer the information available.”*

(Te-CA-03) Exploring qualitative models in educational activities allows for integrating knowledge from different areas (for example, natural sciences and humanities) and for implementing interdisciplinary projects.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	5 (45,4%)	6 (54,6%)	0	0	0	0

Comments: *"I fully agree. It would be possible to mention environmental education as an interdisciplinary subject."*; *"I agree. I understand that qualitative models could be applicable to a number of areas of knowledge."*

(Te-CA-04) Given a situation/ problem related to the well being of urbanites, the use of a qualitative model contributes for the student to organize strategies and to select adequate methodologies for solving the problem.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	2 (18,2%)	9 (91,8)	0	0	0	0

Comments: *"I fully agree. The student can think about a problem and make proposals with solutions of social or environmental intervention."*; *"I agree. If the teacher previously develops a concept map about the problems, it is possible to motivate the student to look for possible solutions."*; *"I agree. It supports the development of concrete actions in a constructivist [Piagetian] way."*; *"I agree. The qualitative model provides a broad overview of the problem, leading the students to draw their own solutions."*

(Te-CA-05) Qualitative models about the interaction between flood, garbage, polluted water and human well being in urban areas may help the students to produce explanations, to formulate hypotheses and to make predictions about the results.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	3 (27,3%)	8 (72,7%)	0	0	0	0

Comments: *"I fully agree, further adding that it also helps integrating competences and abilities defined for the secondary education level."*; *"I agree. The models shown in the present evaluation activity made it clear."*

(Te-CA-06) Using elements of the simulations about the importance of a urban drainage system (influence on the volume of uncontrolled water, size of flooded area, economic damage, amount of transported garbage, mosquitos and pathogens), qualitative models may provide support to the students in writing well structured texts (essays, reports) in which a thematic progression is found.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	4 (36,3%)	6 (54,6%)	1 (9,1%)	0	0	0

Comments: *"I fully agree. The student may describe the whole process in a system starting with the direct influence or agent that propagated the influences to the other quantities in the system."*; *"Fully agree, because the ideas are already organized in qualitative models."*

(Te-CA-07) Understanding the main concepts included in qualitative models (dynamics of continuous variables, processes, causality) contributes for the student to recognize limitations in this approach and to search for alternatives for situations in which these concepts cannot be applied.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	2 (18,2%)	4 (36,4%)	3 (27,3%)	1 (9,1%)	0	1

Comments: *"I fully agree. The models may stimulate scientific research."*; *"I agree. It may happen if there is a broad discussion on the topic."*

(Te-CA-08) The use of a qualitative model about engineered drainage system in urban areas contributes for the student to formulate and adequately articulate arguments about the importance of this theme.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	2 (18,2%)	8 (72,7%)	1 (9,1%)	0	0	0

Comments: *"I fully agree, because by exploring the qualitative model, the student can understand real problems that happen in the rainy season."*; *"I fully agree. Qualitative models stimulate an investigative attitude and the use of scientific methodology."*; *"I agree, because the student will have a global view of the system and will be able to make predictions about the results."*

(Te-CA-09) The analysis of the urban drainage system represented in a qualitative model allows the student to make inferences about the effects of uncontrolled water (floods) on the well being of the human population.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	4 (36,4%)	7 (63,6%)	0	0	0	0

Comments: *"I fully agree. The relations between uncontrolled water, flooded areas, garbage and pathogens are clear in the model."*; *"I fully agree. The model shows causes and consequences."*; *"I fully agree. Dealing with the models, the students find possibilities of improving the well being of human population."*; *"I agree. The student can run simulations with the model and compare the results."*

(Te-CA-10) The study of a qualitative model about erosion in rural areas contributes for the student to make analogies about erosion in areas undergoing urbanization.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	1 (9,1%)	9 (81,8%)	1 (9,1%)	0	0	0

Comments: *"I fully agree. The model shows causes and consequences.*

(Te-CA-11) Given a situation/ problem related, for example, to the implementation of the drainage system, the use of qualitative models may help the student to critically analyze the situation and point out positive and negative aspects of possible solutions.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	3 (27,3%)	8 (72,7%)	0	0	0	0

Comments: *"I agree. Having a general view of the process, it is possible to make a critical analysis."*

(Te-CA-12) With the support of qualitative models, the student may analyse and compare possible solutions for a given situation/ problem, such as the loss of fertility in rural areas.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	4 (36,4%)	7 (63,6%)	0	0	0	0

Comments:
"I fully agree. The analysis of qualitative models lead the student to a solution of a problematic situation."; *"I agree. The student can modify the model to produce different behaviour paths and to investigate whether the results are adequate."*

(Te-CA-13) The use of qualitative models may help the student to assess the adequacy of technical, social and political options in decision-making about the Riacho Fundo basin sustainability.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	5 (45,4%)	3 (27,3%)	3 (27,3%)	0	0	0

Comments: *"I agree. Even in the case that the student does not know exactly what the problem is, the model will show it."*

(Te-CA-14) The use of qualitative models helps the student to develop a vocabulary about a specific phenomenon, or a class of phenomena.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	2 (18,2%)	5 (45,5%)	4 (36,3%)	0	0	0

Comments: *"I fully agree. It is easier, because the phenomenon is presented in parts and in details."*; *"Deforestation: regeneration – degradation – vegetation cover / soil – aggregation – integrity – removal – erosion."*

(Te-CA-15) Qualitative models allow for combining different languages, such as the technical language and everyday vocabulary, in order to describe an specific phenomenon.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	2 (18,2%)	3 (27,3%)	6 (54,6%)	0	0	0

Comments: *"I fully agree. Precipitation and rain, pathogens, diseases, flood and uncontrolled water are examples."*; *"I agree, provided that the teacher uses the appropriate vocabulary.*; *"Maybe. Sometimes we do not use everyday vocabulary."*

(Te-CA-16) Qualitative models provide support for students to identify processes and values of quantities, and to understand how different factors are articulated in natural, social and economic phenomena.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	2 (18,2%)	7 (63,6%)	2 (18,2%)	0	0	0

Comments: *"I fully agree. For example, [it shows] how precipitation influences productive activities, the availability of natural resources and the well being of human population."*; *"Maybe. This is a process that takes much time to happen."*

(Te-CA-17) Given a situation/ problem, the use of qualitative models may help the student to take the relevant decisions in order to solve the problem.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	2 (18,2%)	7 (63,6%)	2 (18,2%)	0	0	0

Comments: *"I fully agree. For example, given problems such as lack of drainage systems or erosion, the student may change his/her own attitudes."*

(Te-CA-18) Qualitative models may provide elements for the students to build up consistent argumentation when defending a certain viewpoint, to use adequate vocabulary, identify variables and explore causal relations.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	1 (9,1%)	8 (72,7%)	2 (18,2%)	0	0	0

Comments: *"I fully agree. When the student is aware of the problems, (s)he can be more active (…)"*; *"I agree. However intensive work exploring the problem is required in order to achieve this goal."*

(Te-CA-19) Qualitative models may be useful to describe proposals of interventions in the real world, taking into account the diversity of human beings with respect to education, interests, and degree of understanding of the system represented in the model.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	1 (9,1%)	7 (63,6%)	3 (27,3%)	0	0	0

Comments: *"I agree. The model accepts inferences for further evaluation."*

(Te-CA-20) The qualitative models presented in this session provide elements to support educational projects about sustainable development of the Riacho Fundo basin and other basins, with similar features.

Options	Fully agree	Agree	Maybe	Disagree	Fully disagree	Did not answer
Number (%)	2 (18,2%)	8 (72,7%)	1 (9,1%)	0	0	0

Comments: *"I agree. Such projects should involve the community in which the student is inserted and the model should address such problems."*

5.5.7. The use of the models and the software at school

(Ta-G-1) Is the use of Garp3 compatible with the reality of the school where you work?

Options	Yes	Maybe	No	Did not answer
Number (%)	5 (45,5%)	4 (36,3%)	2 (18,2%)	0

Comments:

(a) "Yes, because laboratories of informatics are being implemented in state schools."
(b) "Yes, because the teacher in charge of the laboratory of informatics agreed in implementing the software Garp3."
(c) "Yes, I believe it is important to develop a project related to this software."
(d) "Maybe, I am not sure that teachers and students will feel comfortable with the interface and the language used in this software."
(e) "No, my school has no laboratory of informatics."
(f) "No, our reality is that people is below the level of understanding this software."

(Ta-G-2) Which are the positive aspects of Garp3?

(a) "I need to be more familiar with the program in order to point out the positive aspects."
(b) "The program simplifies conceptual representations and facilitates the overall understanding of the aspects addressed in the model."

(c) "Easy access and use. The first contact is scary, because there is so much information involved, but after using the program, I noticed that it is easy to interact with Garp3."
(d) "Free download, easy to manipulate."
(e) "Easy to install, easy to manipulate."
(f) "It stimulates the cognitive development and, with a constructivist [Piagetian] approach, it helps in learning and applying concepts to everyday life."
(g) "It makes it easy to build models; I believe that using models facilitates the acquisition of the notion of causality and gives flexibility to thinking."
(h) "It reduces the amount of information required to explain a phenomenon; it also provides more coverage to possible solutions."
(i) "It offers many options (variables) to organize or to evaluate phenomena."
(j) "It offers the possibility to recognize relations and different interactions between concepts."

(Ta-G-3) Which aspects of Garp3 should be modified?
(a) "The simulations."
(b) "Full translation into Portuguese of the vocabulary used by the program is needed."
(c) "I need to manipulate the program more in order to answer this question."
(d) "Full translation into Portuguese is needed."
(e) "It is necessary to translate parts of the software into Portuguese."
(f) "The program has to be presented in Portuguese. Some graphical aspects have to be improved."

(Ta-G-4) Which aspects of the qualitative models presented in this session should be modified in order to use the model in the classroom?
(a) "[It is necessary] To present few concepts in each model."
(b) "[It is necessary] To be clearer with the symbols and, in some cases, with the causal relations."
(c) "As I mentioned answering previous items, some parts of the vocabulary could be easier and more significant."
(d) "[It is necessary] To make things simpler for the student to handle less concepts in each interaction with the model."
(e) "It is necessary to have training sessions with teachers and students before using the model."
(f) "It is necessary] To select simpler problems and to describe details in captions or footnotes."

(Ta-G-5) Please, make suggestions of topics/ curricular contents that could be approached in new qualitative models so as to be used in your subject or in interdisciplinary work.
(a) "Cytology; botanic; genetics."
(b) "Ecological interactions between populations; the effects of the fire."
(c) "Environmental education; sustainability; ecosystems, ecological equilibrium."
(d) "Energy, its transformations and uses."
(e) "Food chain; photosynthesis; cellular respiration; ecosystems; environmental impacts."
(f) "Human physiology; ecology (interactions between populations); healthy nurture."
(g) "Some items in the chemistry curriculum, such as garbage treatment, water treatment, chemical equilibrium, residues produced in the school laboratory."

(h) "Functions; logarithm function; progressions."
(i) "Heat and temperature; the nature and properties of light."
(j) "Subjects that require analysis of results, such as socio-economic phenomena."
(k) "Soil pollution by heavy metals; global warming; fertilization of the cerrado soil."

(Ta-G-6) Final remarks

(a) The program is very interesting and also the project in which it has been developed. I think that given the strong impact of technology, we have a lot of things to do in order to improve our classes. Maybe the program requires some adaptations to be used in secondary schools."
(b) "I always used diagrams and schemata that resemble the qualitative approach. I found it a clear and visually interesting way to present relations between concepts and between causes and effects."
(c) "I think the material is excellent; the guidelines, the glossary and the supporting material are very didactic. However, we had a lot of new information to work out in only two afternoons. This can be scary and make it difficult to find new partners."
(d) "Garp3 represents the possibility of doing interdisciplinary and contextualized educational activities. However, I still think that the internal logic and the amount of work will scare the teachers."

5.6. Discussion

Section 5 of this Deliverable presents the evaluation of the Riacho Fundo model carried out by three types of stakeholders: experts in hydrology and water resources, managers of water resources and secondary school teachers. Nine meetings were held in order to disseminate the results of NaturNet-Redime and to evaluate the model. Different questionnaires were used, in order to explore different aspects of the model, in accordance with the profile of the evaluator group. This way, experts assessed the scientific contents of the model; teachers assessed the representation of concepts; and three groups[3] (experts, teachers and managers) assessed the usability of the model. Representative numbers of experts (16) and teachers (11) answered the questionnaires, but unfortunately, despite our efforts, a low number of managers (3) produced written answers about the Riacho Fundo model. The positive aspect is that these groups were very heterogeneous, and included a diverse set of people with different backgrounds and professional experience (section 5.2). The results are discussed here as follows: the conceptual validation is discussed in section 5.6.1, and topics on operational validation are discussed in section 5.6.2.

5.6.1. Conceptual validation

Conceptual validation (Rykiel, 1996) of the Riacho Fundo model was made by experts and by teachers. The former provided a scientific view on the model concepts, and the latter, the educational view on the model concepts.

5.6.1.1. Entities and quantities

Experts mainly evaluated the scientific aspects of the knowledge included in the model (section 5.3). Entities and configurations used to represent the Riacho Fundo system structure were mostly approved, although a number of experts would like to see other entities included in the model. One of the most difficult tasks for the modeller is to select

[3] In the remainder we refer to the cluster of 'experts, teachers and managers' as the 'three groups'.

what to include in the model and what to leave out. In the three groups the overall opinion is that none of the entities was wrongly selected, and that the system structure (entities and configurations) used in the three perspectives are enough to capture fully or partially the most relevant aspects of the Riacho Fundo sustainability (section 5.3.1).

The experts analysed three simulations, exploring models of the three perspectives (Urban, Semi-urban and Rural) concerning Riacho Fundo sustainability. Their opinion was that the quantities (47,9%) and qualitative values (50%) selected in the Riacho Fundo model were the most relevant ones (section 5.3.2).

5.6.1.2. Model fragments

Teachers provided an interesting validation of the representations of concepts in the model fragments and scenarios. Although time was short for mastering the modelling language, the educators were able to criticize the knowledge representation of relevant concepts for the Urban RF perspective. Model fragments about three processes and three views were evaluated: ‘Precipitation’, ‘Control of the water runoff by means of (engineered) drainage system’ and ‘Positive and negative factors that determine human well being’ were taken to represent crucial concepts that generate the dynamics of the system; ‘Uncontrolled water transported garbage’ and ‘Uncontrolled water and mosquito population growth’ and ’Water related diseases’ were taken to describe relations that are necessary for the propagation of the influences to the rest of the system. The teachers considered the model fragments conceptually clear and correct, except the one that represents the concept of improvement of human well being. In fact, this model fragment is highly abstract, as it summarizes a number of different influences. A conclusion then is that further development of the model and use in real settings will require a clearer representation of human well being. The comments provided by the teachers (section 5.5.2) are valuable contributions to the improvement of the model.

5.6.1.3. Simulations

Experts and the teachers also evaluated the correctness and conceptual clarity of the simulations. The teachers who analyzed a scenario of the Urban RF perspective found it clear and correct (section 5.5.2). Causal models, behaviour paths and value history diagrams were evaluated by both the experts (section 5.3.3) and by teachers (section 5.5.3). Discussion about the evaluation of the causal models is presented below, in detail. Experts were asked to provide comments on whether the events shown in the simulations were also found in natural conditions, on whether the values assumed by the quantities in each behaviour path were correct, and on whether these values represent values obtained in studies with the real system. As expected, these were difficult questions because the model abstracts from many features, hence some of them may be found under natural conditions, while others may not be. As commented by an expert, *“given that we are supposing things in the model, nothing is completely correct or wrong”*. Also, some of the quantities are measurable, while others not. Studies about urban drainage, erosion and deforestation hardly combine all these quantities. Important to note is that the vast majority of the system and quantity behaviours were acceptable for the experts. Further studies could better explore these quantities that are important for the understanding of the system behaviour but that are not measurable with quantitative methods available.

5.6.2. Causal model evaluation

Causal model is one of the most important representations of qualitative modelling. It is the basis for the causal reasoning and, as such, it is the basis for explanations and predictions about the system behaviour. In fact, causal model evaluation requires conceptual and operational validation techniques. It has so many impacts on the development and the use of qualitative models that it deserves a dedicated section in this discussion. In this section the results of the three groups are brought together for an overview of this important element.

The results concerning the correctness, and use of the causal models to support explanations and predictions about the system's behaviour, are very relevant for the model evaluation. Accordingly, the causal models obtained in simulations with the three perspectives (Urban RF, Semi-urban RF and Rural RF) were shown to the three evaluation groups. Each group explored different aspects. Experts assessed the scientific aspects, and managers and teachers assessed conceptual clarity and correctness.

The possibility of deriving explanations from qualitative models was evaluated from two points of view: the use the causal model alone - by managers and teachers; and the use of a combination of causal model and values of quantities - by experts and teachers (only with respect to the model of the Urban perspective). The results obtained confirm what has been published in the literature: one of the strongest aspects in the use of qualitative models is the (generation of the) causal model, as it promotes understanding of the system and can be used to support explanations and predictions about the system behaviour. The following details are worth discussing.

Scientific knowledge: A significant part to the causal relations (I's and P's) are well justified, either by scientific knowledge or by educated commonsense:

a. 25,0% of the experts considered that all the causal relations in the Urban RF could be justified in scientific grounds, and 37,5% of the experts considered that all the relations that could not be justified by scientific knowledge, could be justified by educated commonsense.
b. 50,0% of the experts considered that all the causal relations in the Semi-urban RF could be justified in scientific grounds, and 31,3% of the experts considered that all the relations that could not be justified by scientific knowledge, could be justified by educated commonsense.
c. 31,5% of the experts considered that all the causal relations in the Rural RF could be justified in scientific grounds, and 31,3% of the experts considered that all the relations that could not be justified by scientific knowledge, could be justified by educated commonsense.

A more detailed study is required to identify which relations cannot be scientifically justified and why this is so.

Conceptual representation - Managers and teachers both evaluated conceptual clarity and correctness. Their answers can be shown as follows:

a. 66,7% of the managers and 81,9% of the teachers considered the Urban RF causal model conceptually clear and correct;
b. 66,7% of the managers and 91,0% of the teachers considered the Semi-urban RF causal model conceptually clear and correct;
c. 66,7% of the managers and 63,7% of the teachers considered the Rural RF causal model conceptually clear and correct;

Additional studies are required in order to understand the reason why some managers and teachers found problems with the causal models.

Deriving explanations from the causal model - The possibility of deriving explanations about the system behaviour based on I's and P's was evaluated by managers and teachers. Their answers are shown as follows:

a. 100,0% of the managers and 63,7% of the teachers considered that it is possible to derive explanations about the system behaviour based on the Urban RF causal model;
b. 66,7% of the managers and 82,9% of the teachers considered that it is possible to derive explanations about the system behaviour based on the Semi-urban RF causal model;
c. 72,8% of the teachers considered that it is possible to derive explanations about the system behaviour based on the Rural RF causal model; none of the managers has answered this question.

Making predictions based on the causal model - The possibility of making predictions about the system behaviour based on I's and P's was also evaluated by managers and teachers. Their answers are shown as follows:

a. 100,0% of the managers and 72,8 % of the teachers considered that it is possible to derive explanations about the system behaviour based on the Urban RF causal model;
b. 33,3% of the managers and 91,0% of the teachers considered that it is possible to derive explanations about the system behaviour based on the Semi-urban RF causal model;
c. 63,7% of the teachers considered that it is possible to derive explanations about the system behaviour based on the Rural RF causal model; none of the managers has answered this question.

From the evaluation it can be also inferred that the values of the quantities obtained in simulations should be used along with the causal model by stakeholders in explanations and predictions. In order to investigate this aspect of the use of causal models, expert and teachers were asked to answer dedicated questions on this combination. Their answers can be summarized as follows:

Deriving explanations from the causal model and quantity values (experts)

a. 43,8% of the experts considered that the system behaviour can be completely explained based on both the Urban RF causal model and the values of the relevant quantities in a behaviour path;
b. 43,7% of the experts considered that the system behaviour can be completely explained based on both the Semi-urban RF causal model and the values of relevant quantities in a behaviour path;
c. 50,0% of the experts considered that the system behaviour can be completely explained based on both the Rural RF causal model and the values of relevant quantities in a behaviour path.

Explanations and predictions based on causal model and quantity values (teachers) - The teachers assessed the combination of causal model and quantity values only in the Urban RF perspective:

a. 72,8% of the teachers considered that the system behaviour can be explained completely based on both the Urban RF causal model and the values of relevant quantities in a behaviour path.
b. 81,9% of the teachers considered that predictions about the system behaviour can be based completely on both the Urban RF causal model and the values of relevant quantities in a behaviour path.

Comparing explanations and predictions based only in the causal model and based in both the causal model and the quantity values, it is interesting to note that the increased proportion of teachers who believe that the system behaviour could be completely explained it would be possible to make predictions about the system behaviour:

I. for deriving explanations, there was an improvement from 63,7% (causal model alone) to 72,8% of the teachers who believe that the combination of causal model + quantity values completely explain the system behaviour;
II. for making predictions, there was an improvement from 72,8 % (causal model alone) to 81,9% of the teachers who believe that the combination of causal model + quantity values completely support predictions about the system behaviour;

These observations may have an important impact in a number of applications of qualitative models, such as research, management and the development of didactic material.

5.6.3. Operational validation

5.6.3.1. Modelling language

Operational validation is related to the use of the model. Most of this type of validation was provided by managers and teachers. These professionals are potential users of the model in decision-making and in the classroom. First, the use of the modelling language is analysed (sections 5.4.1 and 5.5.1). Managers were asked to assign a mark on the degree of difficulty they had personally had to understand each modelling element. The three respondents did not have much difficulty to understand them. Next, the managers were asked to give their opinion about how difficult it would be for other people, with different backgrounds, to understand the meaning of the modelling elements. From their answers it is possible to see that they found a positive correlation between the years of study and the ability to understand the modelling elements.

The teachers had a similar view on the modelling elements. For this segment, the answers showed the following proportion in the choice of the options: 16 ‘very easy’; 51 ‘easy’; 69 ‘medium’. As expected, the notion of process and the use of I’s is considered the most ‘difficult’ modelling element (10 teachers marked the option ‘medium’). However, as noted by one of the participants, practice may change the situation: *“As soon as we explore models already built and try to build our own models, we become familiar with the modelling elements.”*

For the students, things are taken to be harder: 75% of the teachers’ answers were assigned to the options ‘medium’ and ‘difficult’. Among the modelling elements, processes, direct influences, proportionalities and correspondences were the ones that received more ‘medium’ and ‘difficult’ marks. Some teachers asked for simplifications in the modelling language and noted that training is a way of improving understanding, but it may not help much, as can be inferred from one of their comments: *“it is difficult to say whether these elements are difficult or easy for the students. We have students that are able to deal with scientific concepts, while others rarely deal with them.”*

5.6.3.2. Managerial and educational activities

Considering that operational validation requires to access whether the model output meets the performance standards required for the model purpose, managers and teachers were asked to answer some questions about the use of the model in managerial (section 5.4.3) and educational activities (section 5.5.5). Although many managers showed, during the presentation of the model, that they are familiar with urban drainage problems, only two out of three respondents said they had experienced discussions. The same occurred with the problems addressed in the Semi-urban and Rural RF perspective models. However, the managers said they would use the models to support their future discussions on water resources management: *"it is very useful to make things clear."*

Qualitative models have traditionally been used in educational activities. When asked about the relation between the model contents and the secondary schools curricula, 42,4% of the teachers answered that the model contents (in the three perspectives) are included in the curricula; for 18,2% of the teachers, only partially; and for 12,2%, almost nothing. We conclude that many factors have influenced this result: teachers of different disciplines answered the questionnaire; there are differences from one school to another; some curriculum components receive less attention than others, and so on. These observations explain why a significant number of teachers said they have explored almost nothing of these contents (45,5% for themes in the Urban; 54,6% for Semi-urban; 18,2% for Rural RF model). However, having had contact with qualitative models, an expressive proportion of the teachers would use the qualitative models in the future.

The most interesting aspect of this evaluation effort is the results about the use of qualitative models for the mobilization of abilities while exploring the curriculum contents in order to develop cognitive competences. As mentioned in section 5.5.6, competences and abilities are the main driver for the organization of educational activities in secondary schools, not only in Brasilia, but in the whole country.

Among the most important abilities, qualitative models have the potential to support students in identifying central and peripheral information, integrating knowledge coming from different areas, selecting strategies and adequate methodologies for problem solving, formulating hypotheses, making predictions and producing explanations. They may also provide means to facilitate the organization of ideas, the development of arguments and to improve the ability to criticize ideas and proposals. All in all, the use of qualitative models in the context of competences and abilities may become a major breakthrough for the educational applications of qualitative reasoning in Brazil.

5.6.3.3. Conditions for using qualitative models

The use of qualitative models in the classroom may be limited by a number of factors. As mentioned by the teachers (section 5.5.7), some schools have no computer science laboratory, teachers and students may not be prepared to use the software Garp3 and / or the models. However, things are changing (*"laboratories of informatics are being implemented in public schools"*) and the teachers found a number of positive aspects in Garp3 and the models. One of the most mentioned barriers – the language – is currently being solved, as multiple languages (including Portuguese) facilities are being implemented in Garp3. Other aspects, such as the need from training sessions with the teacher and the students before going into the models for serious use, are not difficult to solve.

5.6.4. Potential for using qualitative models

The three questionnaires end with open questions for the stakeholders to express their views on the potential for using qualitative models. Their comments, summarized in sections 5.3.4, 5.4.4 and 5.5.7 are very interesting and very rich of ideas. The experts' opinions show that qualitative models can be used for education and training, research and natural resources management. The representation of cause-effect relations and the facility to communicate with the public are recognized as valuable features for education and training. The possibility of combining a number of non-measurable quantities, of becoming an additional tool for the development of mathematical and complex models and for the representation and testing of hypotheses were seen as important features to support research. Finally, experts mentioned features that make qualitative models an interesting tool for management, such as *understanding the cause and effect relations involving different factors, and therefore allowing for changing one of them to get consequences in the whole system*. Of course, they send some warnings about using the models, such as the danger of *"generalizations and answers to some problems that apparently can be universally applied"*; and the level of complexity, that may be too high *"to be used with the general public."*

5.7. Final conclusions

Section 5 of the Deliverable 7.2 presents the evaluation of the Riacho Fundo model with stakeholders. The model includes three sets of simulations, each providing a perspective to the sustainability regarding the basin, namely the Urban, Semi-urban and Rural perspectives. Each of these viewpoints, created to represent problems identified by stakeholders in the region, explore different but complementary aspects. Urban areas deal with the lack of drainage system and, as consequence, with flooding, mosquitos, transported garbage, pathogens and other aspects that affect the human well being. Semi-urban areas show the effects of urbanization on soil particles aggregation and consequences of erosion and reduction of water infiltration. The simulations explore how these factors may hamper industrial production and water production from springs and streams. Rural areas are represented in the model as areas that are being deforested and, as a consequence, problems of soil and nutrient loss, deposit of sediments on water bodies, reduction on agricultural production and loss of biodiversity arise.

Three groups took part in the evaluation events, namely experts from the Brasilia Water and Sewage Company (CAESB), water resources managers and secondary school teachers. Nine events of dissemination and evaluation activities were held between July and October 2007, involving in total almost 100 representatives of these groups. Three types of questionnaires were prepared, exploring different aspects of the Riacho Fundo model. The experts were concerned with the conceptual validation of the model on scientific grounds inspecting the clarity and correctness of concepts expressed by the entities and configurations (the representation of the system structure), the quantities and the simulations. The teachers assessed the concepts represented in model fragments. During the operational validation different aspects of the model were assessed. Both the managers and the teachers evaluated the use of the model in their dedicated activities. Special attention was given to the evaluation of the causal model, and to the use of qualitative models to support the development of cognitive competences and abilities.

The results were very positive. The experts agreed that the concepts could be justified by scientific knowledge or educated commonsense. The managers considered the modelling language easy and accessible to different potential water managers. They also mentioned that the causal models are conceptually correct and clear, and could be useful to support the generation of explanations and predictions about the system behaviour. The teachers considered the models very useful for educational purposes and accessible for secondary school students. They also assessed the generation of explanations and predictions about system behaviour based on causal models alone or in combination with qualitative values of relevant variables. The teachers also recognized the potential of qualitative models for supporting the development of cognitive competences and abilities. Finally, the three types of stakeholders were very positive about the potential use of qualitative models in general and in the Riacho Fundo model in particular, for representing problems and maybe suggesting solutions for real world problems in their activities.

6. Evaluation of River Kamp model with stakeholders

The qualitative simulation models related to the sustainable development of the Kamp valley are intended to be used by stakeholders and students as learning material (Zitek et al., 2007). The models explore aspects related to two topics:

- Development and implementation of sustainable actions in a river catchment (stakeholder integration, quality of sustainability plans, development of ecological integrity and human well being, and probability of catastrophic events).
- Hydropower production (water storage and release, and water abstraction) and its effect on fish.

The evaluation of models is an important step in the model building process. Generally a model evaluation basically covers 'validation and verification' of the model as well as the 'acceptance of the chosen approach and model' by the addressed stakeholder groups. Validation proves if the scientific and conceptual contents of the model are acceptable for its intended use, verification proves that the model is correctly implemented by a demonstration of its use. Proving the acceptance of stakeholders mainly evaluates the potential of the model and the modelling approach for broader use.

Our evaluation approach is based on the model evaluation process described by Rykiel (1996):

- Conceptual validation: does de model provide a scientifically acceptable explanation for the cause-effect relationships encoded in the model?
- Operational validation (= verification): do the model outputs meet the performance standards required for the model purpose?
- Data validation: are the data (or information) used to build, calibrate, and test the model of good quality?

6.1. Methods

To evaluate the two models developed (Model A: 'Sustainability Management' and Model B: 'Water abstraction and Fish') a two-phase approach was chosen. A general evaluation of Model A mainly focusing on the 'acceptance of the chosen approach and model' by students and scientists of different domains and an expert evaluation of both models focusing on 'validation and verification' of the models were conducted separately.

The expert evaluation of model A was performed by an expert in 'integrated management of riverine landscapes and the development and definition of integrated restoration activities related to the requirements of the EU-WFD'. Model B was evaluated by an expert in 'assessing the ecological status of running waters related to the effects of human pressures on aquatic ecosystems also dealing with the development and definition of restoration measures related to the requirements of the EU-WFD'. Both evaluations took place at the Institute of Hydrobiology and Aquatic Ecosystem Management at the University of Natural resources and Applied Life Sciences, BOKU, Vienna. The expert evaluations lasted about 2 hours each and were run in October 2007 as face-to-face discussions using the printed causal maps, and an exploration of important model fragments and simulations using GARP3 on a laptop.

The general evaluation used a power point presentation to explain general aspects, and personal laptops for a collective exploration of parts of the model using GARP3. It was held in October 2007, and lasted 2 hours. In total, 11 persons, divided into students and experts of different aquatic resource domains, participated the event.

After the presentation, and the collective interactive inspection of important scenarios and model fragments, the participants were asked to fill out questionnaires about GARP3 software. The following questions were used for the general evaluation of Model A (Sustainability Management), dealing mainly with a general evaluation of the model represented in QR language, proving the acceptance of the modelling approach and the model and its potential for broader use.

At the beginning of the evaluation process, the attendees were asked, whether they are an expert in a specific scientific field or a student.

- Expert - Please add, which kind of experience you have (teaching, water resources management, research…)
- Student -Please add the type of study you are conducting!

Next, they were asked to answer a set of evaluative questions (Appendix 3).

6.2. Results

6.2.1. General evaluation of Model A ‘Sustainability Management’

In this section, the most important results of the general evaluation of Model A ‘Sustainability Management’ are highlighted. In total the dominating answer was ‘I largely agree’ (n=43 times) followed by ‘I fully agree’ (n= 34 times) and ‘I somewhat agree/disagree’ (n=18 times). 6 times the statements were commented with ‘I largely disagree’ and 2 times with ‘I fully disagree’ (see Figure 1). For students the dominating statement was ‘I fully agree’, whereas experts answered most of the questions with ‘I largely agree’. As one student ‘largely disagreed’ or ‘fully disagreed’ with all of the statements, it could be, that he probably understood the evaluation scheme in the wrong way.

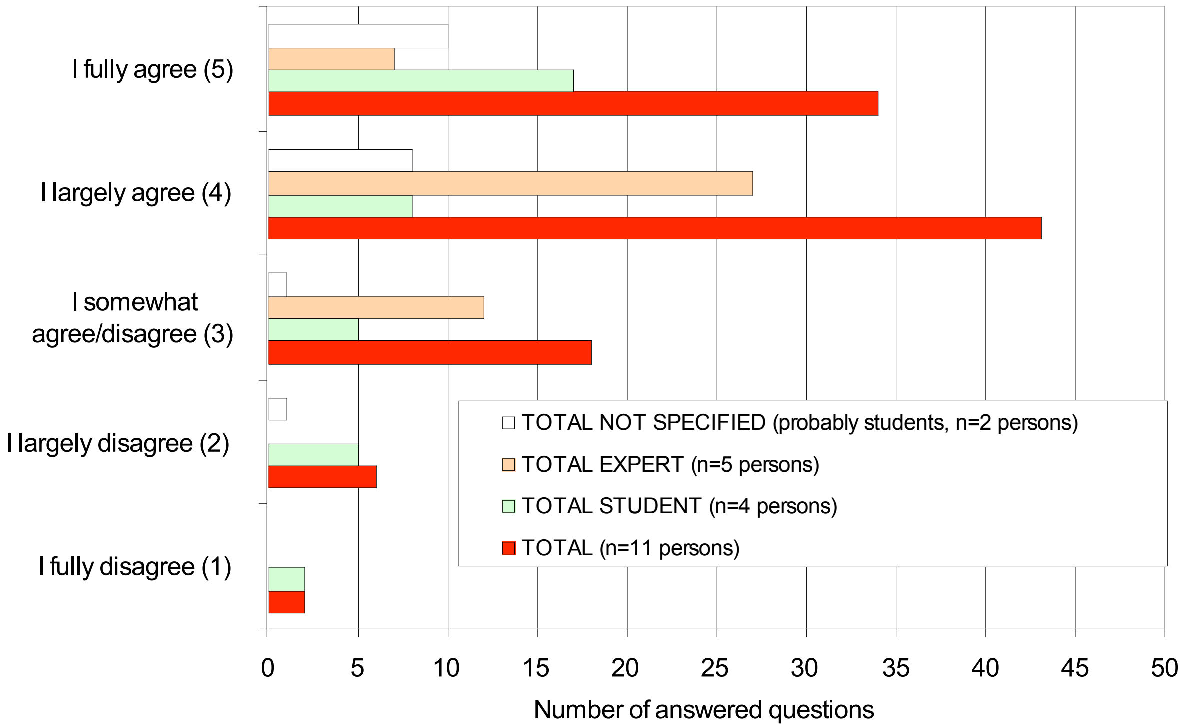


Figure 6.1: Summarized answers of the different groups of persons that attended the general evaluation of model A ‘Sustainability Management’.

Below the details are enumerated for each of the questions (see also Table 6.1, Appendix 3).

Q1: QR models present complex knowledge in an understandable manner.

- 7 persons 'agreed largely', 3 persons 'agreed fully' and one person 'largely disagreed'. All experts 'largely agreed' whereas 50% of the students fully agreed.

Q2: The QR approach allows for a clear representation of real-world phenomena like a sustainable development of the riverine landscape Kamp.

- Most persons (n=6) 'somewhat agreed/disagreed' with this statement, one person 'largely agreed', two persons 'fully agreed' and one 'largely disagreed'. One person did not answer this question.

Q3: QR and GARP3 can be seen as a valuable learning tool for real-world causal relationships related to a sustainable development of riverine landscapes.

- 5 persons 'agreed fully' with this statement, four 'agreed largely', one 'somewhat agreed/disagreed' and one 'largely disagreed'. Most experts (n=3) 'largely agreed' and most students (n=3) 'fully agreed' with this statement. One student added an additional statement: *'QR and GARP are well suited to visualize complex problems related to human affairs that are hard to capture with other approaches, but for a real clear and precise representation quantitative models will be needed'.*

Q4: The presented QR model might significantly contribute to the understanding of students and stakeholders which entities and processes drive a sustainable development of a riverine landscape and therefore enhances their capability of making decisions.

- All experts (n=5) 'largely agreed' and most students (n=3) 'fully agreed'. One student added an additional statement: *'Adaptation of the model to the addressed stakeholder group (e.g. simplification) is probably necessary; people should understand the model (…related to the quantity 'Ways to inform and integrate the public')'.*

Q5: The causal map of the model reflects important information related to a sustainable development of the Kamp valley.

- 3 experts 'largely agreed' and in total 3 persons 'agreed fully' (1 expert, 1 student and one person not specified) and 3 persons (2 students and 1 expert) 'somewhat agreed/disagreed' One student added an additional statement: *'More variables should be used to describe the situation more realistic'.*

Q6: Which part of the model was most interesting for you?

- Students added the following statements:
    - *'Most interesting to see was that private interests might negatively influence the sustainability process and that the combined influence of planners, science and local population (stakeholders) defines the quality of sustainability plans and the whole sustainability process'.*
    - *'Most interesting was to see the interrelatedness of the involved entities of the Kamp management system'.*
    - *'It was interesting to see the possibility of different potential intervention options to reach the goal of a sustainable development'.*
    - *'It was interesting to see that ecological integrity AND human well-being are represented in the sustainability model'.*

- Experts added the following statements:
  - *'Generating stock/flow elements in a qualitative way and relating quantities via P's and I's was very interesting'.*
- One person (expert/student status not specified)[4]
  - *'Specific scenarios were very interesting especially the catastrophic event as trigger for government action for sustainable development, the interrelatedness and influences of the different model components'.*

Q7: Which part of the model most should be enhanced?

- Students added the following statement: '*Private interests should be better represented, as a basis to minimize them and achieve sustainable development'.*
- Experts added the following: '*The full causal model looks a bit complex and heterogeneous (sorting?); pointing out better the most important variables for a sustainable development'.*

Q8: The model can be used for the targeted purpose of teaching students and other interested stakeholders on sustainability issues on a catchment level.

- 3 persons (2 students, 1 expert) 'fully agreed', 3 persons (1 expert, 1 student, and 1 person (status not specified)) 'largely' agreed and 3 persons (all experts) 'somewhat agreed/disagreed'.

Q9: For which purpose do you think the presented QR approach is most suited?

a. Stakeholder integration
   - 5 persons (3 experts, 1 student and 1 status not specified) 'largely agreed' and 4 persons (2 students 1 expert and 1 status not specified) 'fully agreed'
b. University lectures
   - 4 persons (3 experts, 1 student) 'largely agreed' and 4 persons (2 students 1 expert and 1 status not specified) 'fully agreed'
c. Decision-making
   - 6 persons (3 experts, 1 student and 2 status not specified) 'largely agreed' and 3 persons (2 experts and 1 status not specified) 'fully agreed'
d. Others (to be added e.g. technical staff from the government, researchers, secondary school students).
   - The following statements were added: *'The model and the general approach could be also suited for NGO's and other interest groups (stakeholders)'*, by a student. *'The model and the general approach are suited also for research'*, by an expert. '*The model and the general approach could be suited for general others and school teachers'*, added by 2 persons with status not specified.

Q10: Additional comments:

The following statement were added by a student:

- *'Keeping the overview over the model and the P's and I's is a bit complicated'.*

The following statement were added by a person with status not specified:

- *'The modelling approach allows for a communication of new viewpoints of existing problems and facts; the application within different other fields and domains, mainly related to education, is thinkable; it is interesting to take different mental models (also culture-specific) as a starting point for developing models and discussions'.*

[4] The participant did not specify this detail on the form.

6.2.2. Expert evaluation of Model A 'Sustainability Management'

Out of 16 possible answers, 14 were answered with 'I largely agree' during the expert evaluation of Model A 'Sustainability Management'. Many additional statements were given. Below the details are enumerated for each of the questions (see also Table 6.2, Appendix 3).

Q1: QR models present complex knowledge in an understandable manner.

- 'I largely agree'. Additional statement: *'If you are not part of the modelling process, it is not so easy to understand all definitions of the terms used within the model'.*

Q2: The QR approach allows for a clear representation of real-world phenomena like a sustainable development of the riverine landscape Kamp.

- 'I largely agree'. Additional statement: *'The model shows interdependencies and causal relationships very transparent; but if all of the model assumptions are really true, further real-world assessments and studies have to be done (more case studies)'.*

Q3: QR and GARP3 can be seen as a valuable learning tool for real-world causal relationships related to a sustainable development of riverine landscapes.

- 'I largely agree'. Additional statement: *'It is important to address the needs of the specific stakeholder group; sometimes these models might be too complicated (people need to have some education e.g. to deal with complexity and causal relationships – to understand I's and P's for example, in a modelling approach like this)'.*

Q4: The presented QR model might significantly contribute to the understanding of students and stakeholders which entities and processes drive a sustainable development of a riverine landscape and therefore enhances their capability of making decisions.

- 'I largely agree'.

Q5: The causal map of the model reflects important information related to a sustainable development of the Kamp valley.

- 'I largely agree'.

Q6[5]: The entities and configurations are relevant and sufficient to support a representation of the system structure.

- 'I largely agree'.

Q7: The quantities used capture the most interesting properties of the entities.

- 'I largely agree'.

Q8: The quantity spaces and values capture the most interesting qualitative states of the entities.

- 'I largely agree'.

Q9: The (important) model fragments are conceptually correct and clear.

- 'I largely agree'.

[5] Notice that the order of questions differs compared to the order used for the general evaluation.

Q10: The presented scenarios describe a real situation that it is good enough to trigger an interesting/good simulation.

- 'I largely agree'.

Q11: The general behaviour (how it develops trough the simulation) of the presented model is in accordance to what is already known (or accepted).

- 'I somewhat agree/disagree'. Additional statement: 'Other scientific fields (sociology, political science…) should be integrated and/or asked to deal with these questions properly'.

Q12: Which part of the model was most interesting for you?

- Additional statement: *'Causal relationships in general'. 'To model the money for community driven development as an own quantity; the idea, that money spent for measures can only be treated as money spent for a community driven development, if the community is involved in the process of developing and implementing measures (stakeholder participation); otherwise the money spent is not a community driven investment!'*

Q13: Which part of the model most should be enhanced?

- Additional statement: *'The government action for sustainable development should be better described, as in reality this is of high complexity, being also driven by the general political structure, difficulties between different organization units with regard to their competences (personal behaviour!) and differences in financial resources; additionally very often policies with complementary aims are existing, as policies often lack behind the social development. That means, a more detailed study and representation of the internal political structures determining the implementation process is needed'.*

Q14: The model can be used for the targeted purpose of teaching students and other interested stakeholders on sustainability issues on a catchment level.

- 'I largely agree'.

Q15: For which purpose do you think the presented QR approach is most suited?

a. Stakeholder integration: 'I somewhat agree/disagree'.
b. University lectures: 'I largely agree'.
c. Decision-making: 'I largely disagree'.
d. Others (to be added e.g. technical staff from the government, researchers, secondary school students).
    e. Planners, for their understanding of their social role (e.g. as shown in Model A) ('I largely agree').

Q16: Additional comments

- *'A good approach for social learning and conflict management (mediation), when applied together with students or stakeholders in a mediated or group modelling process'.*
- *'Identifying dependencies and causal relationships is of high interest and importance for understanding a system'.*
- *'It could also be of relevance, to think about which degree of fulfilment the three pillars should have (the weighting of the individual factors) to really reach a sustainable development; who defines the relationships between the pillars? Often the focus is more on the ecological side, sometimes more on the economic side that is currently closely linked to human well being in industrial societies;*

human wishes or controlling paradigms of society are often not sustainable; probably the currently controlling paradigm of integrating all the needs and wishes of the human population does not always lead to a sustainable development from an ecological point of view'.

- *'It is important to very well define the terms and their use within the model'!*

6.2.3. Expert evaluation of Model B 'Water abstraction and Fish'

Out of 16 possible answers, during the expert evaluation of Model B 'Water abstraction and Fish', 9 were answered with 'I fully agree', 2 with 'I largely agree', 4 with 'I somewhat agree/disagree' and 1 with 'I largely disagree'. Many additional statements were given. Below the details are enumerated for each of the questions (see also Table 6.3, Appendix 3).

Q1[6]:

- 'I somewhat agree/disagree'. Additional statement: *'There are still OR domain specific ingredients, semantics and behaviours (e.g. the quantity spaces as points and intervals), that might conflict with the intuitive way of stakeholders to express things)'.*

Q2:

- 'I somewhat agree/disagree'.

Q3:

- 'I somewhat agree/disagree'. Additional statement: *'Sometimes the model does not reflect a 'real' causal relationship; here it would be good to point out more specific the difference between real causal parameters and 'surrogate' parameters ('latent' parameters, a definition used in structural equation modelling – SEM) that interact with variables in a 'correlating' way)'.*

Q4 – Q8:

- 'I fully agree'.

Q10:

- 'I largely agree'. Additional statement: *'Some behaviours related to intervals in quantity spaces might not be true in real-world systems (e.g. that they stay within an interval for a certain time steps before they change). This should be avoided, when not explicitly defined as model target!'*

Q11:

- 'I fully agree'.

Q12:

- Additional statement: *'That it is easy to change the content of a scenario by using and exchanging different assumptions that simply allows to model the effects of the same human pressure on different guilds of fish (positive and negative effects of flow velocity and water temperature on different guilds).'*

Q13:

- Additional statement: *'A more realistic representation of the natural variability of the river discharge (probably by using the random function in the scenario editor)*

[6] See for questions details the previous subsection discussing model A.

and the amount of abstracted water related to mean annual flow as this defines the frequency of water overflow events at weirs that are suspected to have a significant effect on fish. A more realistic representation of the influence of the length of the water abstraction stretch on the temperature development within the river (at the moment the river stretch is treated as a 'container' with the same abiotic factors everywhere) and an integration of the effect of morphology on fish and on water temperature'!

Q14:

a. 'I fully agree'.

Q15:

b. Stakeholder integration: 'I somewhat agree/disagree'.
c. University lectures: 'I fully agree'. Additional statement: *'The approach is very well suited for interactive learning'.*
d. Decision-making: 'I largely disagree'.
e. *Others (to be added e.g. technical staff from the government, researchers, secondary school students).*
   - *'Adult education and environmental education' ('I fully agree').*

Q16:

- *'The software now can be used very intuitively, which is a prerequisite for the target, to motivate stakeholders and students to put their conceptual knowledge in causal models'!*
- *'It takes time and engagement, to establish approaches like that in society and (university) education teaching such approaches are the basis for their broader use and application by the upcoming generation(s)'.*
- *'To further enhance the modelling process itself it would be helpful to always see the consequences of my model definitions and implemented model fragments (configurations, proportionalities and influences) on the fly in an accompanying window of the software (for example as they can be explored by the 'show entities and configurations' button, by the 'show dependencies' button)'.*
- *'It also could be helpful to have the full model shown in a screen like in the 'show entities and configurations' window with the opportunity to select parts of the model to be run in a simulation (running only parts of the model by simply drawing a window over a certain part of the model)'.*
- *'To link the outcomes of causal models to a GIS would open a new field of promising applications!'*

6.3. Concluding remarks

Both evaluations, the general evaluation of Model A 'Sustainability Management' and the expert evaluations of Model A and B 'Water abstraction and Fish' yielded a very positive feedback with regard to the QR approach, the GARP3 software used to build models, and the models themselves representing important issues related to the sustainable development of the riverine landscape Kamp. For example most people 'largely agreed' or 'fully agreed' that QR models represent complex knowledge in an understandable manner and that QR and GARP3 can be seen as a valuable learning tool for understanding real-world causal relationships related to a sustainable development of riverine landscapes. Also most people 'largely agreed' or 'fully agreed' that the presented QR models might significantly contribute to the understanding of students and stakeholders which entities and processes drive a sustainable development of a riverine landscape and therefore enhances their capability of making decisions. So the general aim, to produce software and models in QR language that allow people to interact with and learn about sustainable development clearly can be seen as fully achieved.

Experts were a bit more conservative in agreeing with the approach than students. That could be because (i) that experts know better about problems of model building and therefore do not agree full with many things (they only agree ‘largely’), or (ii) that students can be more influenced by the opinion of the presenter being on fire with QR modelling. On the other hand some students gave sometimes answers like ‘I fully disagree’ or ‘largely disagree’, which did not occur that often with the persons considering themselves as experts. That means in our opinion, that these students probably have not yet understood the potential of the approach or they simply made a mistake when answering the questions (they probably misinterpreted the rating scheme).

Important additional statements related to the QR approach, the software, and the models were also collected. Most interesting for the attendees was to see the interrelatedness of the system presented and the use of qualitative ‘stock-flow’ dynamics known from the System Dynamics approach. Only some added that they sometimes get a bit lost when confronted with the total view of the causal model describing a sustainable development of the Kamp valley. It was also stated that when showing these models to other user groups, their general ability to deal with complexity should be accounted for; meaning that for each user group the way of presenting the model should be adopted. Probably sometimes these models might be too complicated for certain stakeholder groups (people need to have some education e.g. to deal with complexity and causal relationships – to understand I’s and P’s for example, in a modelling approach like this).

Participants suggested a high potential for the application of QR models in various fields, particularly in education but also in decision-making and research. The potential of the GARP3 software, and the QR approach to sustain collective, interactive social learning was clearly pointed out. Particularly, the identification of dependencies and causal relationships can be seen as a prerequisite for understanding a system and therefore also for learning and decision-making. With regard to a broader use of QR models in society especially for decision-making it was stated, that it might take some time and engagement to establish approaches like that in society. (University) education using and teaching such approaches can be seen as an important basis for a further application.

Parts of the Model A that were most interesting for the evaluators were:

- To see the causal interrelatedness of the involved entities of the Kamp management system and especially that private interest might negatively influence the sustainability process and that the combined influence of planners, science and local population (stakeholders) defines the quality of sustainability plans and the whole sustainability process. This understanding opens up the possibility of different potential intervention options to reach the goal of a sustainable development.
- To see that ecological integrity AND human well-being are represented in the sustainability model.
- Specific scenarios showing the catastrophic event as trigger for government action for sustainable development.

- The idea that money spent for measures can only be treated as money spent for a community driven development, if the community is involved in the process of developing and implementing measures (otherwise the money spent is not a community driven investment!).
-

Parts of the Model A that should be enhanced in the eyes of the evaluators were:

- Private interests should be better represented, as a basis to minimize them and achieve sustainable development
- The government action for sustainable development should be better described, as in reality this is of high complexity, being also driven by the general political structure, difficulties between different organization units with regard to their competences (personal behaviour!) and differences in financial resources; additionally very often policies with complementary aims are existing, as policies often lack behind the social development. That means, a more detailed study and representation of the internal political structures determining the implementation process is needed.

Parts of the Model B that were most interesting for the evaluators were:

- That it is easy to change the content of a scenario by using and exchanging different assumptions that simply allows modelling the effects of the same human pressure on different guilds of fish (positive and negative effects of flow velocity and water temperature on different guilds).
-

Parts of the Model B that should be enhanced in the eyes of the evaluators were:

- A more realistic representation of the natural variability of the river discharge (probably by using the random function in the scenario editor) and the amount of abstracted water related to mean annual flow as this defines the frequency of water overflow events at weirs that are suspected to have a significant effect on fish.
- A more realistic representation of the influence of the length of the water abstraction stretch on the temperature development within the river (at the moment the river stretch is treated as a ‘container’ with the same abiotic factors everywhere) and
- An integration of the effect of morphology on fish and on water temperature.

With regard to the presented models but also to the QR approach some further interesting statements were collected. For example it was stated, that some behaviours of simulations might not be true in real-world systems (e.g. that they stay within an interval for a certain time steps before they change). This should be avoided, when not explicitly defined as model target! That means on the one hand that the simulation behaviours of final models to be presented should be restricted as much as needed to avoid outcomes that are not intended (although one also might also significantly learn from unwanted outcomes of a simulation) and on the other hand that there are still QR domain specific ingredients, semantics and behaviours (e.g. the quantity spaces as points and intervals), that might conflict with the intuitive way of stakeholders to express things. Therefore we suggest that the end user should (i) only be confronted with simulations and scenarios that exactly show the intended behaviour, and (ii) as less as possible confronted with QR domain specific features not to irritate an intuitive modelling building practice by domain specific restrictions.

There were also some suggestions specific to the GARP3 software produced within the project. With regard to the software packages available for building QR models prior to the project, GARP3 can now be used very intuitively to build QR models representing a prerequisite for the target, to motivate stakeholders and students to use the software and put their conceptual knowledge in causal models! To further enhance the modelling process itself, it could be helpful to always see the consequences of model definitions and implemented model fragments (configurations, proportionalities and influences) ‘on the fly’ in an accompanying window of the software (for example as they can be explored by the ‘show entities and configurations’ button, and by the ‘show dependencies’ button). It could also be helpful to have the full model shown in a window like the ‘show entities and configurations’ window with the opportunity to select parts of the model by hand to be run in a simulation (running only parts of the model by simply selecting parts of the model by drawing a window). Finally, to link the outcomes of causal models to a GIS would open a new field of promising applications!

7. Evaluation of Curriculum for learning about SD using QR

The curriculum for learning about Sustainable Development (SD) using Qualitative Reasoning (QR) is implemented in Moodle (http://moodle.org), an online course management system (see also D6.10 for details) (Nuttle and Bouwer, 2007). This curriculum is centred on QR models that explore SD situations in the context of the five case studies, in addition to other models that explore specific sustainability issues. The 'lesson' activity in Moodle is used to create a series of questions that guide learners stepwise through different scenarios for each model. Additionally, the 'Glossary' resource tool in Moodle is used to provide information that helps learners with help tips and definitions of SD and QR terms. Web pages in Moodle are used to present background information and suggest themes and learning routes for learners to explore the content of the curriculum.

We evaluated the structure of the Lesson activities for one of the modelled scenarios from one of the SD case studies. Each Lesson for each scenario is structured as follows:

1. Background about the case study;
2. Introduction and learning goals for the scenario;
3. Exploration of the system structure;
4. Exploration of causality of the system;
5. Exploration of the dynamics of the system;
6. Application to real-world situation and to sustainable development.

Hence, questions are organised to move from structure to causality to dynamics to evaluation, while also achieving the six levels of cognitive ability described in Bloom's Taxonomy (namely: knowledge, comprehension, application, analysis, synthesis, and evaluation) (Bloom, 1965; Krathwohl et al., 1973). Each question is accompanied by a hint link, which instructs the learner how to interact with the model in Garp3 to discover the answer to the question. The Moodle tool collects learners' typed answers to the questions and reports them to the instructor (if desired). Before moving on to the next question, learners are able to compare their answer with an answer provided by the instructor. This allows learners to assess whether they have understood the intended content before moving on to higher-level questions. Specific questions for the evaluated lesson are provided below. Students were also asked about their impressions of the lesson format and using QR models to learn about sustainability.

7.1. Setting

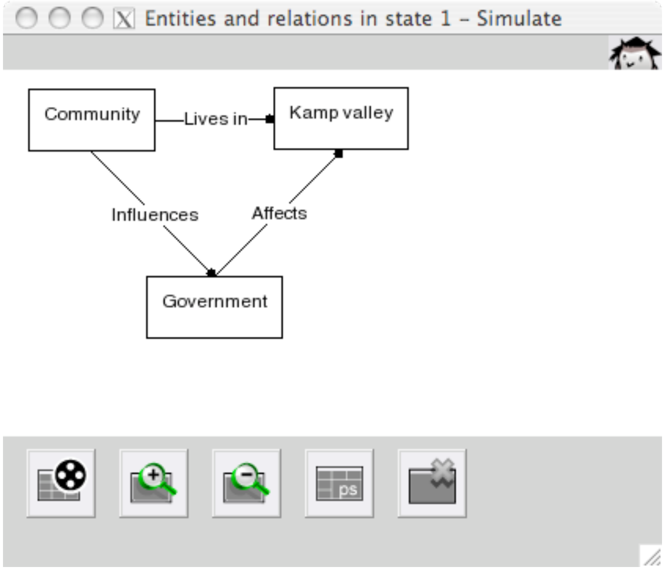
We evaluated one of the lessons on the Curriculum page, as implemented in Moodle, with a university Botany class of 19 students. The class was composed of students ranging from second- to fourth-year Biology majors at Indiana University of Pennsylvania, USA. Students gave informed, written consent to participate in the study. None of the students had previously had a course in Ecology or SD. Some of the students had worked with concept maps, but none had experience with QR models.

7.2. Methods

We presented the students with River Kamp Lesson 1. There was no access to separate computers for each student. Instead, the appropriate computer screen shots were displayed as if the students had run the model themselves, using a computer and video projector. Students were allowed to view the diagram and write their responses until it looked like most had finished writing. Then, the facilitator solicited student responses to the question. If it looked like they did not understand how to answer the question, the presenter showed the appropriate help tip (see section on ‘Hints’ below). After a few students had responded, the facilitator presented a prepared ‘correct’ answer before moving on to the next question. Here, we provide the questions and diagrams that were presented to the students.

Question 1.

- Explain in your own words the system being modelled by this scenario. *Hint0. Hint1.*

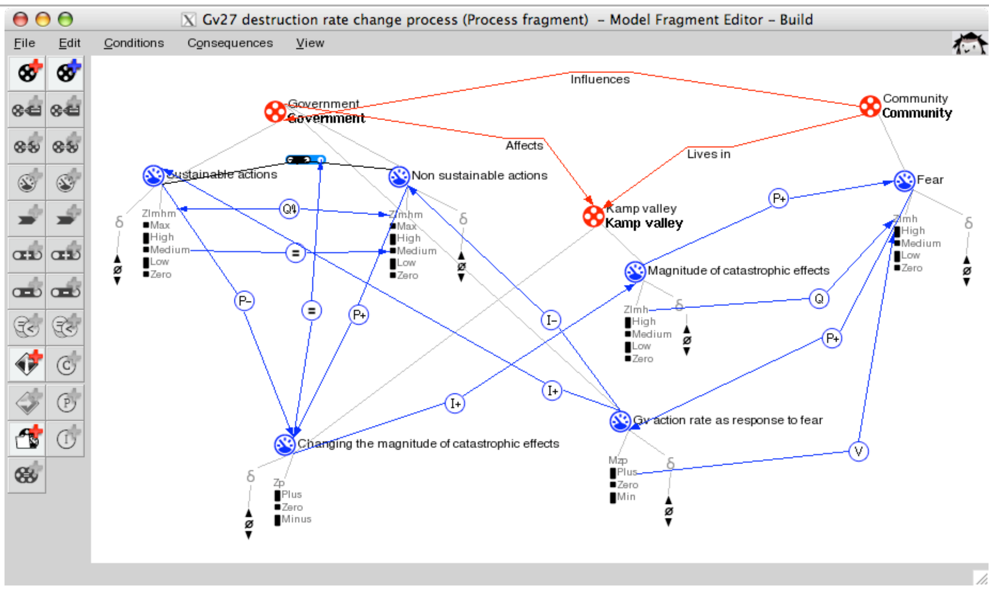


Answer to question 1.

- The system that is modelled in this scenario includes three entities, which are related in the following way. There is a community living in the Kamp valley. According to this scenario, the community influences the government, and the government affects the Kamp valley.

Question 2.

- Explain in your own words what processes and external influences are likely to be important in this system. *Hint2*.



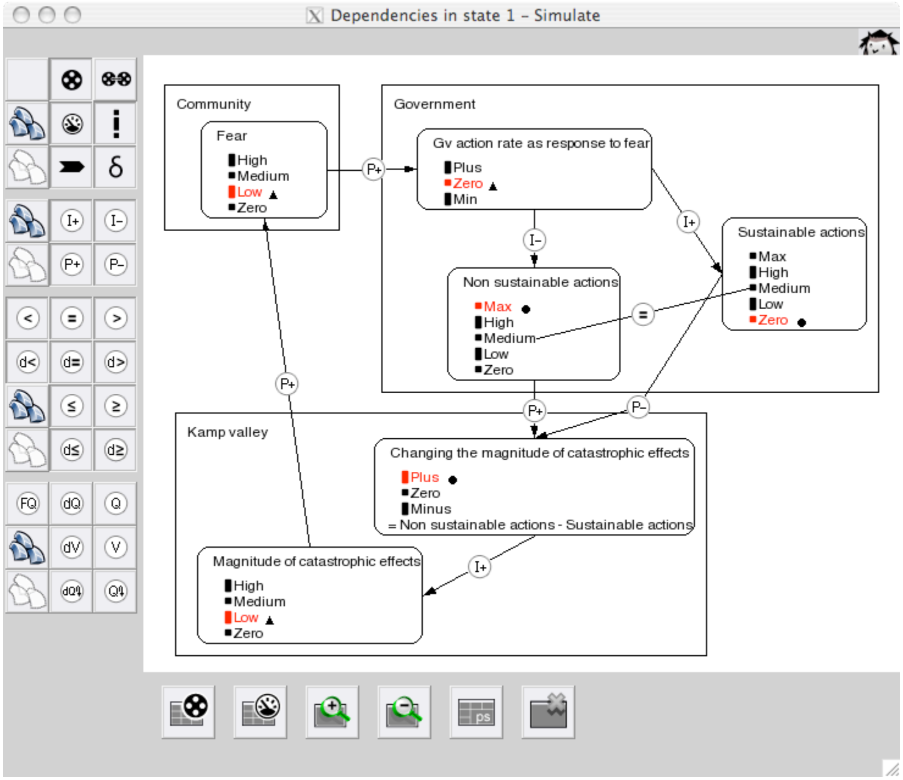
Answer to question 2.

- The Government action rate influences sustainable actions (I+) and non-sustainable actions (I-).

- The net effect of sustainable and non-sustainable actions determines the Change of Risk (notice the P+ and P- dependencies and the calculation in the model fragment view or the dependencies view).
- The Change of Risk influences the Magnitude of catastrophic effects (I+).
- The Magnitude of catastrophic effects propagates to Fear of the community (P+).
- Fear of the community propagates back to Government action rate (P+).
- The other information in the model fragment (The Q- and V-correspondences) helps to reduce the complexity of the resulting simulation.

Question 3.

- What are the state variables, the rates, and their starting magnitudes and derivatives for the initial state? What entities are the variables (quantities) associated to? *Hint3.*

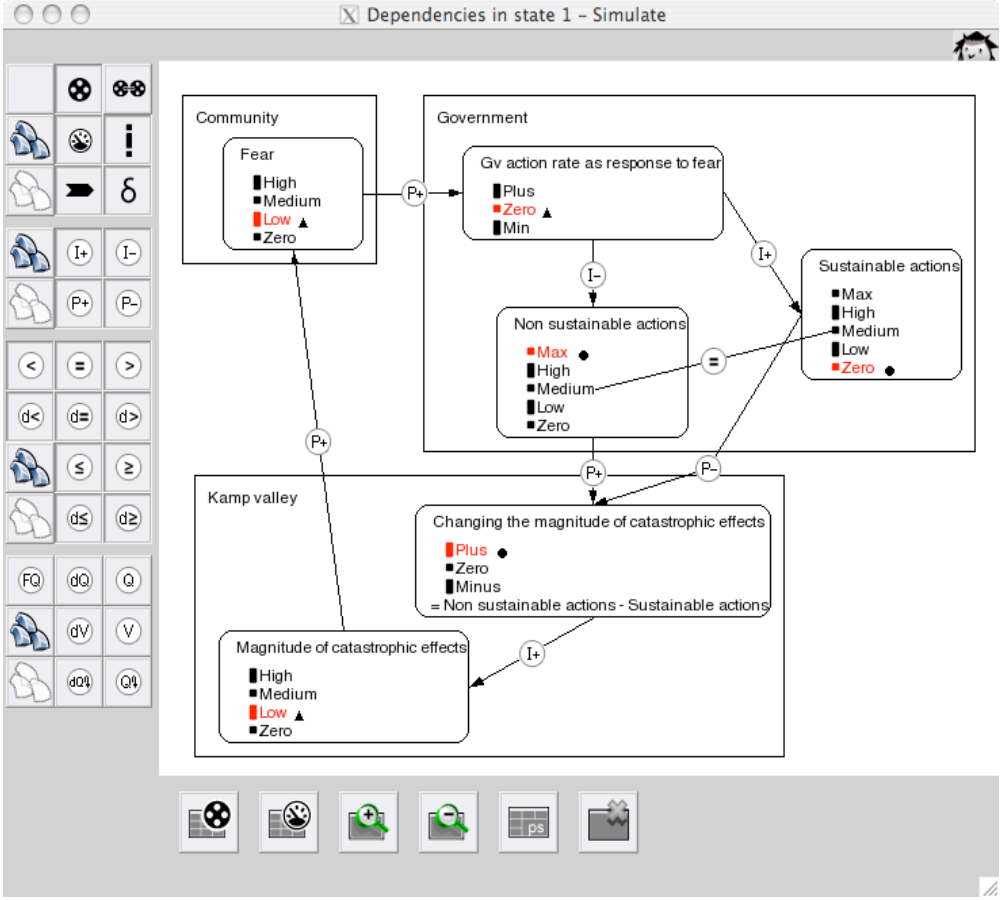


Answer to question 3.

- In the dependencies view for state 1, we see:
  - Rates with initial magnitudes and derivatives:
  - Gv action rate: Zero, increasing
  - Change of Risk: Plus, steady
  - State variables with initial magnitudes and derivatives:
  - Fear: Low, increasing
  - Sustainable actions: Zero, steady
  - Non sustainable actions: Max, steady
  - Magnitude of catastrophic effects: Low, increasing
- Which quantities belong to which entities?
  - Fear belongs to the entity Community
  - Gv action rate, Sustainable actions (I+) and Non-sustainable actions belong to the entity Government
  - Change of Risk and Magnitude of catastrophic effects belong to the entity Kamp valley

Question 4.

- Inspect each quantity's derivative (in state 1) and explain why it is increasing, decreasing or stable. *Hint4.*

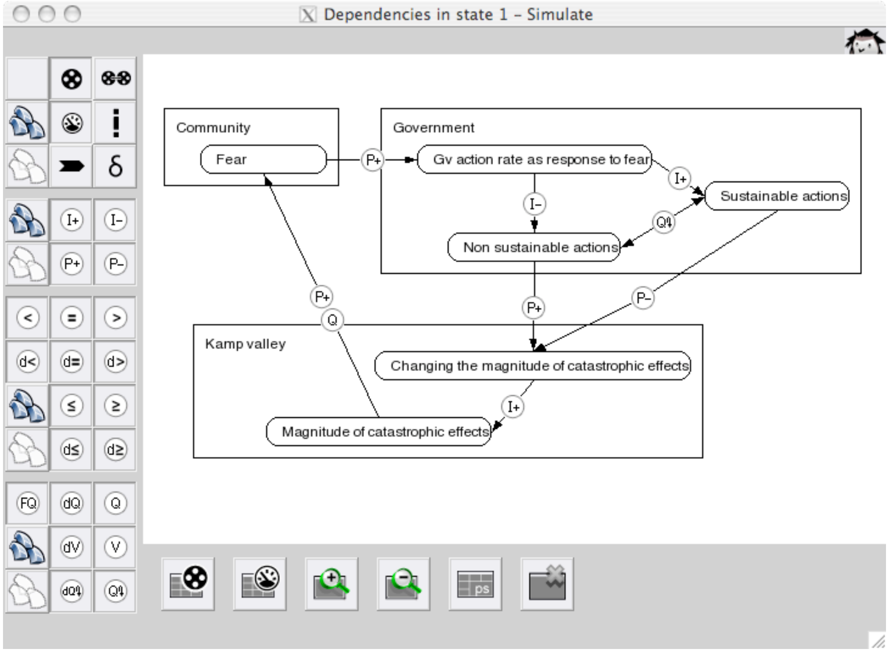


Answer to question 4.

- Inspecting the Dependencies view in state 1, we see:
  - Fear is increasing because Magnitude of catastrophic effects is increasing (P+).
  - Government action rate as response to fear is increasing because fear is increasing (P+)
  - Non-sustainable actions is currently steady because Government action rate is zero; therefore the negative influence (I-) currently does not have an effect.
  - Sustainable actions is also currently steady because Government action rate is zero; therefore the positive influence (I+) currently does not have an effect.
  - Change of Risk is currently steady because both quantities affecting it are also steady (P+/P-).
  - Magnitude of catastrophic effects is increasing because the quantity influencing it (I+) has a positive value.

Question 5.

- What are the focal quantities in the system? How can you tell? *Hint5.*

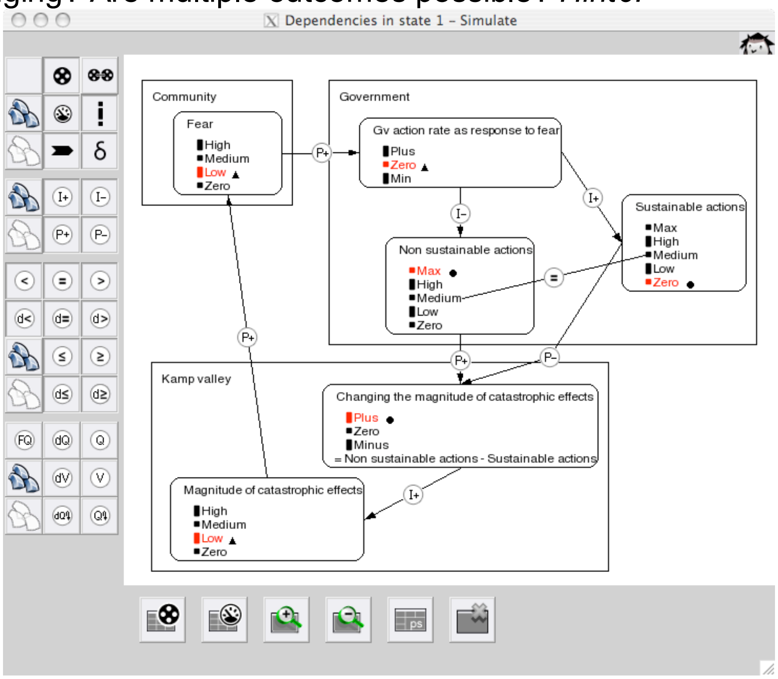


Answer to question 5.

- In this scenario, we see a feedback loop which results in cyclic behaviour. Therefore, all of the quantities can be considered important.
- Specifically, Government action rate and Change of Risk are the quantities that directly influence the Sustainable actions and Non-sustainable actions, and the Magnitude of catastrophic effects, respectively. Although its behaviour directly corresponds to the Magnitude of catastrophic effects, the quantity Fear of the community is also important because it provides a causal feedback loop to the Government's action rate.

Question 6.

- Predict what will happen to the focal quantities (your answer to question #5) and why. Will they stay the same, or will they reach another value, e.g., a maximum/minimum? Will a new equilibrium be reached, or might the system keep changing? Are multiple outcomes possible? *Hint6.*

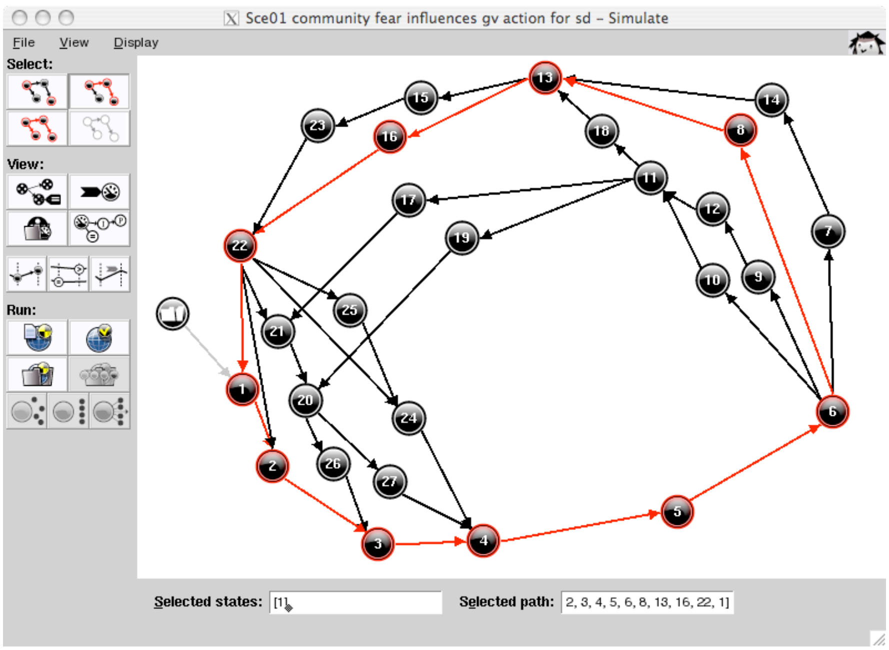


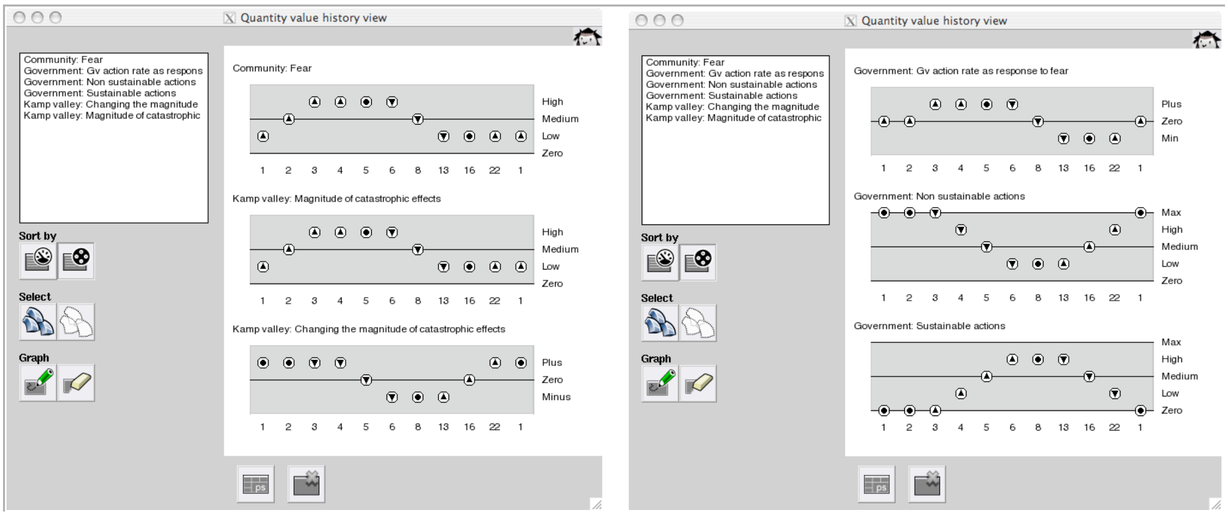
Answer to question 6.

- The quantity ‘Change of Risk’ is calculated as Non-sustainable actions (current value: Max) minus Sustainable actions (value: Zero). This difference is positive (value: Plus), which causes the magnitude of Catastrophic effects to increase (via the positive influence I+). This propagates to an increase of Fear. This propagates to an increase of Government action rate. The value of Government action rate will therefore become positive, which will positively influence Sustainable actions (I+), and negatively influence Non-sustainable actions (I-). Therefore, in the following states, Sustainable actions will increase and Non-sustainable actions will decrease. Change of Risk will decrease and magnitude of Catastrophic effects will stabilize, and then start to decrease. This will in turn propagate further, so that Government actions will decrease again, in which case the whole cycle starts again.

Question 7.

- Describe what the model predicts to happen in the system over time. *Hint7*.
    - In the state-transition graph, select state 1 and press the button ‘Select a path’.
        - Does the behaviour end at a certain state?
        - Describe what happens to the focal quantities during this behaviour path.
    - Now look at the state-transition graph again.
        - Are alternative behaviour paths possible?
        - Pick one of these alternative paths and describe how the behaviour differs from the first one.





Answer to question 7.

- When selecting state 1 and pressing the button ‘Select a path’, a cyclic path is selected, namely: [1, 2, 3, 4, 5, 6, 8, 13, 16, 22, 1]. In this cyclic path, the value of Fear and Magnitude of catastrophic effects go up to High first, and down again to Low. These values lag behind a few states after the Actions of the government, which also go up and down (or vice versa, for the Sustainable actions). This lag can be interpreted as the time that it takes for government actions to have an effect on the environment.
- Note that the actual maximum or minimum value reached by the various quantities differs for the various paths that are possible.
- For example, in the cyclic path [1, 2, 3, 4, 5, 6, 8, 13, 15, 23, 22, 1] (which you get when selecting state 23 in path selection mode), you’ll see that both the Fear and Magnitude of catastrophic effects reach Zero for a moment.
- In the path [1, 2, 3, 4, 5, 6, 10, 11, 19, 20, 26, 3] (which you get when you select state 20 in path selection mode), you’ll see that from state 3 onwards, the turning point occurs within the interval High, so that both the Fear and Magnitude of catastrophic effects is High all the time - not a very positive prediction!
- If you chose to describe a different path, that’s ok.

Question 8.

- Compare the model predictions (your answer to #7) with your predictions (your answer to #6). Do they match? What in your predictions is missing in the model predictions and vice versa?

Answer to question 8.

- Probably, your predictions were not as complete as those produced by running the full simulation. Perhaps you found some surprising results. Here is a checklist of things to consider when comparing your own predictions with the model’s predictions:
    - Did you predict that Fear and Magnitude of catastrophic effects would go up at first?
    - Did you also predict that Fear and Magnitude of catastrophic effects would go down again later?
    - Did you consider that cyclic behaviour was possible?
    - Did you predict the time lag between Government action and Magnitude of catastrophic effects (and fear)?
    - Did you consider that different values could be reached as a maximum/minimum by the various quantities?

Question 9.

- Now that you have inspected the simulation, consider how this relates to the real-life situation as described in the introduction. More specifically,
    - Where in the cycle do you think the system was prior to the flood that happened in 2002, in terms of values for Fear and Government action rate?
    - Where is it now?
    - Can you give some examples of Sustainable actions and Non-sustainable actions?
    - Can you think of ways to maintain a High level of Sustainable actions and a Low level of Non-sustainable actions even when Fear of catastrophes is Low?

Answer to question 9.

- Where in the cycle do you think the system was prior to the flood that happened in 2002, in terms of values for Fear and Government action rate?
    - Fear: low, Action rate: zero/minus
- Where is it now?
    - Fear: high, Action rate: plus
- Can you give some examples of Sustainable actions and Non-sustainable actions?
    - Sustainable actions: planting forested buffer areas, building only on high ground, removing unnecessary dams
    - Non-sustainable actions: cutting forest in buffer areas, building on low ground, canalising the river
- Can you think of ways to maintain a High level of Sustainable actions and a Low level of Non-sustainable actions even when fear of catastrophes is low?
    - This requires a continuous lobbying and public support to attract the attention of decision makers for the importance of sustainable actions

Question 10.

- What do you think people can learn from studying this lesson and scenario (including new facts, ideas, or ways of thinking about things)?

Answers to question 10

- Were open-ended.

Hints

Hint0: In the model building environment, open the scenario and arrange the contents if you need to (don't change anything else!). The scenario describes the system of interest. Try completing this sentence: 'The scenario investigates what happens if...'

Hint1: Run the simulation to the initial states. When you do this, the program looks in the digital library for information that matches what is in the scenario and assembles all the pieces of information together to create a more complete model of the system.
Next, select one of the initial states and open the Entity Relations view. Arrange the contents so you can see everything clearly and describe what you see. How is the information described here similar/different to that contained in the scenario?

Hint2: With the initial state still selected, open the model fragments list to see what model fragments fire. Open the model fragment editor for each model fragment to see which ones contain causal dependencies (P, I). These will describe the main processes. Agents, or external influences, will be evident by their specific agent icons, and also be connected via a causal dependency to other quantities.

Hint3: You can double click the state to display values (magnitudes and derivatives) for all quantities in a state or you can open the dependencies view and choose the button to display values (quantity magnitudes and derivatives).

- Rates often have the word ‘rate’ in them, or describe some action (generally via direct influences).
- State variables are the remaining variables, and are generally an amount of something.

Hint4: Look at the causal dependencies in the dependencies view. Make sure I’s and P’s are displayed along with the values quantity magnitudes and derivatives.

Hint5: Look at the causal chain (or chains) in the dependencies view. Are there any quantities at the end of a causal chain or that are influenced by several dependencies? Also, what (if any) quantities are influenced by agents?

Hint6: Reason through the causal dependencies (including magnitudes and derivatives) to see how they cause and propagate change through the system.

Hint7: Run the full simulation. You may need to spend some time arranging the state graph to optimise the layout of states and transitions (circles and arrows). Note the general nature of the state graph (cycles, branching paths, etc.). Select a path starting with the initial state. View the value history, focusing on what happens to the state variables. The causes of any changes in quantity values are due to causal dependencies (influences), and the net effect of ‘competing’ influences on any one quantity depends on which influence is larger. To see which influence is larger, view the Equation History for the selected path.
Repeat for several paths (if there are more than one).

7.3. Results

Based on the written responses to questions in the lesson, all students made a good effort to answer the questions to the best of their ability, despite some technical difficulties with reading the displayed diagrams from across a large room. Some students at first had difficulty grasping the meaning of some of the model diagrams during the first few questions. For example, instead of basing their answers on the diagram for question 2, several students based their answers on their background knowledge about natural catastrophes. This may have been due to their inability to access the help tips that would ordinarily be available if they were working on their own computers, which would guide the students in how to go about answering the question based on what is presented in the model. However, once they had the opportunity to hear responses of other students and the class went over the ‘correct’ answer provided in the Lesson, most were able to catch up and were able to understand and appropriately respond to later questions. Students successfully used the diagrams to reason about a complex system involving causal feedback loops and multiple possible outcomes, including cyclic behaviour. For example, many students correctly predicted that multiple outcomes were possible and that the system would never stabilize (questions 6, 7, and 8). Finally, students responded for questions 9 and 10 that people could learn the government should institute policies that maintain sustainable actions even when community fear level is low. They also named several types of sustainable actions (e.g., maintaining natural ecosystems in the watershed to reduce flood levels) and unsustainable actions (e.g., building too close to the river or in areas are that lie too low).

Students were also asked about their impressions of the lesson format and using QR models to learn about sustainability. Many students responded that they thought the simulation model provides useful support for learning about the behaviour of complex systems. Several students suggested more use of colour in the presentation, perhaps to hold their attention better or draw their attention to specific parts of the various diagrams. Again, when students interact with the models themselves on their own computers, this will probably be less of an issue because they will be moving screens around, be able to view photographs from other pages if they want, and can access help tips to draw their attention to the focal points of each diagram.

7.4. Conclusion and Discussion

Overall, students responded positively to the experience and several expressed interest (orally or in their written evaluations) in participating in future activities using the online curriculum. We are currently planning additional evaluation of the curriculum using this and other groups of students, where each student will have the opportunity to complete lessons on individual computer work stations.

8. Conclusions

This report has presented a comprehensive set of evaluation studies within the NaturNet-Redime project addressing the software developed in work-package 4, and the case studies and curriculum developed in work-package 6. The results confirm the findings obtained in earlier studies (e.g., Bouwer et al., 2006; Nuttle et al., 2006; Liem et al., 2007). The main points can be summarised as follows.

- The Garp3 workbench scores well on usability factors (Section 3), and learners in formal education easily learn about Sustainable Development (SD) issues when interacting with Qualitative Reasoning (QR) Models using the Garp3 workbench (section 2 and 4).
- Different groups of stakeholders are impressed and react highly positive when evaluating the potential and usefulness of QR models for performing their tasks (Section 5 and 6).

It seems fair to conclude that due to the NaturNet-Redime project a lot of progress has been made on making QR technology available and usable for stakeholders dealing with SD. However, still more can be accomplished. Future research could further simplify the process of adequately capturing conceptual knowledge using QR technology. In addition, more qualitative knowledge on SD issues needs to be captured in order to be able to use it. Finally, the technology could be further enhanced to facilitate easy switching between ‘levels of detail’ and ‘points of view’ on the subject matter to accommodate users better in continuously changing circumstances, such as collaborative decision-making processes.

Acknowledgements

All authors and evaluators would like to thank the users who participated in the evaluation studies. Particularly, the authors of the Section 5, Paulo Salles, Gisele Morisson Feltrini, Ana Luiza Rios Caldas and Monica Maria Pereira Resende, would like to thank the stakeholders who participated of the dissemination and evaluation activities. The authors are also grateful to the Centro Educacional 6, the Centro Educacional Ave Branca, the Brasilia Water and Sewage Company (CAESB) and the University of Brasilia for their support to the activities of the NaturNet-Redime project.

9. References

Bertels, D. (2007). Ontwikkeling van een visueel vocabulaire voor modelprimitieven en operaties in Garp3, MSc Thesis, University of Amsterdam, Amsterdam, The Netherlands.

Bloom B. S. (1956). Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. New York: David McKay Co Inc.

Bouwer, A., Bredeweg, B., Liem, J., and Salles, P. (2006). 1st NNR user group workshop on using QR technology, Naturnet-Redime, STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2006), Project no. 004074, Project Deliverable Report D7.1.1.

Bredeweg, B., Salles, P., Bouwer, A., Liem, J., Nuttle, T., Cioaca, E., Nakova, E., Noble, R., Caldas, A. L. R., Uzunov, Y., Varadinova, E. and Zitek, A. (2007). Towards a structured approach to building qualitative reasoning models and simulations. Ecological Informatics (in press).

Bredeweg, B., Bouwer, A., and Liem, J. (2006a). Single-user QR model building and simulation workbench , Naturnet-Redime, STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2006), Project no. 004074, Project Deliverable Report D4.1.

B. Bredeweg, A. Bouwer, J. Jellema, D. Bertels, F. Linnebank, and J. Liem. (2006b). Garp3 - A new Workbench for Qualitative Reasoning and Modelling. 20th International Workshop on Qualitative Reasoning (QR'06), C. Bailey-Kellogg and B. Kuipers (eds), pages 21-28, Hanover, New Hampshire, USA, 10-12 July.

Bredeweg, B., Salles, P., Bouwer, A., and Liem, J. (2005). Framework for conceptual QR description of case studies, Naturnet-Redime, STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2006), Project no. 004074, Project Deliverable Report D6.1.

Krathwohl, D. R., Bloom, B. S., and Masia, B. B. (1973). Taxonomy of Educational Objectives, the Classification of Educational Goals. Handbook II: Affective Domain. New York: David McKay Co., Inc.

Konikow, L.F. and Bredehoeft, J.D. (1992). Ground water models cannot be validated, *Advances in water resources* 15: 75-83.

Lee, L. and Ghanime, L. (2004). Millennium Development Goal 7 Summary Review: 67 Country MDG Reports. Bureau for Development Policy Energy and Environment Group.

Liem, J., Bouwer, A. and Bredeweg, B. (2006). Collaborative QR model building and simulation workbench, Naturnet-Redime, STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2006), Project no. 004074, Project Deliverable Report D4.3.

Liem, J., Bouwer, A., Bredeweg, B., Salles, P. and Bakker, B. (2007). 3rd NNR user group workshop on Qualitative Reasoning and Modelling, Naturnet-Redime, STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2006), Project no. 004074, Project Deliverable Milestone M7.2.3.

Nuttle, T. and Bouwer, A. (2007). Curriculum for learning about Sustainable Development using Qualitative Reasoning, NaturNet-Redime, STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2007), Project no. 004074, Project Deliverable Report D6.10.

Nuttle, T., Liem, J., Bouwer, A., Bredeweg, B., and Salles, P. (2006). 2nd NNR user group workshop on Qualitative Reasoning and Modelling, Naturnet-Redime, STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2006), Project no. 004074, Project Deliverable Report D7.1.2.

Oreskes, N., Shrader-Frechette, K. and Belitz, K. (1994). Verfication, validation and confirmation of numerical models in the earth sciences, *Science*, 263: 641-646.

Rafalowicz, J. (2007). Evaluating Qualitative Reasoning as a Support Tool for the Transfer of Conceptual Knowledge, MSc Thesis, University of Amsterdam, Amsterdam, The Netherlands.

Rykiel, E.J. (1996). Testing ecological models: the meaning of validation, Ecological Modelling, 90(3): 229–244.

Salles, P. (2007). *QR models and documentation for learning about sustainable development, focusing on basic biological, physical, and chemical processes related to the environment in the Riacho Fundo*, Naturnet-Redime, STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2006), Project no. 004074, Project Deliverable Report D6.4.2.

Salles, P., Bredeweg, B., Caldas, A.L.R. and Nuttle, T. (2007). Modelling sustainability in the Riacho Fundo water basin (Brasília, Brazil). In Chris Price (ed.) *Proceedings of the 21st International Workshop on Qualitative Reasoning (QR'07)* Aberystwyth, (Wales, U.K.), 26-28 June, 2007, pages 147-160.

Salles, P. and Caldas, A.L.R. (2006). *Textual description of Riacho Fundo case study focusing on basic biological, physical, and chemical processes related to the environment*, Naturnet-Redime, STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2006), Project no. 004074, Project Deliverable Report D6.4.1.

Salles, P. and Bredeweg, B. (2003). Qualitative Reasoning about Population and Community Ecology, AI Magazine, 24(4): 77-90.

Salles, P., Bredeweg, B. and Nuttle, T. (2005). Qualitative Models of Indicators for Environmental Sustainability Related to the Millennium Development Goals. In Proceedings of 2nd MONET Workshop on Model-Based Systems (MONET 05) at the 19th International Joint Conference on Artificial Intelligence (IJCAI-05), pages 66–72.

Salles, P. (2001). Comitê de Gestão da Bacia do Paranoá. In: Fonseca, F.O. (org.) *Olhares sobre o Lago Paranoá*. Brasília: Secretaria de Meio Ambiente e Recursos Hídricos, pages 296-307.

Sargent, R. (2004). Validation and verification of simulation models. *In* Ingalls, R.G., Rossetti, M.D., Smith, J.S. and Peters, B.A. (eds.). *Proceedings of the 2004 Winter Simulation Conference*, IEEE.

Sharp, H., Rogers, Y. and Preece, P. 2007. Interaction Design: Beyond Human-computer Interaction by (2nd edition), Wiley & Sons, Ltd.

Smeets, E. and Weterings, R. (1999). Environmental indicators: Typology and overview. Technical report, TNO Centre for Strategy, Technology and Policy, The Netherlands.

United Nations Development Group (2001). Reporting on the Millennium Development Goals at the Country Level. Guidance Note.

White, B.Y. and Frederiksen, J.R. (1990). Causal model progressions as a foundation for intelligent learning environments, Artificial Intelligence 42: 99–157.

Zitek, A., Muhar, S., Preis, S., Schmutz, S., Bredeweg, B. and P. Salles, P. (2007). QR models and documentation for learning about sustainable development, focusing on basic ecological and socio-economic features for an integrative and sustainable development of the riverine landscape of the Kamp valley. Naturnet-Redime, STREP project co-funded by the European Commission within the Sixth Framework Programme (2002-2006), Project no. 004074, Project Deliverable Report D6.6.2.

Appendix 1

Garp3 – Visual language questionnaire

Investigating the meaning of icons

[Location] – [Date]

In computer software icons are often used to create a graphical interface. The icons refer to actions the software will perform when pressing the icon with the mouse (thus by clicking on the icon). With this questionnaire we want to investigate the meaning of icons. You are therefore asked to select the most appropriate meaning for each icon enumerated below. You should select your answer out of the four options given for each icon. Give your answers by writing/typing '1', '2', '3', or '4' in the boxes on the right hand side.

Be aware that we will be using these icons for building and simulating qualitative models. The following main tasks are typical for that:

- File manipulations (opening, saving, etc.)
- Building the model
- Simulating the model
- Inspecting the simulation results

Notice that answers are 'good' or 'wrong'. The goal is to find out the ideas people associate to icons.

Thank you very much for participating!

Please fill out:

Name:

Date:

Choose one answer for each of the following questions:

Question 1

Icon	What action do you associate with this icon?	Your answer:
	1. Delete selected item 2. Remove above 3. Undo changes 4. Layout list	

Question 2

Icon	What action do you associate with this icon?	Your answer:
	1. Open model fragments editor 2. Find successors for selected states 3. Save current model to a different file 4. Start a new model	

Question 3

Icon	What action do you associate with this icon?	Your answer:
	1. Open quantity space editor 2. Undo changes 3. Open configurations definitions editor 4. Split interval	

Question 4

Icon	What action do you associate with this icon?	Your answer:
	1. Open model fragments editor 2. Draw 3. Add a quantity 4. Close model fragments editor	

Question 5

Icon	What action do you associate with this icon?	Your answer:
	1. Add selected item to list 2. Add value low 3. Open agent hierarchy editor 4. Close current model	

Question 6

Icon	What action do you associate with this icon?	Your answer:
	1. Show item properties 2. Simulate ‘current’ scenario 3. Full simulation ‘current’ scenario 4. Select all items	

Question 7

Icon	What action do you associate with this icon?	Your answer:
	1. Show value history 2. Open assumption hierarchy editor 3. Show transition history 4. List by quantities	

Question 8

Icon	What action do you associate with this icon?	Your answer:
	1. Open quantity definitions editor 2. Select a path 3. Select all states 4. Layout hierarchy: vertical	

Question 9

Icon	What action do you associate with this icon?	Your answer:
	1. Full simulation ‘current’ scenario 2. Open model fragment editor 3. Clear screen 4. Select a scenario to simulate	

Question 10

Icon	What action do you associate with this icon?	Your answer:
	1. Terminate selected states 2. Open model fragment editor 3. Change layout entities 4. Find successors for selected states	

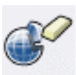
Question 11

Icon	What action do you associate with this icon?	Your answer:
	1. Select all items 2. Add new item 3. Open model from file 4. Save changes to model	

Question 12

Icon	What action do you associate with this icon?	Your answer:
	1. Start a new model 2. Copy selected item 3. Combine intervals 4. Show long names	

Question 13

Icon	What action do you associate with this icon?	Your answer:
	1. Show transition history 2. Save changes to model 3. Cancel changes 4. Clear screen	

Question 14

Icon	What action do you associate with this icon?	Your answer:
	1. Zoom out 2. Zoom in 3. Save changes to model 4. Open de simulator to its current state or a saved simulation	

Question 15

Icon	What action do you associate with this icon?	Your answer:
	1. Change layout entities 2. List by entities 3. Clear screen 4. List by quantities	

Question 16

Icon	What action do you associate with this icon?	Your answer:
	1. Print diagram to file (postscript) 2. Show dependencies 3. Add value high 4. Combine intervals	

Question 17

Icon	What action do you associate with this icon?	Your answer:
	1. Open quantity definitions editor 2. Copy selected item 3. Start a new model 4. Close window	

Question 18

Icon	What action do you associate with this icon?	Your answer:
	1. Delete value low 2. Show dependencies 3. Delete dependencies 4. Delete value high	

Question 19

Icon	What action do you associate with this icon?	Your answer:
	1. Show model fragment in context 2. Select all states 3. Layout hierarchy: vertical 4. Select all items	

Question 20

Icon	What action do you associate with this icon?	Your answer:
	1. Add value high 2. Delete quantities 3. List by quantities 4. Change layout quantities	

Question 21

Icon	What action do you associate with this icon?	Your answer:
	1. Clear screen 2. Start a new model 3. Terminate selected states 4. Add new item	

Question 22

Icon	What action do you associate with this icon?	Your answer:
	1. Show quantity values 2. Clear screen 3. Open quantity space editor 4. Select all items	

Question 23

Icon	What action do you associate with this icon?	Your answer:
	1. Find successors for selected states 2. Add new item 3. Edit current / selected scenario 4. Open quantity space editor	

Question 24

Icon	What action do you associate with this icon?	Your answer:
	1. Find successors for selected states 2. Cancel changes 3. Save changes to model 4. Start a new model	

Question 25

Icon	What action do you associate with this icon?	Your answer:
	1. Delete value high 2. Add value high 3. Add selected item to list 4. Split interval	

Question 26

Icon	What action do you associate with this icon?	Your answer:
	1. Show long names 2. Add selected item to list 3. Draw 4. Close current model	

Question 27

Icon	What action do you associate with this icon?	Your answer:
	1. Delete value low 2. Terminate selected states 3. Delete value high 4. Add new item	

Question 28

Icon	What action do you associate with this icon?	Your answer:
	1. Edit current / selected scenario 2. Start a new model 3. Show transition history 4. Find successors for selected states	

Question 29

Icon	What action do you associate with this icon?	Your answer:
	1. Save changes to model 2. Delete quantity 3. Show quantity values 4. Save changes to model	

Question 30

Icon	What action do you associate with this icon?	Your answer:
	1. Add new item 2. Zoom in 3. Open quantity space editor 4. Close window	

Question 31

Icon	What action do you associate with this icon?	Your answer:
	1. Change layout quantities 2. Start a new model 3. Clear screen 4. Close current model	

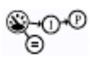
Question 32

Icon	What action do you associate with this icon?	Your answer:
	1. Delete value low 2. Add value high 3. Find successors for selected states 4. Cancel changes	

Question 33

Icon	What action do you associate with this icon?	Your answer:
	1. Layout hierarchy: vertical 2. Terminate selected states 3. Show model fragment in context 4. Open assumption hierarchy editor	

Question 34

Icon	What action do you associate with this icon?	Your answer:
	1. Open quantity definitions editor 2. Select a path 3. Select all states 4. Show dependencies	

Question 35

Icon	What action do you associate with this icon?	Your answer:
	1. Open quantity definitions editor 2. Combine intervals 3. Show long names 4. Zoom in	

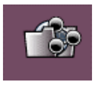
Question 36

Icon	What action do you associate with this icon?	Your answer:
	1. Cancel changes 2. Clear screen 3. Show dependencies 4. Print diagram to file (postscript)	

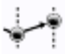
Question 37

Icon	What action do you associate with this icon?	Your answer:
	1. Change layout entities 2. Open model from file 3. Zoom in 4. List by entities	

Question 38

Icon	What action do you associate with this icon?	Your answer:
	1. Copy selected item 2. Find successors for selected states 3. Start a new model 4. Open de simulator to its current state or a saved simulation	

Question 39

Icon	What action do you associate with this icon?	Your answer:
	1. Open model from file 2. Add selected item to list 3. Show transition history 4. Select a path	

Question 40

Icon	What action do you associate with this icon?	Your answer:
	1. Start a new model 2. Terminate selected states 3. Clear screen 4. Close current model	

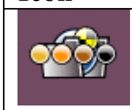
Question 41

Icon	What action do you associate with this icon?	Your answer:
	1. Select all items 2. Select a scenario to simulate 3. Open assumption hierarchy editor 4. Find successors for selected states	

Question 42

Icon	What action do you associate with this icon?	Your answer:
	1. Show item properties 2. Open agent hierarchy editor 3. Show quantity values 4. Open model fragment editor	

Question 43

Icon	What action do you associate with this icon?	Your answer:
	5. Select a scenario to simulate 6. Select a path 7. Full simulation ‘current’ scenario 8. Add new item	

Question 44

Icon	What action do you associate with this icon?	Your answer:
	1. Add value low 2. Open quantity definitions editor 3. Change layout quantities 4. Select all items	

Question 45

Icon	What action do you associate with this icon?	Your answer:
	1. Select a scenario to simulate 2. Draw 3. Open assumption hierarchy editor 4. Show details	

Question 46

Icon	What action do you associate with this icon?	Your answer:
	1. Show item properties 2. Save current model to a different file 3. Show value history 4. Add new item	

Question 47

Icon	What action do you associate with this icon?	Your answer:
	1. Open configurations definitions editor 2. Add value low 3. Terminate selected states 4. Delete value low	

Question 48

Icon	What action do you associate with this icon?	Your answer:
	1. Clear screen 2. Select all states 3. Draw 4. Delete selected item	

Question 49

Icon	What action do you associate with this icon?	Your answer:
	1. Combine intervals 2. Remove above 3. Split interval 4. Open quantity definitions editor	

Question 50

Icon	What action do you associate with this icon?	Your answer:
	1. Edit current / selected model fragment 2. Cancel changes 3. Close model fragments editor 4. Layout list	

Appendix 2

Tests on Deforestation (DEF) and Fuel and Global Warming (FGW)

(Answers are in italic)

Test A

1. When deforestation happens,
    a. land with no vegetation increases.
    b. water reservoir increases.
    c. production of new food and medicines increases.
    d. GDP (wealth) increases.

2. When the production of smoke in poor households increases,
    a. the production of greenhouse gases increases.
    b. the use of petroleum for industry increases.
    c. the petroleum available in global economy increases.
    d. the occurrence of chronic respiratory diseases increases.

3. When land with vegetation decreases,
    a. production of new food and medicines increases.
    b. erosion decreases.
    c. GDP (wealth) decreases.
    d. water reservoir increases.

4. When biodiversity increases,
    a. GDP (wealth) decreases.
    b. production of new food and medicines increases.
    c. erosion increases.
    d. agricultural production decreases.

5. The occurrence of chronic respiratory diseases increases because
    a. the petroleum available in the global economy increases.
    b. the use of petroleum for transportation increases.
    c. the use of solid fuel in poor households increases.
    d. there is a surplus of petroleum in the global economy.

6. New food and medicines increases because
    a. biodiversity increases.
    b. GDP (wealth) increases.
    c. agricultural production decreases.
    d. deforestation is active.

7. When erosion increases,
    a. uses of water increase.
    b. (GDP) wealth increases.
    c. amount of removed soil decreases.
    d. population without water increases.

8. When the use of petroleum for industry increases,
    a. the overall atmosphere temperature decreases.
    b. the occurrence of generic respiratory diseases increases.
    c. the use of solid fuel in poor households increases.
    d. the use of petroleum for transportation decreases.

9. Agricultural production decreases because
    a. erosion decreases.
    b. land with no vegetation decreases.
    c. deforestation stopped.
    d. removed soil increases.

10. The use of petroleum in poor households decreases because
    a. the overall atmosphere temperature increases.
    b. the use of petroleum for transportation increases.
    c. there is a shortage of petroleum in the global economy.
    d. the occurrence of generic respiratory diseases increases.

11. The water reservoir decreases because
    a. erosion increases.
    b. deforestation stopped.
    c. amount of removed soil decreases.
    d. land with no vegetation decreases.

12. The overall atmosphere temperature increases because
    a. the use of petroleum for industry decreases.
    b. there is a surplus of petroleum in the global market.
    c. the occurrence of chronic respiratory diseases increases.
    d. the use of solid fuel in poor households increases.

13. Population without water increases because
    a. land with no vegetation decreases.
    b. uses of water increase.
    c. erosion increases.
    d. amount of removed soil decreases

14. GDP (wealth) decreases because
    a. uses of water increase.
    b. biodiversity increases.
    c. erosion is decreasing
    d. land with no vegetation increases.

15. The use of solid fuel in poor households decreases because
    a. the petroleum available in the global economy increases.
    b. the occurrence of chronic respiratory diseases increases.
    c. the overall atmosphere temperature decreases.
    d. the use of petroleum for industry decreases.

Test B

1. When deforestation stops,
   a. agricultural production increases.
   b. biodiversity decreases.
   c. erosion increases.
   d. uses of water decrease.

2. The occurrence of generic respiratory diseases decreases because
   a. the production of greenhouse gases increases.
   b. the overall atmosphere temperature decreases.
   c. the production of smoke in poor households decreases.
   d. the use of petroleum for transportation increases.

3. When land with no vegetation increases,
   a. uses of water increase.
   b. removed soil increases.
   c. new food and medicines decrease.
   d. wealth (GDP) increases.

4. When biodiversity decreases
   a. soil erosion decreases.
   b. new food and medicines decreases.
   c. population without water decreases.
   d. uses of water increase.

5. The occurrence of chronic respiratory diseases decreases because
   a. the petroleum available in the global economy increases.
   b. the overall atmosphere pollution decreases.
   c. the occurrence of generic respiratory diseases decreases.
   d. the use of petroleum for industry increases.

6. When new food and medicines decreases,
   a. biodiversity increases.
   b. GDP (wealth) increases.
   c. deforestation is active.
   d. agricultural production increases.

7. When erosion decreases,
   a. amount of removed soil increases.
   b. uses of water decrease.
   c. water reservoir decreases.
   d. agricultural production increases.

8. The production of smoke in poor households decreases because
   a. the use of petroleum for transportation decreases.
   b. the use of solid fuel in poor households increases.
   c. the occurrence of chronic respiratory diseases increases.
   d. the overall atmosphere temperature decreases.

9. When removed soil decreases,
    a. GDP (wealth) decreases.
    b. agricultural production increases.
    c. uses of water increase.
    d. erosion decreases.

10. When the petroleum available in the global economy increases,
    a. the overall atmosphere temperature increases.
    b. the use of petroleum for transportation decreases.
    c. the use of solid fuel in poor households increases.
    d. the production of smoke in poor households increases.

11. When uses of water decrease,
    a. erosion decreases,
    b. agricultural production increases.
    c. GDP (wealth) decreases.
    d. new food and medicines decreases.

12. When the use of petroleum for industry decreases,
    a. the production of smoke in poor households decreases.
    b. the production of greenhouse gases decreases.
    c. the use of petroleum for transportation increases.
    d. the occurrence of chronic respiratory diseases decreases.

13. When population without water decreases,
    a. agricultural production decreases.
    b. biodiversity decreases.
    c. land with no vegetation decreases.
    d. GDP (wealth) decreases.

14. GDP (wealth) increases because
    a. water reservoir decreases.
    b. uses of water decrease.
    c. new food and medicines decreases.
    d. erosion decreases.

15. When the use of petroleum in poor households increases,
    a. the occurrence of chronic respiratory diseases decreases.
    b. the production of greenhouse gases decreases.
    c. the use of solid fuel in poor households increases.
    d. the use of petroleum for industry decreases.

Treatment assignments

(Answers are in italic)

1. Impact on Vegetation

1. Which quantity is influenced negatively by deforestation?
   Land with vegetation
2. If land with vegetation decreases, what will happen to biodiversity?
   Decreases
3. What is the value of biodiversity in state 1 and 4?
   Large and zero
4. Which quantity is increasing?
   Land without vegetation

2. Impact on Food and Medicines

1. The indirect effect of deforestation on biodiversity via land with vegetation is negative. Is the effect via land without vegetation also negative, or is it positive?
   Also negative
2. Will the production of new food and medicine increase or decrease because of deforestation?
   Decrease
3. In which state is the value of biodiversity equal to medium?
   State 2
4. What is the value of food and medicine in that state? And deforestation?
   Food and medicine = medium, Deforestation = plus

3. Impact on Land

1. If there is deforestation, will erosion increase or decrease?
   Increase
2. Does agricultural production increase or decrease because of erosion?
   Decrease
3. In which state is deforestation equal to zero?
   State 4
4. What is the value of land with vegetation in that state?
   Zero

4. Impact on Land and Water

1. Via which two quantities does erosion affect uses of water?
   Removed soil and water reservoirs
2. Does agricultural production have an effect on uses of water according to this model?
   No, not according to this model
3. When removed soil = medium, is agricultural production increasing or decreasing?
   Decreasing
4. And when removed soil = max, is agricultural production increasing or decreasing?
   Steady

5. Impact on Land, Water and Human

1. How should water reservoirs change to make the population without (access to) water decrease?

Water reservoirs has to increase

2. If there is erosion, what will happen to water reservoirs?
   Decrease
3. When removed soil is small, what is the value of water reservoirs?
   Large
4. Is it true that population without water changes in the same direction as water reservoirs?
   No

6. Impact on GDP (Wealth)

1. What are the three quantities affecting GDP (human wealth) in this model?
   New food and medicine, agricultural production, and uses of water
2. Are these quantities influenced by deforestation, or by erosion?
   Food and med, agricultural production and uses of water are al l influenced by deforestation
3. When biodiversity decreases, what happens to the production of new food and medicine?
   Decrease
4. What is the value of GDP (human wealth) in the end?
   Zero

7. Global Economy

1. What are the entities included in the model?
   Global economy, Transportation, Industry, Households, Atmosphere and Human
2. What type of relation is established between 'global market' and 'petroleum available'?
   A direct influence (I+)
3. What are the values of 'global market' in states [19, 16, 17]?
   (surplus, decreasing); (shortage, stable) and (surplus, stable), respectively
4. How can the changes in the values of 'global market' be described?
   A cyclic behaviour, oscillating between 'surplus' and 'shortage'

8. Uses of petroleum

1. Which quantities are positively influenced by 'petroleum available'?
   'Use of petroleum' for industry, for transportation, and in households
2. Compare the behaviour of the 'global market' and the 'use of petroleum' for industry, transportation, and households. Are they the same or different?
   The four quantities show the same cyclic behaviour

9. Effects on the atmosphere

1. Via which two quantities does 'petroleum available' affect 'greenhouse gases'? And 'pollution'?
   Both 'greenhouse gases' and 'pollution' are affected via 'use of petroleum' for industry and for transportation
2. What are the values of the quantities 'use of petroleum' for the industry and 'greenhouse gases' in states [16, 21, 12]?
   High, medium, low, respectively, for the two quantities
3. Which quantities influence the overall atmosphere 'temperature'?
   'Greenhouse gases' and 'pollution'
4. In which states is the overall atmosphere 'temperature' in the zone of global warming?
   11, 14, 16

10. Effects on households

1. Which quantity is negatively influenced by 'petroleum available'?
   'Use of solid fuel'
2. What is the relation between the values of quantities 'use of petroleum' in households and 'use of solid fuel'?
   They have exactly opposite values (inverse correspondence)
3. What happens with 'smoke' in states [16, 21, 12]? And how does it compare to the behaviour of 'global market' in the global economy?
   'Smoke' increases from low to high, while the 'global market' shows a shortage of petroleum in the global economy

11. Effects on human health

1. What happens with 'chronic respiratory diseases' when 'smoke' decreases?
   'Chronic respiratory diseases' also decreases
2. How do the values of 'chronic respiratory diseases' relate to 'use of solid fuel'?
   These quantities have the same values in every state
3. If 'generic respiratory diseases' increases, what is the reason for that behaviour?
   'Generic respiratory diseases' increases because atmospheric 'pollution' increases
4. How do the behaviour of 'chronic respiratory diseases' and 'generic respiratory diseases' compare in the selected behaviour path?
   These quantities present opposite behaviours: when 'chronic respiratory diseases' increases, 'generic respiratory diseases' decreases

12. Global effects of the petroleum use

1. According to this model: is smoke produced in households related to global warming?
   No, there is no relation between 'smoke' and the overall atmosphere 'temperature' in this model
2. What happens to the overall atmosphere temperature when petroleum available increases, remains stable and decreasing?
   'Temperature' changes in the same direction, that is, it also increases, remains stable and decreases
3. What happens to generic respiratory diseases and to chronic respiratory diseases when petroleum available increases, remains stable and decreases?
   'Generic respiratory diseases' changes in the same direction as 'petroleum available'; 'chronic respiratory diseases' changes in opposite direction when compared to 'petroleum available'

Background and attitude questionnaires

PARTICIPANT DATA

Please fill in or circle the relevant data

1 – **Participant name:**

2 – **Participant number:**

3 – **Male / Female**

4 – **Age:**

5 – **Topic of Study:**

6 –**Degree of Study:**
BSc / MSc / …

7 – **Year of Study:**
1st / 2nd / 3rd / 4th / …th

8 – Amount of computer experience:

1	2	3	4	5	6	7
none						much

9 – Amount of expertise about Ecology:

1	2	3	4	5	6	7
none						much

10 – Amount of expertise about Conceptual Modelling:

1	2	3	4	5	6	7
none						much

11 – Amount of expertise about Qualitative Reasoning:

1	2	3	4	5	6	7
none						much

PARTICIPANT FEEDBACK

1 – How much do you feel you have learned during the session?

1	2	3	4	5	6	7
nothing						much

2 – Did you find it difficult or easy to answer the test questions?

1	2	3	4	5	6	7
very easy						very difficult

3 – Did you find it difficult or easy to use the software interface?

1	2	3	4	5	6	7
very easy						very difficult

4 – Did you find the diagrams difficult or easy to understand?

1	2	3	4	5	6	7
very easy						very difficult

5 – Did you enjoy today's session?

1	2	3	4	5	6	7
not at all						very much

6 – Did you enjoy using the software?

1	2	3	4	5	6	7
not at all						very much

7 – If you want, you can leave remarks about today's session:

Appendix 3

River Kamp study questionnaires

Questionnaire used for the *general* evaluation study of the river Kamp models A and B.

1. **QR models** present **complex knowledge** in an **understandable manner**.

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

Why not?

2. **The QR approach** allows for a **clear representation** of real-world phenomena like **a sustainable development of the riverine landscape Kamp.**

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

Why not?

3. **QR and GARP3** can be seen as a **valuable learning tool** for **real-world causal relationships** related to **a sustainable development of riverine landscapes.**

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

Why not?

4. **The presented QR model** might significantly **contribute to the understanding of students and stakeholders which entities and processes drive a sustainable** development of a riverine landscape and therefore **enhances their capability of making decisions**.

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

Why not?

5. The **causal map** of the model reflects **important information** related to **a sustainable development of the Kamp valley.**

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

What is missing?

6. **Which part of the model** was **most interesting** for you?
Please write down in short words.

7. **Which part of the model** most should be **enhanced**?
Please write down in short words.

8. The **model** can be used **for the targeted purpose of teaching students and other interested stakeholders** on **sustainability issues on a catchment level**.

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

If necessary, specify your answer.

9. For **which purpose** do you think the presented **QR approach is most suited?**
   a. **Stakeholder integration**

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

   b. **University lectures**

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

   c. **Decision-making**

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

   d. **Others** (**to be added** e.g. technical staff from the government, researchers, secondary school students)…
   *Please add:*_______________

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

10. Additional comments:

Questionnaire used for the *expert* evaluation study of the river Kamp model A.

11. The **entities and configurations** are **relevant and sufficient** to support a representation of the system structure.

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

If necessary, specify your answer

12. The **quantities used** capture the **most interesting properties** of the entities.

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

If necessary, specify your answer

13. The **quantity spaces** and **values** capture the **most interesting qualitative states of the entities**.

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

If necessary, specify your answer

14. The (important) **model fragments** are conceptually correct and clear.

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

If necessary, specify your answer

15. The **presented scenarios** describe a **real situation** that it is good enough to trigger an **interesting/good simulation**.

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

What is missing?

16. The general behaviour (how it develops trough the simulation) of the **presented model** is **in accordance to what is already known** (or accepted).

○	○	○	○	○
1	2	3	4	5
I fully disagree	I largely disagree	I somewhat disagree/agree	I largely agree	I fully agree

If necessary, specify your answer

Tables enumerating river Kamp evaluating study results

Table 6.1 – part 1

Questions	I fully disagree (1)	I largely disagree (2)	I somewhat agree/disagree (3)	I largely agree (4)	I fully agree (5)
1) **QR models** present **complex knowledge** in an **understandable manner** .					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)		1		1	2
Expert (water resources mangagement/n=1 , water resources research/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)				5	
Not specified				1	1
TOTAL	0	1	0	7	3
2) **The QR approach** allows for a **clear representation** of real world phenomena like **a sustainable development of the riverine landscape "Kamp".**					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)		1	2		1
Expert (water resources mangagement/n=1 , water resources research/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)			3	1	
Not specified			1		1
TOTAL	0	1	6	1	2
3) **QR and GARP3** can be seen as a **valuable learning tool** for **real world causal relationships** related to a **sustainable development of riverine landscapes.**					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)		1			3
Expert (water resources mangagement/n=1 , water resources research/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)			1	3	1
Not specified				1	1
TOTAL	0	1	1	4	5
4) **The presented QR model** might significantly **contribute to the understanding of students and stakeholders which entities and processes drive a sustainable** development of a riverine landscape and therefore **enhances their capability of making decisions** .					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)			1		3
Expert (water resources mangagement/n=1 , water resources research/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)				5	
Not specified				1	1
TOTAL	0	0	1	6	4

Table 6.1 – part 2

5) The **causal map** of the model reflects **important information** related to a **sustainable development** of the **Kamp valley.**					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)			2	1	1
Expert (water resources mangagement/n=1 , water resources research,/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)			1	3	1
Not specified				1	1
TOTAL	0	0	3	5	3
6) **Which part of the model** was **most interesting** for you?					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)	Most interesting was to see: that private interests trigger sustainability process, the combined influence of planners, science and local population (stakeholders) on the quality of sustainability plans and the whole sustainability processs; the interrelatedness of a system; the possibility for different potential intervention options; that ecological integrity AND human well being are represented in the sustainability model.				
Expert (water resources mangagement/n=1 , water resources research,/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)	Generating stock/flow elements in a qualitative way and relating quantities via P`s and I`s was very interesting.				
Not specified	Specific scenarios, the catastrophic event as trigger for gv action for sd, the interrelatedness and influences of the different model components were most interesting.				
7) **Which part of the model** most should be **enhanced** ?					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)	Private interests should be better represented, as a basis to minimize them and achieve sustainable development				
Expert (water resources mangagement/n=1 , water resources research,/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)	The full causal model looks a bit complex & heterogenous (sorting?); pointing out better the MOST important variables for a sustianable development would be good.				
Not specified					
8) The **model** can be used **for the targeted purpose** of **teaching students** and **other interested stakeholders** on **sustainability issues on a catchment level** .					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)	1			1	2
Expert (water resources mangagement/n=1 , water resources research,/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)			3	1	1
Not specified				1	
TOTAL	1	0	3	3	3
9) For **which purpose** do you think the presented **QR approach** is **most suited?**					
a. **Stakeholder integration**					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)		1		1	2
Expert (water resources mangagement/n=1 , water resources research,/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)			1	3	1
Not specified				1	1
TOTAL	0	1	1	5	4
b. **University lectures**					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)	1			1	2
Expert (water resources mangagement/n=1 , water resources research,/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)			1	3	1
Not specified		1			1
TOTAL	1	1	1	4	4
c. **Decision making**					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)		1		3	
Expert (water resources mangagement/n=1 , water resources research,/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)			2	1	2
Not specified				2	1
TOTAL	0	1	2	6	3
d. **Others** (**to be added** eg. technical staff from the goverment, researchers, secondary school students)…					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)					1
Expert (water resources mangagement/n=1 , water resources research,/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)				2	
Not specified					2
TOTAL	0	0	0	2	3
10) Additional comments:					
Student 1 , Student 2 (ecology/limnology), student 3 (landscape planning/social ecology), student 4 (Phd/aquatic ecosystem research)	Keeping the overview over the model and the P's and I's is a bit complicated				
Expert (water resources mangagement/n=1 , water resources research,/n=2, hydraulic modeling & fish ecology/n=1, not specified/n=1)					
Not specified	The presented models allow for a communication of new viepoints of existing problems and facts; application within different other fields & domains, mainly related to education, is thinkable; interesting is to take different mental models (also culture-specific) as a starting point for models & discussion.				
TOTAL (n=11 persons)	2	6	18	43	34
TOTAL STUDENT (n=4 persons)	2	5	5	8	17
TOTAL EXPERT (n=5 persons)	0	0	12	27	7
TOTAL NOT SPECIFIED (probably students, n=2 persons)	0	1	1	8	10

Table 6.2

Questions	I fully disagree (1)	I largely disagree (2)	I somewhat agree/disagree (3)	I largely agree (4)	I fully agree (5)
1) **QR models** present **complex knowledge** in an **understandable manner**.					
				1	
TOTAL	0	0		1	0
2) **The QR approach** allows for a **clear representation** of real world phenomena like **a sustainable development of the riverine landscape "Kamp".**					
				1	
TOTAL	0	0	0	1	0
3) **QR and GARP3** can be seen as a **valuable learning tool** for **real world causal relationships** related to **a sustainable development of riverine landscapes.**					
				1	
TOTAL	0	0	0	1	0
4) **The presented QR model** might significantly **contribute to the understanding of students and stakeholders which entities and processes drive a sustainable** development of a riverine landscape and therefore **enhances their capability of making decision**					
				1	
TOTAL	0	0	0	1	0
5) The **causal map** of the model reflects **important information** related to **a sustainable development of the Kamp valley**					
				1	
TOTAL	0	0	0	1	0
6) The **entities and configurations** are **relevant and sufficient** to support a representation of the system structure.					
				1	
TOTAL	0	0	0	1	0
7) The **quantities used** capture the **most interesting properties** of the entities.					
				1	
TOTAL	0	0	0	1	0
8) The **quantity spaces** and **values** capture the **most interesting qualitative states of the entities**.					
				1	
TOTAL	0	0	0	1	0
9) The (important) **model fragments** are conceptually correct and clear.					
				1	
TOTAL	0	0	0	1	0
10) The **presented scenarios** describe a **real situation** that it is good enough to trigger an **interesting /good simulation**.					
				1	
TOTAL	0	0	0	1	0
11) The general behaviour (how it develops trough the simulation) of the **presented model** is **in accordance to what is already known** (or accepted).					
			1		
TOTAL	0	0	1	0	0
12) **Which part of the model** was **most interesting** for you?					
	Causal relationships in general , to model the money for community driven devleopment as an own quantity - that can only be treated as "spent by the community", when the community nvolvement (stakeholder participation) is really implemented; otherwise it is not a community driven investment!				
13) **Which part of the model** most should be **enhanced** ?					
	Govemment action, as this is in reality of high complexity, as it is also driven by the general political structures , difficulties between different organization units with regard to their competences (personal behaviour!) and financial resources, and policies with complementary aims as policies and their integration often lack behind the social development. That means a more detailed study of how internal political sructures determine the implementation process is needed.				
14) The **model** can be used **for the targeted purpose of teaching students and other interested stakeholders** on **sustainability issues on a catchment level**.					
					1
TOTAL	0	0	0	0	1
15) For **which purpose** do you think the presented **QR approach is most suited?**					
a. Stakeholder integration					
			1		
TOTAL	0	0	1	0	0
b. University lectures					
				1	
TOTAL	0	0	0	1	0
c. Decision making					
		1			
TOTAL	0	1	0	0	0
d. **Others** (**to be added** eg. technical staff from the goverment, researchers, secondary school students)...					
				1	
TOTAL	0	0	0	1	0
16) Additional comments:	A good approach for socialearning and confict managment (mediaton) when appled togehter wth students or stakeholders n a mediated or group modelng process. Identfyng dependences and causalrelatonships is of high interest and importance for understaning a system. It could also be of relevance, to think about which degree of fulfilment of the three plars (the weighting of the indicidualfactors) of sd realy is "sustainable" and who defines how the relatonships between the plars should look ke (oftem the focus is more on the ecologicalside, sometimes more on the economic side often closely inked to human welbeing in industrial societies; human wishes or controlng paradigms often are not sustainable; probably the currently controlng parading of intgerating the needs and wishes of the human populaton is does not lead to a realsustainable development form an ecologicalviewpoint). It is important to very weldefine the terms and their use within the model				
TOTAL domain expert evaluation	0	1	2	12	1

Table 6.3 – part 1

Questions	I fully disagree (1)	I largely disagree (2)	I somewhat agree/disagree (3)	I largely agree (4)	I fully agree (5)
1) **QR models** present **complex knowledge** in an **understandable manner**.					
			1		
TOTAL	0	0	1	0	0
2) **The QR approach** allows for a **clear representation** of real world phenomena like **water abstraction and ist effects on fish.**					
TOTAL	0	0	1	0	0
3) **QR and GARP3** can be seen as a **valuable learning tool** for **real world causal relationships** related to **water abstraction and its effects on fish.**					
			1		
TOTAL	0	0	1	0	0
4) **The presented QR model** might significantly **contribute to the understanding of students and stakeholders of how different modes of water abstraction might affect fish** and therefore **enhances their capability of making decision.**					
					1
TOTAL	0	0	0	0	1
5) The **causal map** of the model reflects **important information** related to **different modes of water abstraction and ist effects on fish.**					
					1
TOTAL	0	0	0	0	1
6) The **entities and configurations** are **relevant and sufficient** to support a representation of the system structure.					
				1	
TOTAL	0	0	0	1	0
7) The **quantities used** capture the **most interesting properties** of the entities.					
					1
TOTAL	0	0	0	0	1
8) The **quantity spaces** and **values** capture the **most interesting qualitative states of the entities**.					
					1
TOTAL	0	0	0	0	1
9) The (important) **model fragments** are conceptually correct and clear.					
					1
TOTAL	0	0	0	0	1
10) The **presented scenarios** describe a **real situation** that it is good enough to trigger an **interesting /good simulation**.					
				1	
TOTAL	0	0	0	1	0
11) The general behaviour (how it develops trough the simulation) of the **presented model** is **in accordance to what is already known** (or accepted).					
					1
TOTAL	0	0	0	0	1
12) **Which part of the model** was **most interesting** for you?					
	That it is easy to change the content of a scenario by using different assumptions to model the effects of the same human pressure on different guilds (positive and negative effetcs of flow velocity and water temperature on different guilds).				
13) **Which part of the model** most should be **enhanced**?					
	A more realistic representation of the natural variability of the river discharge (probably by using the random function in the scenario editor) and the amount of abstracted water related to mean annual flow as this defines the frequency of water overflow events at weirs that are suspected to have a significant effect on fish. A more realistic representation of the influence of the length of the water abstraction stretch on the temperature devleopment within the river (at the moment the river stretch is treated as a "container" with the same abiotic factors everywhere) and an integration of the effect of morphology on fish and on water temperature!				
14) The **model** can be used **for the targeted purpose of teaching students and other interested stakeholders** on **water abstraction and ist effects on fish**.					
					1
TOTAL	0	0	0	0	1

Table 6.3 – part 2

15) For **which purpose** do you think the presented **QR approach is most suited?**					
a. Stakeholder integration					
			1		
TOTAL	**0**	**0**	**1**	**0**	**0**
b. University lectures					
					1
TOTAL	**0**	**0**	**0**	**0**	**1**
c. Decision making					
		1			
TOTAL	**0**	**1**	**0**	**0**	**0**
d. **Others** (**to be added** eg. technical staff from the goverment, researchers, secondary school students)…					
					1
TOTAL	**0**	**0**	**0**	**0**	**1**
16) Additional comments:					
	"The software now can be used very intuitively, which is a prerequisite for the target, to motivate stakeholders and students to put their conceptual knowledge in causal models"! "It takes time and engagement, to establish approaches like that in society and (university) education teaching such approaches are the basis for their broader use and application by the upcoming generation(s)". "To further enhance the modelling process itself it would be helpful to always see the consequences of my model definitions and implemented model fragments (configurations, proportionalities and influences) on the fly in an accompanying window of the software (for example as they can be explored by the "show entities & configurations" button, by the "show dependencies" button)". "It also could be helpful to have the full model shown in a screen like in the "show entities & configurations" window with the opportunity to select parts of the model to be run in a simulation (running only parts of the model by simply drawing a window over a certain part of the model)". "To link the outcomes of causal models to a GIS would open a new field of app.".				
TOTAL domain expert evaluation	**0**	**1**	**4**	**2**	**9**