# Visualization of Complex Systems

**Magnhild Viste**

Dissertation for the degree philosophiae doctor (PhD)
at the University of Bergen

September 24, 2007

## Acknowledgements

I would like to thank my advisor Associate Professor Katherine Goodnow for invaluable support throughout the work in this thesis. Without her encouragement, advice, and belief in me the end to this thesis may not have been so joyous. In addition to being an involved and enthusiastic advisor she has been an excellent role model of someone who manages to balance between academic and family life.

My second advisor, Professor Pål I. Davidsen, initiated the VOCS project and due to his engagement the project came into being. He has also provided advice particularly regarding complex systems and system dynamics theory.

My dear friend and colleague, Hanne-Lovise Skartveit, has been a great inspiration and support throughout the work on the thesis. Although our friendship endures, I will miss having her as an office mate.

Several students have been part of the project group and have contributed to the work described in this thesis. Krisitan Andre Kastet worked on the first part of the enhanced two-shower model. I would particularly like to thank him for his commitment and hard work before and during the CSCL2003 conference. I would also like to thank Laila Frotjold for her development work and for being available even after finishing her Master's thesis and starting a new job. Nils Magnus Djupvik has also been of great help and together with Laila Frotjold he developed the web service on which the prototypes are running.

My fellow PhD students (of which some are no longer PhD students) have participated in numerous encouraging conversations and lunches: Eskil Andreassen, Frode Guribye, Ingerid Rødseth, Lars-Jacob Hove, Sonja Wilhelmsen, Tone Lønning, and Vibeke Vold.

I would further like to thank my family and friends for always believing in me. Particular thanks to my friend Manne. I miss the days of our morning coffees at

Dromedar. We were usually busted by some advisor who thought that mornings should be spent in the office rather than a coffee shop, but I am still convinced that our morning coffee breaks helped us manage our work and lives better.

At last, I would like to thank the two persons who are dearest to me, Olve and Ask. Olve for his endurance with me and invaluable support during the ups and downs of the process of writing a doctoral thesis. I would also like to thank Ask for enduring this period and I promise him that from now on his mom will (mostly) take the weekends off and that I will not miss any more vacations.

## Abstract

Complex systems are difficult to understand. To aid understanding of complex dynamic systems the field of system dynamics has developed a set of visualization methods for graphic representation of the simulation models of complex systems. The resulting visualizations, however, may sometimes be difficult to understand for audiences without a background in the scientific investigation of complex dynamic systems. It is therefore necessary to find new ways of representing complex systems with the purpose of making them more accessible. This thesis describes how system dynamic models may be visualized with a particular audience in mind – novice system dynamics students. The use of narrative elements and multimedia in the development of the user interface is explored. A variety of theories and research fields, such as narrative theory, system dynamics, human computer interaction, and computer support for collaboration and learning (CSCL) are utilized in the study. The project includes analysis and development of visualizations of system structure, behavior, and narratives that link and explain the relationship between structure and behavior of complex dynamic systems.

Two prototypes of interactive learning environments are developed and evaluated in the project: *The Two-Shower prototype* and *The Quito prototype*. The Two-Shower prototype is based on a small model of a complex system and portrays two showers that share a hot water tank. The model is a metaphorical model and represents a simple system for resource sharing. The second model is a model of Quito, the capital of Ecuador, and describes the dynamics of the citizens, tourists, pollution, and the maintenance of Quito's UNESCO enlisted historic buildings. The Two-Shower prototype is evaluated by usability tests and suggestions for new design are made. The experiences from the development of the Two-Shower prototype are then discussed considering how they may be transferred to the Quito prototype, which is based on a larger model. Being a model of a social system, it portrays issues that are more complex. Modeling of social systems raises some particular issues regarding

choice and representation of variables. The thesis attempts to make bridges between complex system theory, system dynamics, visualization, and learning theory – as related to design.

# Contents

# List of Figures

# 1. Introduction

There is a need to understand complex dynamic systems within a variety of disciplines. What are, for example, the causes of population and economic growth and decline within a city? Why do industrial nations become richer while developing countries become relatively poorer? Which forces drive the fluctuations of the labor market? What are the processes that influence inflation? How do environmental factors affect human migration? Why does the information gap between members of a society increase? What are the economic forces that drive the fluctuations of commodity prices? These are examples of questions that we may try to answer through a complex dynamic systems investigation.

Research indicates that even small complex dynamic systems may be difficult to understand and that people have problems understanding basic elements of such systems (Moxnes, 2004; Sterman & Sweeney, 2002; Sweeney & Sterman, 2000). Both size and complexity are in this case of course relative terms. (I will discuss this in detail later in this introduction (Section 1.3).)

To aid understanding of complex dynamic systems the field of system dynamics has a set of visualization methods for graphic representation of the underlying simulation model. The resulting visualizations may sometimes be difficult to understand for audiences without a background in the scientific investigation of complex dynamic systems. This constitutes a challenge for the designers of models of complex dynamic systems. In several of the previous studies of complex dynamic systems there is little focus on how the complex dynamic system is represented (see for example Diehl & Sterman, 1995; Kleinmuntz, 1985; Moxnes, 1998a, 1998b, 2000, 2004; Paich & Sterman, 1993; Sterman & Sweeney, 2002; Sweeney & Sterman, 2000).

There is a need for tools that represent complex dynamic systems in new ways that are created with the purpose of making complex dynamic systems more accessible. This thesis constitutes a small step in that direction.

In this thesis I explore how narrative and multimedia may be used in the development of prototypes of learning tools that represents a complex dynamic system. A variety of theories and research fields, such as narrative theory, system dynamics, human computer interaction, and computer support for collaboration and learning (CSCL) are utilized to that end. My project includes analysis and development of visualizations of system structure, behavior, and narratives that link and explain the relationship between structure and behavior of complex dynamic systems.

In this thesis I describe the development of two interfaces – prototypes for learning environments based on system dynamic models - and two initial tests of the functionality of one of these prototypes with a focus on the effectiveness of the visualizations.

The two prototypes developed and evaluated in the research project of which this thesis is part are *The Two-Shower prototype* and *The Quito prototype.* The Two-Shower prototype is based on an existing system dynamic model described by Morecroft, Larsen, Lomi, & Ginsberg (1995). The Quito prototype is inspired by the works of Jay Forrester (1969). Contrary to the Two-Shower prototype, however, the system dynamic model on which the Quito prototype is based was developed by the project group.

The Two-Shower prototype has been the main focus of my work. This is a relatively small model with few, however closely connected, variables with a high degree of interaction. In the case of this relatively small model it was important and possible to seek to obtain a one-to-one relationship between the simulation model and the interface. For larger models a one-to-one relationship would not be possible or desirable and the organization of the user interface views and the navigation system between these views would require additional efforts.

Our second prototype, which is less developed, is based on a larger model. Through this prototype I seek to focus on some of the problems that occur when visualizing larger models where the structure of the model, its behavior, and the links between them must be organized into multiple interface views.

Three beta-tests were carried out by members of the VOCS project. One on the Quito prototype and two on the Two-Shower prototype. All three were aimed at evaluating the functionality of the user interface. The testing of the Quito prototype was carried out by Master student Nils Magnus Djupvik (2006). With my focus on the Two-Shower prototype I discuss this in less detail in this thesis. The first test of the Two-Shower prototype was of a single-user version of the prototype carried out by Master student Laila Frotjold (2005). I carried out the second test of a further developed two-user version of the Two-Shower prototype. The second test was aimed at evaluating whether the participants felt that the visualizations aided them in understanding and controlling the system. This does not necessarily indicate whether the participants understood the system or that they learned something by using it. It is rather an evaluation of the interface based on their own experience and of how they themselves felt the interface and its various components aided them in controlling and understanding the underlying model. This testing is part of a first step towards creating a fully-fledged learning environment. A second step would be a study of whether the system actually supports understanding and learning, but that is not considered in the thesis.

## 1.1 The VOCS (Visualization of Complex Systems) Research Project

This thesis is part of the research of the Visualization of Complex Systems (VOCS) research group. VOCS was a collaboration project between the Department of Information Science and Media Studies and InterMedia at the University of Bergen and Powersim AS. InterMedia is a multidisciplinary research center for researchers interested in information and communication technology, ICT-based information and knowledge dissemination, and learning. Powersim AS is a consultant company that develops simulation-modeling software based on the principles of system dynamics, such as one finds in economical, social and ecological systems. The aim of the VOCS project was to develop prototypes for interactive learning environments where the link between structure and behavior was visualized using graphics, animation, text, and video.

The VOCS project and this thesis was funded by the Norwegian Research Council. The project was led by Professor Pål I. Davidsen, Head of the System Dynamics Group, currently at the Department of Geography, University of Bergen. Professor Katherine Goodnow at the Department of Information Science and Media Studies was also a part of the VOCS project and has been, together with Davidsen, the advisor for this thesis.

In addition to my research, the project involved research by PhD student Hanne-Lovise Skartveit. Skartveit focused on the integration of video in interactive learning environments, interactive documentaries, and games.[1]

A number of Masters students have also contributed to the project in different ways. They have all worked in close collaboration with Skartveit and myself. Details of the collaboration are described in Section 1.5.

Porfirio Guevara-Chaves and Antonio Perez-Bennet (2002) developed the first version of the underlying simulation model for the Quito prototype. Their work was a term project as part of the Masters program in system dynamics at the Department of Information Science.

Kristian Kastet participated in the development of a first two-user version of the Two-Shower prototype.

Nils Magnus Djupvik and Laila Frotjold (2003) continued the work on the simulation model of Quito. Their work was also a term project and part of the Master's program at the Department of Information Science. Together they also developed the web-

[1] Skartveit delivered her thesis in May 2007.(Skartveit, 2007, forthcoming).

service on which both the Two-Shower prototype and the Quito prototype are running.

Laila Frotjold (2005) developed the framework for a text-based story generator for the Two-Shower prototype as part of her Master's thesis.

Nils Magnus Djupvik has worked on the development and technical implementation of the user interface for the current version of the Quito prototype and has written his Master's thesis in relation to this (2006).

Olve Sæther Hansen was hired as an external consultant to improve some features of the story generator and to further develop the Two-Shower prototype.

## 1.2 The Intended Audience of this Thesis

As stated earlier this thesis is a prototype of a small, complex dynamic system to introduce students to the domain of complex dynamic systems. The intended audience of the prototype, if further developed, is first year students in system dynamics. The intended audience of the thesis would therefore be teachers in system dynamics. The thesis should also be of interest to others who are developing interactive learning environments for complex dynamic systems.

## 1.3 Why the Two-Shower Model was Chosen

The model was chosen for a variety of reasons related to both content and complexity. First, the model represents both *human* as well as *physical* elements that could be illustrated by multimedia. By *human* elements, I refer to the representation of the *decision processes* of a *simulated person* in the model who *reacts* to the temperature of the water in the shower by changing the tap setting. A representation of the results of physical impact on a person, such as hot and cold water, and the resulting pain in contrast to comfort, may be easier to visualize than other emotions such as, for example, jealousy, sorrow, or excitement. The model is suitable for *a narrative approach* as it contains two actors that each strive to reach a goal, and in doing so, disturbs the context of the other actor. This forms a good point of departure for creating stories about their interaction and experience with narrativity for introducing complex dynamic systems.

The representation of the *physical* construction of the pipe system is relatively uncomplicated. The straightforwardness of the physical system and the human elements makes the model suitable for visualization of the processes that take place in the interplay between humans and the physical system.

A further, important reason for choosing the Two-Shower model relates to its size and complexity. A small model makes *it easier to develop an interface with a relatively strong link between the underlying simulation model and the user interface.* Relatively few variables must be visualized and it is easier to develop a simple user interface that includes richer information about the underlying model.

The Two-Shower model is suitable also as a starting point because, although the model consists of a relatively small structure, it is *sufficiently comprehensive to encompass the main characteristics that make complex, dynamic systems difficult to understand*. A small system may be complex and difficult to understand even though it has few variables. The variables may be coupled tightly and there may be a high degree of interaction between the variables. Great variations in the type of relations between the variables also make it more complex because this means that the argument that follows from structure to behavior must change its character for every relation. As mentioned at the start of this introduction, small models have been shown to be quite complex in studies carried out by Sterman and Sweeney (Sterman & Sweeney, 2002; Sweeney & Sterman, 2000) and Moxnes (2000; 2004). Moxnes has also shown that models, though slightly larger than the Two-Shower model, have a complexity sufficiently great to be of significance when it comes to understanding and controlling the models (Moxnes, 1998a, 1998b). Size and complexity are therefore not the same.

Let me explain this issue of size and complexity in more detail through a first short description of the Two-Shower model.

## 1.4 The Two-Shower Prototype

### 1.4.1 Introduction to the Original Two-Shower Model

The original Two-Shower model by Morecroft et al. (1995) describes two showers that share one hot water resource. The underlying structure (the pipe system) renders control of the temperature difficult. Although seemingly simple, the model contains some of the basic characteristics that make a complex, dynamic system difficult to understand and control.

There is a *delay* from a change is made in the tap setting until the water with the new temperature reaches the showerhead. The changes in the tap setting may be exaggerated if the delay in the system is disregarded. One of the goals in the development of the Two-Shower prototype was therefore to visualize this delay.

The model also contains *feedback*. If there were an undesired temperature of the water at the showerhead the person in the shower would change the tap setting. After some time this would change the temperature of the water at the showerhead (see figure 1-1).

In addition, the model is *nonlinear*. The implication of this is that a decision by Shower 1 to change the tap setting will affect other variables in the system, in this case the conditions of Shower 2. Shower 2 will then change its tap setting which again influences the conditions of Shower 1.

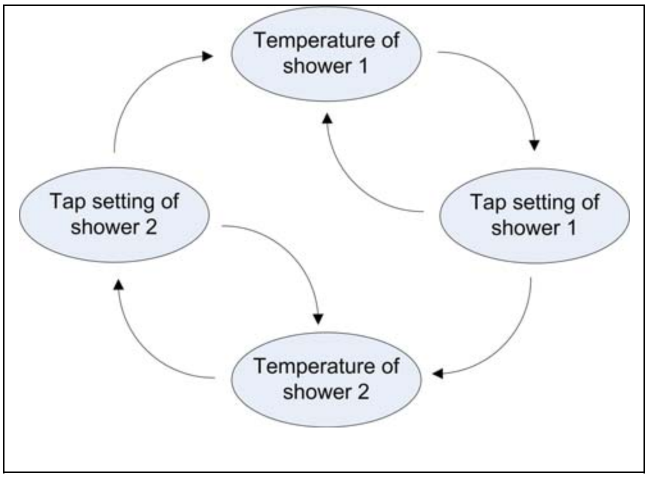


*Figure 1-1: Overview of the structure of the Two-Shower model.*

More specifically, the effect of a change in tap setting for one shower depends on the current tap settings of both showers. This is because each shower's access to the hot water resource depends on the tap settings of both showers. Both, internal and external factors set the terms of the systems (the tap setting and the tap setting of the other shower). Subsystems thus influence the conditions for how the system may be controlled internally.

Controlling the system based on previous experience may be difficult because the same tap setting (or change in such a setting) will not always result in the same temperature. Nonlinearity is part of what Milrad, Spector and Davidsen (2003) saw as general problems that people face in their attempt at understanding complex dynamic systems. Their concern was particularly with problems people have in considering the system as a whole when trying to solve a problem that seems to be local. The visualizations of the Two-Shower prototype had to be developed with the intention of describing such nonlinearity.

Finally, there are two *stakeholders* in the original Two-Shower model. An enduring comfortable temperature for both showers is only feasible when they both take into account the reactions of the other, the system structure, and system state when changing the tap setting. The decisions regarding changes in tap settings of both showers had to be coordinated. One of the main problems is that the characters of the two showers are not able to communicate directly. They *experience* the changes in the conditions through feedback from the system. This could lead them to misperceive the structural characteristics of the system, thinking for example that blue represented hot water because the more they turned the tap setting towards red the colder the water at the showerhead became, although the reason it turned colder would be the actions of the character in the other shower. The system is in this way based on experience. Because the system is nonlinear it is difficult to base decisions on previous experience. A solution could then be to turn to structural explanations of the behavior and to challenge the people in the showers (or those who control the showers) to question their hypothesis of what the structure of the system looks like.

The user interface of the prototype had to be developed with the purpose of supporting the coordination of the decisions between the participants and to support them in communicating about how their showers were linked through the structural characteristics of the system.

The original Two-Shower model and its complexity will be described in greater detail in Chapter 2.

### 1.4.2 The Purpose of the Two-Shower Prototype

The purpose of our development and use of the Two-Shower prototype is to find new ways of presenting complex dynamic systems. Prototype-based research is utilized and new visual and narrative representations have been developed. These visual representations are intended to create a basis upon which participants can communicate about complex dynamic systems and to support them in discussing and expressing their understanding of the structure of the system, although the extent to which the user interface supports this communication is not studied in the interface test described in this thesis.

A further goal is that the experiences from this project may be transferred to future research to other models of complex dynamic systems. I do this in a limited way with regards to the Quito prototype described in Chapter 10.

The Two-Shower prototype, fully developed, could be used as part of an introduction course to system dynamics. As mentioned in Section 1.2, the target audience for the Two-Shower prototype is novice system dynamics students. The students would likely be undergraduate or graduate students at a university. It is important to note also that the prototype, fully developed, is not intended as a stand-alone product but that additional information must be provided to the participants.

## 1.5 Method

As mentioned previously, prototype-based research is utilized in this project. Prototyping traditionally involves a four-step procedure where, first, the basic requirements are identified, a simple version of the product is developed and then implemented (Naumann & Jenkins, 1982). The product is revised based on feedback from users and an improved version is implemented. The last two steps are repeated until the users find the product satisfactory. For this project this has meant that several versions of the Two-Shower prototype have been made, tested, and changed based on the test results. The prototype was tested by the research team several times and two formative evaluations or beta tests of the model were conducted. To offer a better understanding of the process and method used, I include here a chronological overview of the work on the prototype and this thesis. Details of the methods used for each phase of this work are described in further detail in later chapters.

### 1.5.1 Spring 2002: Initial Decisions and Visualization

The work on the VOCS project started in the spring of 2002 when several meetings were held between the project members, a developer, and the general manager of Powersim. One of the first tasks was to select the problems of the models that would be used in the project and to carry out initial reviews of literature.

It was decided to develop two prototypes where one would deal with problems of resource sharing and be based on the Two-Shower model (Morecroft et al., 1995).

The other prototype, the Quito prototype, would deal with problems regarding the dynamics of a city, in this case with particular emphasis on preservation of Quito's historic buildings. Three initial publications were written (Skartveit, Goodnow, & Viste, 2002, 2003; Skartveit & Viste, 2003), in which we described the plans for the two models.

I began the initial work on the Two-Shower prototype by implementing the equations that were listed in Morecroft et. al (1995) into Powersim Constructor and studying the structure and behavior of the model. The next step involved removing the variables that represented the *human decision processes* as the decisions to change the tap setting would be made by a participant in the prototype. Some additional variables were also added in order to ensure coherence between the variables that were sent between the simulation model and the user interface. Skartveit joined me in developing the plans for a user interface. As will be discussed in Chapter 3, it was decided that the user interface should include narrative characteristics and in part be based on narrative theory (see for example Bordwell & Thompson, 1997).

Note again that narrative theory, interactive documentary, and games, in the division of labor between Skartveit and myself, is primarily dealt with by Skartveit. My theoretical perspective is from a complex dynamic systems and information science perspective.

It was decided that multimedia should be used to represent the characteristics of the underlying simulation model. This was in part based on theories of Mayer (Mayer, 2001, 2005; Mayer & Anderson, 1991), Alessi & Trollip (2001), Alessi (2002), Dix, Finlay, Abowd, & Beal (1998) and Tufte (1997) and will be further discussed in Chapter 4.

### 1.5.2 Fall 2002 to Fall 2003: A Prototype of a Collaborative Interactive Learning Environment

I implemented the first version of the user interface of the Two-Shower prototype in Flash MX. The system dynamic model had to be adapted, some variables were removed, and some help-variables were added in order for the values to be communicated between Powersim SDK and Flash MX. Master student Kristian Kastet joined the project and together we developed a first two-user prototype. I did the work in Powersim and Flash. Part of this work was the design and programming of animations and the response to user input and the consequent reflection of changes in values of the model. Hanne-Lovise Skartveit took the photos for the Flash application and designed some of the graphics for the user interface. Kristian Kastet developed a Visual Basic program that transferred data between Powersim and Flash. This program controlled the progression of the simulation model based on user input to the Flash user interface. This version of the Two-Shower prototype was presented and demonstrated at CSCL 2003 (Viste & Skartveit, 2003). A further version was described in Viste & Skartveit (2004).

At this stage, particular emphasis was made on theories of learning and collaboration and on model transparency and how this could be used in design (see for example Alessi, 2002; Koschmann, 1996a, 1996b; Koschmann, Kelson, Feltovich, & Barrows, 1996; Lipponen, 2002; Stahl, 2002). Theory, method, and development of the Two-Shower prototype as a prototype of a collaborative interactive learning environment will be discussed in Chapter 5.

### 1.5.3 Spring 2004 to Spring 2005: The Story Generator and Test 1 of the Two-Shower Prototype

Laila Frotjold joined the project in the spring of 2004 to work on the Two-Shower prototype as part of her Master's thesis. Her Master's thesis (2005) was written under the supervision of Associate Professor Weiqin Chen at the Department of Information Science and Media Studies, University of Bergen.

The development of a story generator had been discussed from the outset of the project and Frotjold was given the task to implement such a generator into the model. She developed a text-based story generator, and at the same time, I made some additional changes to the representation of structure in the user interface developed in Flash.

At this time it was decided to change the technological framework for the project as Visual Basic seemed not to be ideal to handle the flow of data between Powersim SDK and Flash MX. Part of the existing code, developed by Kristian Kastet, was transformed into C# and developed further by Laila Frotjold and Nils Magnus Djupvik. The Microsoft .NET framework was now used to communicate between the simulation model and the user interface.

Frotjold and Djupvik developed a web-service that controlled the transactions of data between the user interface and the underlying simulation model. The user interface was partly based on the Flash application of the previous interface, but some additional features were added in C#, such as message boxes, pop-up messages and a time slider that had previously been part of the Flash user interface was improved and implemented in C#. The development of the story generator will be discussed in Chapter 3.

A first evaluation, *Test 1*, of the Two-Shower prototype was conducted on a single-user version of the prototype by Frotjold (2005). This will be discussed in Chapter 6 of this thesis. Frotjold completed her Master's thesis in June 2005.

### 1.5.4 Fall 2005 to Spring 2006: Further Development of the Story Generator and the Return to a Collaborative System

Based on the findings from Frotjold's test I decided to make some changes to the user interface of the Two-Shower prototype and to develop it further. Frotjold had used a single-user version of the prototype and, because support for understanding complex dynamic systems through collaboration was part of the intention of the user interface design, it was decided to turn it back into a collaborative system. During the fall of 2005 and spring of 2006 the model was further developed by myself and Olve Sæther Hansen who was hired as an external programmer. The prototype was developed further and several of the technical difficulties were resolved. The story generator, which had previously only been text-based, was improved using graphics and animations. Some of the graphics, which at this stage were implemented partly by use of Flash MX and C#, were improved. The text generator was translated from Norwegian to English. The prototype that is primarily described in this thesis is this latest version developed in 2006 with the latest changes made by myself and Olve Sæther Hansen. Reference to earlier versions of the prototype will be indicated explicitly.

### 1.5.5 Fall 2005 to Spring of 2006: Test 2 of the Two-Shower Prototype

In parallel to the work on further development, plans were made for a second test of the prototype. I started with a theoretical study of how system dynamics models had previously been evaluated.

I also studied how *mental models* was used as the common concept of knowledge in system dynamics and how learning was seen as *changes* in mental models (see for example Doyle & Ford, 1998, 1999; Doyle, Ford, Radzicki, & Trees, 2001; Doyle, Radzicki, & Trees, 1998; Forrester, 1961, 1971a, 1980, 1994a; Forrester, 1994b, 1997; Lane, 1999; Morecroft, 1994; Richardson & Pugh, 1981; Sterman, 2000). I found there to be some problems in the use of the concept of mental models for my test as the goal of my test was not what participants learned, or how they had changed their preconceived ideas or *mental models*. My concern was to test the *efficiency* or *ease-of-use* of the prototype. I sought therefore to find alternative ways of evaluating the user interface of the prototype.

*Test 2* was formulated as a test to improve the user interface based on methods described in Nielsen (1993), Alessi & Trollip (2001), and Dix et al. (1998) (described in Chapter 7). The test was performed in the spring of 2006. Based on the results, which are discussed in Chapter 8, some suggestions for improvements were developed.

### 1.5.6 The Development of the Quito Prototype

In parallel to the development and evaluation of the Two-Shower prototype, work on the Quito prototype was carried out (see Skartveit et al., 2002). In the division of work, I have mainly concentrated on the Two-Shower prototype, while Hanne-Lovise Skartveit has focused on the Quito prototype. We have both been involved, however, in discussions and decisions regarding the two models throughout the development process.

Skartveit and I were responsible for determining the requirements for both versions of the Quito model and were involved in supervising the development. The first version of the system dynamic Quito model was, as mentioned previously, developed within the VOCS project by Masters students Porfirio Guevara-Chaves and Antonio Perez-Bennet (2002). Before Nils Magnus Djupvik and Laila Frotjold (2003) developed the second version of the model, I went through the model structure and behavior to find sections that needed improvement and further development.

The Quito prototype is currently implemented with a user interface that was developed by Hanne-Lovise Skartveit and Nils Magnus Djupvik. In Chapter 10 I provide some suggestions for how this user interface may be changed for the purpose of creating an application that could be used as part of a course in system dynamics based on the experiences from developing and testing the Two-Shower prototype.

## 1.6 Chapter Outline

This brief *Introduction* chapter lays out the goals of this thesis. The two prototypes in the VOCS project are introduced and the method on which this thesis was based is described.

*Chapter 2: Complex Dynamic Systems and the Original Two-Shower Model* begins with a further introduction to complex dynamic systems and the characteristics that make complex dynamic systems difficult to understand and control. This is important for the design of the prototypes. A short literature review of research on the understanding of complex dynamic systems within the system dynamic literature illustrates the difficulties that people have of understanding complex dynamic systems and how even small systems may be difficult to understand (see for example Diehl & Sterman, 1995; Dörner, 1997; Ford, 1999; E. Jensen & Brehmer, 2003; Moxnes, 1998a; Paich & Sterman, 1993; Sterman & Sweeney, 2002; Sweeney & Sterman, 2000). The chapter provides a further and more detailed introduction to the system dynamics model that is used in the first prototype, the original Two-Shower model, described by Morecroft et al. (1995). There is then an introduction to the field of system dynamics as this is the research field on which the development of the underlying simulation models of both, the Two-Shower prototype and the Quito prototype, is based (see for example Richardson, 1991; Richardson & Pugh, 1981). The chapter ends with a discussion of some limitations of the existing visualization methods in system dynamics.

*Chapter 3: Narrative Design* describes how the original Two-Shower model may be enhanced into a prototype for an interactive learning environment that includes elements of narrative. The first part of the chapter is a discussion of the considerations that must be made before developing an interface for a system dynamics model and how this may be done from both a technical and design point of view. Narrative and computer game theory is briefly discussed (see for example Bordwell & Thompson, 1997; Juul, 1998; Meadows, 2002). The major design decisions regarding narrativity are then considered. The last part of Chapter 3 describes a story generator which was developed as part of the model. The participants can here go back and study how their actions affected the behavior of the system.

*Chapter 4: Enhancing the Two-Shower model by a Multimedia User Interface* describes how a multimedia user interface with the purpose of supporting understanding was developed for the Two-Shower prototype. Theoretical issues of user interface design, visualizations, and the use of multimedia are discussed (see for example Alessi, 2002; Alessi & Trollip, 2001; Dix et al., 1998; Mayer, 2001, 2005; Mayer & Anderson, 1991; Spence, 2001; Tufte, 1997). There is then a presentation of how multimedia is used to represent structure and behavior in the user interface of the Two-Shower prototype.

*Chapter 5: Designing for Understanding* is concerned with how the original Two-Shower model was designed with the intention to support the participants in understanding the underlying model. Cognitive and constructivist theories of understanding and learning are discussed as are the difference between the view of mental models in cognitive theory and system dynamics (see for example Doyle & Ford, 1998, 1999; Doyle et al., 2001; Doyle et al., 1998; Forrester, 1961, 1971a, 1980, 1994a; Forrester, 1994b, 1997; Johnson-Laird, 1989; Lane, 1999; Miller & Johnson-Laird, 1976; Morecroft, 1994; Richardson & Pugh, 1981; Seel, 2003; Sterman, 2000). The constructivist view on learning and understanding is then discussed (Jonassen, Hernandez-Serrano, & Choi, 2000; Lave & Wenger, 1991; Salomon & Perkins, 1998) followed by description of a method for developing constructivist learning environments by Milrad, Spector and Davidsen (see for example Milrad, 2006; Milrad et al., 2003; Spector, 2000). Theoretical issues regarding model transparency and interactivity are raised (see for example Alessi, 2002; Davidsen, 2000; Davidsen, Spector, & Milrad, 1999; Größler, Maier, & Milling, 2000; Jensen, 2000; Machuca, 2000; Manovich, 2001; Spector & Davidsen, 1997b). The last part of the chapter deals with how the original Two-Shower model was transformed into a prototype of a *collaborative* interactive learning environment designed to support grounding through visualizations (see for example Arnseth & Solheim, 2002; Baker, Hansen, Joiner, & Traum, 1999; Dyck, Pinelle, Brown, & Gutwin, 2003; Feltovich, Spiro, Coulson, & Feltovich, 1996; Koschmann, 1996b; Koschmann et al., 1996; Timothy Koschmann & C.D. LeBaron, 2003).

*Chapter 6: Test 1 of the Two-Shower Prototype* describes the first test of the prototype. The chapter is a summary and discussion of the test that was performed by Frotjold (2005). The chapter ends with a discussion of lessons learnt from this test and how these were used in the planning of the second test, Test 2 of the Two-Shower prototype.

*Chapter 7: Test 2 of the Two-Shower Prototype* begins with a discussion of usability evaluation methods (Mack & Nielsen, 1994; Nielsen, 1993) before turning to *query techniques* (Dix et al., 1998) and *beta testing* (Alessi & Trollip, 2001). Changes that were made to the prototype between Test 1 and the second test Test 2 are then described The last section of the chapter describes the test design for Test 2 of the prototype.

*Chapter 8: The Results from Test 2 of the Two-Shower Prototype* deals with the results from Test 2 of the Two-Shower prototype. The results from the questionnaires and feedback from the participants are analyzed and discussed with respect to parts of the prototype that may be improved.

*Chapter 9: Future Work on the Two-Shower Prototype* is divided into two main sections. The first section is a discussion of how the design of the Two-Shower prototype may be improved based on the results from Test 2. The second part deals with lessons learnt from the test design of Test 2 and ways forward before additional issues regarding future testing is discussed.

*Chapter 10: Applying the Experiences from the Two-Shower Prototype to a Larger Model: The Quito Prototype* presents the second prototype in the project. As mentioned previously this is not as fully developed as the first model. I discuss how experiences from the Two-Shower prototype may apply to this model. The chapter begins with an introduction to the underlying system dynamic model. The Quito model is a larger and more complex model of a social system. This brings forward additional design issues that must be considered during development both regarding the organization of the user interface and the presentation of model characteristics. The chapter ends with a description of the current user interface of the Quito prototype, which was developed by Skartveit and Djupvik, and some suggestions for how the model may be made into a prototype for system dynamics students.

*Chapter 11: Lessons Learnt: Transferability* is a discussion of how the experiences and findings from the Two-Shower prototype and the Quito prototype may be used in the development of other interactive learning environments based on models of complex systems. The chapter ends with ways forward and some lessons learnt from the project.

### 1.6.1 Issues of Style and Terminology

I end this chapter with a short note on some issues of style and terminology. In this thesis I try to avoid extensive use of abbreviations. Some abbreviations, however, are used to facilitate reading. I will, for example, mainly refer to the abbreviation for Computer Support for Collaboration and Learning (CSCL) because the original term is long and there has been some dispute within the CSCL community regarding the exact wording (Koschmann, 1996b).

Regarding terminology, there is a need to discuss what I mean by the term *model* in this thesis. A model is a representation of a theory, in this case a theory of the structure of a system, how components of the system affect one another through causal relationships to generate the resulting behavior of the system. The term *model* in this thesis refers to a *simulation* model that can be run to generate a representation of the systems' behavior over time.

By *interactive learning environment* I refer primarily to the technological software products that are developed with the purpose of supporting learning. This definition does not include, for example, the context in which it is used or the learners that participate in using the interactive learning environment. The term interactive learning environment in my use here also does not refer to a complete course such as, for example, an online course. It refers rather to software products that may be used as part of a course but that needs additional supplement provided by a facilitator or teacher.

When I refer to *simulation-based interactive learning environments* I refer to learning environments based on simulation models that are developed with the purpose of learning about or understanding complex dynamic systems.

By *system dynamic interactive learning environments* I refer to interactive learning environments where the purpose is to learn about system dynamics. These interactive learning environments may be used as part of a course where the purpose is to learn system dynamics.

In the text, references to web addresses are footnoted, while the APA standard is used for literature references.

With this brief comment on style and terminology I move on to Chapter 2 and a further introduction to the Two-Shower model and complex dynamic systems.

# 2. Complex Dynamic Systems and the Original Two-Shower Model

## 2.1 Introduction

The purpose of this chapter is to provide a more thorough description of the Two-Shower model. A presentation of its structure and complexity provides also a further introduction to some of the common problems related to the understanding of such systems. The goal of the chapter is to provide a background to some of the problems and issues relating to the model and the understanding of complex dynamic systems that had to be dealt with when developing the interfaces for the prototypes described in this thesis.

The chapter is divided into this short introduction, 2.1, and five main sections. Section 2.2 summarizes the characteristics of complex dynamic systems and point to some difficulties in understanding them. Section 2.3 considers previous research on the understanding of complex dynamic systems and how people have problems understanding even basic components of complex dynamic systems. Section 2.4 is a description of the original version of the Two-Shower model that was presented in Morecroft, Larsen, Lomi, & Ginsberg (1995). Section 2.5 is a description of the system dynamic method, the process of developing a system dynamic model and the various visualization methods that traditionally have been utilized in system dynamics. The section draws on the seven steps of the system dynamic method, presented by Richardson & Pugh (1981). Section 2.6 is a discussion of the traditional visualization techniques and potential problems regarding the way they portray complex dynamic systems.

## 2.2 Complex Dynamic Systems

Complex dynamic systems and some of the problems of complex dynamic systems were briefly introduced in Chapter 1. This section is a further explanation of the characteristics of complex dynamic systems and the challenges one faces when dealing with such systems. These challenges have to be taken into account when developing a prototype based on a model of a complex dynamic system. The main structural characteristics that make a complex dynamic system difficult to understand are the relationship between stocks and flows and the resulting time delays, feedback, and nonlinearity. The dynamics of the system arise as a consequence of the interaction between coupled negative and positive feedback loops caused by multiple time delays, nonlinearities, and accumulations that constitute the feedback structures. It is important to look at complex dynamic systems as a synthesis of these characteristics.

### 2.2.1 Stocks and Flows

*Stocks* are accumulations that are modified by its *inflows* and *outflows*. Such an accumulation may, for example, be the population of a city. A population changes based on its birth/immigration rate (inflow) and its death/emigration rate (outflow). The changes in the level of the accumulation, which is determined by its rates, can be expressed in a graph that shows the level over time.

Any understanding of a dynamic system, however complex, must be grounded in an understanding of the fundamental process of integrating flows into stocks. The state of any system is *qualitatively* represented by stocks - each of which represents a state variable, and *quantitatively* by the levels of those stocks. In a static system, all stock levels remain the same. In a dynamic system they change over time. These changes are *qualitatively* represented by flows affecting (accumulating into or draining) the stocks, - causing the stock levels to change and *quantitatively* represented by the rates of these flows, i.e. the rates at which these changes take place. This impact of flow (-rates) on stock (-levels) is the key component in the structure of a dynamic system.

Although the flow rates may be set in an instance, they characterize changes that can only take place over *time*. Stock levels take time to change. Accordingly, time must pass for the stock level to respond. Instantaneously there is no change in the stock level. The impact applies over time.

Thus the primary challenge in understanding complex, dynamic systems is to understand the origin of the dynamics, i.e. the transition that takes place from flow rates that describe what takes place over time, to stock levels that respond accordingly to form the state of the system at time instances, - the integration of flows into stocks over time. One of the primary challenges for developing a user interface of a system dynamic model is thus to visualize this process.

### 2.2.2 Delays

Stocks create *delays* in the system by accumulating the difference between their inflows and outflows. As mentioned before, the accumulation of a stock is something that occurs over *time* and delays are therefore implied by the relationship between flows and stocks.

The consequence of the delays in the system is that the effect of a change in one variable is not seen instantaneously. Due to the intervening integration processes it takes *time* for the variable influenced to react to changes in the influencing variable. For example, an increase in the demand for housing in one area of a city usually does not occur immediately after the prices decrease. It takes time for the market to respond to the lower prices. The problem of understanding delays is that the effect of a change is not seen immediately and changes in variables may be caused by previous changes in other variables. It may therefore be difficult to diagnose the origin of a change in behavior. In addition, delays make a complex dynamic system difficult to control because people may not take to a sufficient extent the delays into consideration when performing actions and making decisions. They tend to underestimate rather than overestimate (Sterman & Sweeney, 2002). Delays make it difficult to infer the effects of changing a variable and to make decisions regarding the timing of a change. A user interface of a system dynamic model must therefore portray the delays in the system.

### 2.2.3 Feedback

According to Sterman (2000) the most complex system behavior usually arises from the nonlinearities between the variables in the system, not from the complexity of the structure alone. Feedback is created in response to the behavior of the variables in a system.

Feedback means that a change in one variable causes another variable to change and that the change in the second variable again affects the first variable. The prices of housing in a city may for example influence the demand on housing which again influences the prices of housing (see figure 2-1). If the prices of housing decline in one part of a city, more people may want to move there. After a while there are few vacant houses and the large demand causes the prices to increase. The result may be that the decline in the prices of housing feeds back and after some time causes the prices of housing to fall less, stabilize and possibly increase again.

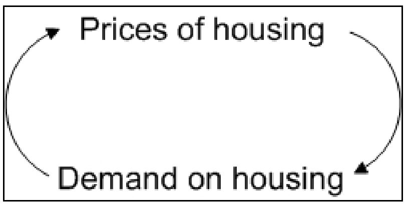


*Figure 2-1: Example of feedback.*

In a feedback system the variables are interconnected in feedback loops, that altogether constitute the structure of the system. *"A feedback loop is a closed sequence of causes and effects, a closed path of action and information. An interconnected set of feedback loops is a feedback system"* (Richardson & Pugh, 1981, p. 4, emphasized in original). In fact, a much more complex feedback system than the one described above controls the prices of housing beyond the demand. Many factors determine the prices and control how they increase and decline. These factors influence each other in various ways often by way of feedback. The prices of housing, for example, may be influenced by the demand of housing. A change in the prices of housing thus cause the demand to change. A change in demand will again affect the prices of housing. This is an example of one feedback loop. There are, however many feedback loops in a complex dynamic system and the prices of housing may be influenced by, for example, the prices of building new houses, the prices of land, and how many new houses are built. All these variables are part of a feedback system that contributes to control the prices of housing and where a change in the prices of housing will cause the other variables to change and after some time change the prices of housing.

A problem of understanding feedback is the circularity in the arguments or reasoning. It may be difficult to determine where to start. Generally stocks are good starting and halting points because they do not change instantaneously since the accumulation process takes time. A representation of feedback must also be embedded into a user interface of a system dynamic interactive learning environment.

### 2.2.4 Nonlinearity

As previously mentioned, each variable is not influenced by a single variable, but by a set of interacting variables. The system is nonlinear. This means that the effect of changing the value of variable A is determined by the state of the system, that is, either by the current value of variable A or by the current values of other variables.

The prices of housing in a city may be influenced by the demand on housing and the costs of constructing new houses. The price effect of a change in demand depends on the current demand and the current price, in addition to all the states of the other variables in the system. If the demand currently were 2000, an increase of 2000 would represent a large change and have a relatively large effect on the prices of housing given that the supply is constant. If the demand were 100 000 then an increase in 2000 would have a comparably small effect on the prices. This is because the state of the variable is different for the two situations. In addition, if the prices of construction are relatively low a change in demand will have a relatively large effect on the prices of housing. If the costs of construction were high, the same change in demand would have a comparably lower effect on the prices of housing. This shows that also the states of other variables in the system are important in determining the effect of a change in a particular variable.

The nonlinearity of the system thus makes it important to consider the current state of the system when changing variables. Because the state changes it becomes difficult to infer the consequences that a change in a variable may have on the system behavior over time. The visualization of nonlinearity must be taken into consideration when developing the user interface.

This section has shown that the characteristics of complex dynamic systems are intricate and that there are, therefore, serious challenges regarding the representation of complex dynamic systems. The next section deals with previous research on the understanding of complex dynamic systems. From this, it may be deduced that there are yet challenges to overcome in the representations and communication about such systems.

## 2.3 Previous Research on the Understanding of Complex Dynamic Systems

Although this thesis does not include an attempt to study people's understanding of complex dynamic systems, I have considered a review of some of the literature on such research as necessary. The literature shows that there are problems in understanding such systems, and although speculative, a step towards improved understanding could be to use new and different forms of visual and multimedia support to endorse understanding. Previous research also justifies the choice of a small, but complex model, like the Two-Shower model, to introduce system dynamics because the literature documents that people have problems understanding even basic components of a complex dynamic system.

Paich & Sterman (1993) studied people's problems of learning in experimental markets. They developed a management flight simulator where subjects were to manage a new product. The model was small, consisting of a company, its market, and its competition. The experiment showed that, even though the model was small, the subjects overlooked the feedback processes, time delays, and nonlinearities.

Diehl & Sterman (1995) observed similar problems in their study of subjects who managed an inventory with stochastic sales. They varied the strength of the feedback and the time delays and found that performance was not optimal across any conditions. The subjects undercontrolled the system as the feedback grew stronger and the delays grew longer. Diehl & Sterman (ibid.) states: *"Subjects' understanding of complex feedback settings declines as delays between cause and effect increase and as actions have stronger side effects"* (p. 198).

Dörner (1997) studied problem-solving in relation to simulation games. The participants were to control a fictive region in West Africa. The participants did not consider long time delays when making their decisions. When the food supply increased the size of the population would gradually increase. Most participants did not anticipate that the growing population would lead to a critical shortage in the food supply, thus leading to a famine. According to Dörner we need what he refers to as *structural knowledge* (p. 41) in order to be able to control complex systems. He argues:

> *"If we want to operate within a complex and dynamic system, we have to know not only what its current status is but what its status will be or could be in the future, and we have to know how certain actions we take will influence the situation (p. 41).*

According to Dörner the problem, however, is that our knowledge about the structure and its related components are often wrong.

Moxnes (1998a) asked people from the fisheries sector in Norway to manage a fish stock in a fish stock model. A decision panel displayed the variables that could be controlled by the subjects. The subjects had a tendency to over-invest leading to a large over-capacity. Moxnes found that the subjects overlooked important feedback in the system, which led to an average over-capacity of 60%.

Sweeney & Sterman (2000) also studied the understanding of stock and flow relationships and time delays of complex dynamic systems. A group of students was given a figure of a bathtub and a trajectory of the inflow and outflows of the bathtub. They were asked to draw the trajectory of the water in the bathtub over time. Another task consisted of drawing a company's cash balance. A figure of receipts and expenditures going in and out from a safe and a trajectory of the flows were provided. Sweeney & Sterman (ibid.) state that *"initial findings indicate that subjects from an elite business school with essentially no prior exposure to system dynamics concepts have a poor level of understanding of stock and flow relationships and time delays"* (p. 249). This study showed that the subjects had problems understanding the system even though the model was small and consisted of a small set of variables.

Sterman & Sweeney (2002) further studied intuitive understanding of $CO^2$ pollution and its effect on global warming. The focus of the experiment was whether people were able to perceive and take into account the delay that exists from the expulsion of $CO^2$ to its effect on the global temperature. The subjects were given a stock and flow diagram of the system and were asked to draw trajectories of the anthropogenic $CO^2$ emissions, the carbon dioxide in the atmosphere, and the global mean temperature. According to Sterman & Sweeny (ibid.) no mathematics were required to solve the tasks. The subjects only needed an understanding of stocks and flows and fundamental facts regarding climate change. Nevertheless highly educated the subjects were not able to understand the system. The study showed that the *mental models* people used were incorrect and would even violate basic laws of physics. One of the problems was that they would consistently underestimate the delay between changes in the $CO^2$ concentration and the temperature. The perception was that the temperature would respond too fast and too much. According to Sterman & Sweeny (ibid.) their study showed that *"...people's intuitive understanding of even the simplest dynamic systems is poor"* (p. 234).

Jensen & Brehmer (2003) studied laypeople's understanding of a simple dynamic system. The subjects were asked to control a model of a simple predator-prey system consisting of one balancing loop and two reinforcing loops. The system was represented by a stock and flow diagram and graphs. Jensen and Brehmer (ibid.) found that the subjects' *"...difficulties deemed to stem from a low ability to apply indirect reasoning and thinking in terms of discrete time steps instead of in terms of continuous time"* (p. 119). They conclude that even though the structure of the system is relatively small, consisting of only two stocks, it is still difficult. The distinction between this study and the ones previously described is that the subjects' understanding of the system is studied through expressions of reasoning and strategies used by the subjects and how this affected performance, not merely on performance.

Studies of complex dynamic systems have revealed that people seem to have a general problem of understanding such systems and to infer their behavior. There is , however, little focus on to what degree the interface contributes to people's understanding or misunderstanding. None of the studies raise critical questions about *the way* they have chosen to portray the systems. I would argue that it would be beneficial to discuss the understanding of these systems with a discussion of how the systems are represented both graphically and textually - which information is given, how, and when. The participants' experience with these presentation tools may also be an important factor for their understanding of the system and should be taken into account when performing such studies. The users' ability to understand the system may be linked to the way it is represented.

This literature review shows that people have problems understanding *small*, complex dynamic systems or even components of small, complex dynamic systems. It is therefore adequate, as argued previously, to use a small, complex dynamic system as a starting point for studying visualization of complex dynamic systems.

The following section is a description of the original two-shower model, which is the small system dynamic model on which our first prototype is based. The model is used as an example of applying the system dynamics approach to the study of complex dynamic systems in Section 2.5.

## 2.4 The Original Two-Shower Model

The original two-shower model is, as described previously, a model about resource sharing, in this case, a hot water tank that is shared between two showers (see figure 2-2). The model can be seen as a metaphor for resource sharing in general.

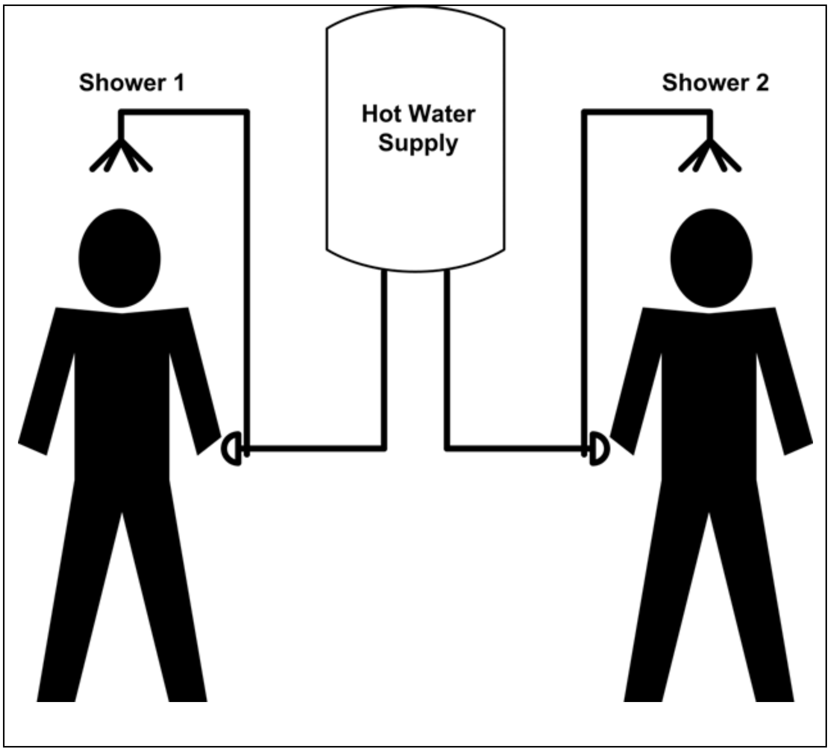


*Figure 2-2: Overview of the original two-shower model.*

The structure of the original model consists of two showers that are connected through a pipe structure to one hot water tank and the decisions and thought processes made by two *simulated people* whom each *control* the tap setting (mixing battery) of a shower. As briefly described in Chapter 1, the system contains *delays*. There is a delay of four seconds from a change is made in the tap setting until the water with the new temperature reaches the showerhead. This delay exists for both showers.

The original two-shower model also contains *feedback*. The change of the temperature at the showerhead triggers a *simulated human decision process* that represents how a person may react to the changes in the shower temperature. If there is a gap in the desired and actual temperature of the showerhead, the tap setting is again changed. A change in tap setting will thus feed back and have consequences for the shower of the person that changed the tap setting. This is nonlinearity in the system, represented by the interaction between two negative feedback loops.

A change in tap setting will also change the conditions of the other shower, which typically will lead to a change in tap setting done by the person in this other shower. This decision will next time around have consequences for the water temperature in the first shower.

In addition, the model is *nonlinear*. The access that each shower has to the hot water resource depends on the tap settings of both showers, that is, how much demand they each put on the resource. The implication is that the effect of a change in the tap setting depends on the current state of both tap settings. If both showers have a low tap setting, a small change in the tap setting of Shower 1 will have a relatively larger effect on the water temperature, than for example if Shower 2 has a high tap setting and already demands a large portion of the hot water resource. As a consequence the same tap setting or change in tap setting will not always lead to the same temperature or the same change in the temperature. It is due to this that a change in tap setting will change the condition of the other shower.

When the two *simulated people* in the showers are *unaware* of each other and do not take the conditions and decisions of the other into account, the temperatures of both showers oscillate, and they may both be unable to sustain a stable, comfortable temperature. A balance in behavior is only possible when there is balance in both subsystems. Balance in this case refers to when the desired temperature equals the actual temperatures of both showers.

The original two-shower model is a metaphorical model. It illustrates the consequences of mutual dependency of decision-makers related to resource allocation. In the paper by Morecroft et al. (1995) about the original model there is also a focus on the changes that can be made in order to improve the coordination of decision-makers' actions. The main focus of the paper is on resource allocation within businesses and how it is difficult for decision-makers to infer the consequences that their decisions have for other decision-makers within or outside of the organization. It also takes time before the consequences of their actions are seen. It may be a problem that decision-makers support local interests (for example within a department) when allocating resources, -and not take into account that the result may turn out to be collectively worse (for example for the company as a whole).

The original two-shower model applies to situations where the decision-makers themselves are part of the system structure and their decisions, after a delay, will feed back and affect their own conditions. In these situations, a demand on a large portion of a resource leaves less to others. When the effect of a decision influences the conditions of another decision-maker, the other decision-maker will respond accordingly and base her decisions on the new situation (the new state of the system). The effect of the other decision-maker's decisions will feed back and change the situation for the first decision-maker.

In the original Two-Shower model it is presumed that the quantity of the shared resource (the flow of hot water) is fixed. The total level of the hot water resource remains the same, although it is divided among the parties based on their claim on the resource. In real life, it is difficult to find examples of resources that are completely fixed because they are usually somehow connected to a greater system of causal relationships. An example of a shared, relatively fixed resource is the world oil reserves (although all oil reserves are not yet discovered and the amount that can be exploited changes based on technological advances). If one country demands a large portion of the oil reserves hostility from other countries may erupt and they may take action against the first country. This action may in some way result in the first country having access to less oil reserves.

In real life there are numerous examples of cases where the level of a shared resource is not fixed, but where decisions of the decision-makers also affect the total level of the resource. An example of a shared resource is an internal monetary resource of a company. The production department of a manufacturing company, for example, may demand so much of a resource that the marketing department does not have enough to market the products that were produced. The great demand from the production department then feeds back and alters their conditions when the products that they have produced are not sold due to lack of money spent on marketing. This will again affect the total level of resources available for the company.

The original Two-Shower model cannot be directly compared to most real-life resource sharing situations. This is also not the intention of the authors:

> *"The Two-shower model provides a bridge from the dynamics of a tangible everyday system to the dynamics of complex organizational systems. Like all good metaphors its effectiveness lies in the power of the images that it conjures up in the mind of the recipient and the comparisons or associations that these images provoke" (Morecroft et al., 1995, p. 305).*

Without knowledge of system dynamics, it may be difficult to understand how this type of model is constructed and how it functions. The following is a step-by-step introduction to system dynamics with the purpose of presenting the method which is used in the development of both underlying simulation models used in the interactive learning environments of this project. This presentation also involves a more detailed description of the original Two-shower model and an introduction to some of the problems faced in our second prototype, the Quito prototype, which will be discussed in Chapter 10.

## 2.5 System Dynamics

This section is a relatively brief introduction to the system dynamics method. Both, the original two-shower model described in Morecroft et al. (1995) and the simulation model on which the Quito prototype is based, were developed by use of the system dynamic method. System dynamics use terms, symbols, and concepts, which are not followed strictly in this thesis.

After a brief introduction to the research field of system dynamics I go through the different steps that modelers generally follow when developing and analyzing system dynamic models. I continue to use the original two-shower model to illustrate this process.

The system dynamic method consists of seven steps:

1. *problem identification and definition*
2. *system conceptualization*
3. *model formulation*
4. *analysis of model behavior*
5. *model evaluation*
6. *policy analysis*
7. *model use or implementation*

*(Richardson & Pugh, 1981, p. 16).*

It is important to note that these steps are not fixed. Changes may for example be made to the model during the model analysis and model evaluation is usually performed to some extent throughout the modeling process. In addition, system dynamic modeling is an iterative process and the different steps may be performed several times (ibid.).

### 2.5.1 A Short Introduction to System Dynamics

For those unfamiliar with system dynamics I offer a brief introduction to the field. System dynamics is a research tradition that aims to identify, understand, and utilize the relationship between structure and behavior in complex, dynamic systems. *"It takes the philosophical position that feedback structures are responsible for the changes we experience over time. The premise is that dynamic behavior is a consequence of system structure"* (Richardson & Pugh, 1981, p. 15, emphasized in original). System dynamics is a research field that studies complex dynamic systems by way of computerized modeling and simulation. *"The expressed goal of the system dynamics approach is understanding how a system's feedback structure gives rise to its dynamic behavior"* (Richardson, 1991).

In short, the system dynamic method starts with the need to examine a problem. The problem is believed to be affected by causal influence between variables within a complex dynamic system over time. The system and the variables are studied in order to map the structure of the problem and a computerized simulation model is developed based on the findings. The model is then studied and policies for how to solve the problem are tested in the model before being implemented in the real system.

*"System dynamics is a method to enhance learning in complex systems"* (Sterman, 2000, p. 4). Because the aim of modeling and using simulation models is to enhance learning, the study of what a person learns through modeling and simulation becomes important. How actors within the system make decisions, which again change the system, is also of importance. Within system dynamics this is addressed by studying what is referred to as people's *mental models*, how they are used as a basis for decision-making, and how they change as a result of experience with modeling or using system dynamic models. This will be discussed further in Chapter 5.

System dynamics emerged from control theory in the late 1950s and early 1960s. The research tradition has its roots at M.I.T., under the leadership of Jay W. Forrester. His book *Industrial Dynamics* (1961) is deemed the first main contribution to the system dynamic literature. Some other important publications are Forrester (1969; 1971b), Richardson and Pugh (1981), Senge (1990), Ford (1999), and Sterman (2000). The focus of the research started with a need to obtain more efficient business-solutions (see Forrester, 1961) and much of the recent literature is still devoted to business-related problems (see for example Akkermans, 2001; Bianchi, 2002; Campbell, 2001; Lyneis, Cooper, & Els, 2001). As the field developed, Forrester applied it to other areas such as the problems of aging urban areas (Forrester, 1969), the socioeconomic system of the world (Forrester, 1971b), and his more recent interest in the introduction of system dynamics to students ranging from kindergarten to high school level (Forrester, 1997). In 2005 the *System Dynamics Society* had more than one thousand members located in 55 countries[2] and recent literature presents studies covering a wide area of topics, such as healthcare (see for example Homer, Hirsch, Minniti, & Pierson, 2004; Liddell & Powell, 2004), environmental issues (see for example Arquitt & Johnstone, 2004), problems regarding natural resources (see for example Faust, Jackson, Ford, Earnhardt, & Thompson, 2004; Moxnes, 1998b), and the introduction of laws in societies (see for example Sridharan & Hunt, 2000).

### 2.5.2 Problem Identification and Definition

The system dynamic method starts with the need to examine a *problem* (Richardson & Pugh, 1981). Initial *questions* arising for this problem may be: *does the introduction of a new law lead to less criminal activity within a certain geographical area? Is the hospital utilizing its resources effectively? Which direction will the prices of housing turn within the next years?* It is important to note that system dynamicists are not interested in finding exact answers such as the exact interest rate in two and a half years. Rather than focusing on point prediction there is an interest in studying the underlying mechanisms that cause changes in the interest rate, be able to predict the direction it will head, and through an understanding of the system influence it in the desired direction.

The *question* that started the process of developing the original two-shower model could be: *Why does the temperature of the hot water oscillate?* Because this is a

[2] http://www.systemdynamics.org/index.html (accessed November 5, 2005).

metaphorical model, however, this was probably not the initial question that was asked before deciding to develop the model. The question could rather be: *How can the oscillations in resource availability be explained in a simple manner with the purpose of encouraging decision-makers to consider and discuss why their access to resources may oscillate?*

The original two-shower model is used as an example throughout this section. However, because the model is metaphorical it does not apply to all steps in the system dynamic method. I therefore use a problem of a polluted city as an additional illustrative example as this can also be related to our second prototype, the Quito prototype. In the Quito model, part of the concern is with air pollution from cars, where the city government wishes to study the effect of traffic on the air pollution level and to find policies that may decrease the air pollution from cars in the city. The *question* may be: *How may the level of pollution in the air be decreased?*

According to the system dynamic method there must also be a statement of the *purpose* of the model in order to define the intended *audience* or users (Richardson & Pugh, 1981). The *purpose* of examining the air pollution problem would probably be to find ways to decrease the pollution. The intended *audience* would then be the people who have the power to influence the problem, that is, the involved city government officials, heads of infrastructure, public transportation leaders, environmental research labs, and maybe also environmental organizations. As for the original two-shower model the intended audience could be anyone who deals with resource sharing.

### 2.5.3 System Conceptualization

When the problem is identified and defined the second step in a system dynamic study is *conceptualization*. The relevant variables are identified, the time horizon for which it is adequate to study the problem is determined, model boundaries are discussed, and reference modes for the important variables in the system are developed. Defining the relevant variables often starts with a consideration of which stocks that exist in the system. Examples of stocks may be the population of a city, number of cars in a city, or the amount of unexploited oil reserves in a country. As discussed previously in this chapter, the level of a stock is controlled by its rates. The birthrate and immigration rate, for example, may represent the inflow rate to a city, while the death rate and emigration rate represent the outflow rate. Reference modes, a set of graphs that are used to show the assumed behavior of the most important variables over time, are developed (Sterman, 2000). In the example of the polluted city where the city government wishes to decrease pollution from cars, relevant variables may be cars, citizens, buses, other public transportation, road infrastructure/capacity, road toll fees, and variables related to the climate of the city. The relevant time horizon could be, for example, anywhere from ten to fifty years depending on the time period in which they want to reach their goal. Reference modes of the assumed behavior over time could be developed for variables such as the development of pollutants in the air, number of cars, number of passengers utilizing public transportation, population, etc.

In the case of the original two-shower model the relevant variables are among others the two simulated people in their showers, their tap settings, the temperature at their showerheads, and various variables determining the decisions of the people in each shower. The time horizon of the original two-shower model is much shorter than for the air pollution example. It will be the time it takes to have a shower, for example anywhere from one minute to twenty. Adequate units of measure will be minutes or seconds. Reference modes for the original two-shower model could be made of changes in tap setting and the temperature at the showerhead. It could also be advantageous to try to infer how the temperatures of the two showers behave in relation to each other. For an example of a reference mode, see figure 2-3.

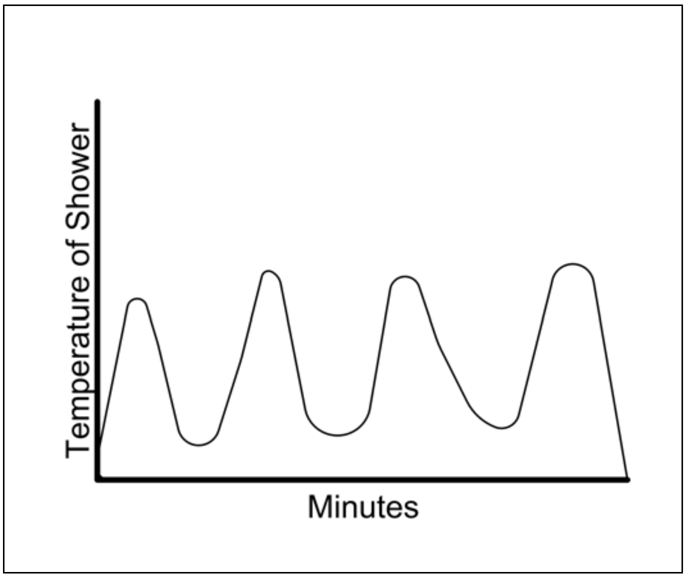


*Figure 2-3: Example of reference mode.*

Determining the system boundary is also part of the conceptualization phase. Within the boundary are the variables that are believed to influence the problem. The variables that are believed not to influence the system, or believed not to influence the system in a manner that has a substantial impact on the problem, are omitted. When modeling the problem of car pollutants in the air of a city, pollution from industry may for example be omitted from the model, not because it does not contribute to the air pollution level, but because the focus of the study may be to find policies to lower pollution from cars through, for example, public transportation.

The system boundary of the original two-shower model includes the decision processes that the persons who control the tap settings are assumed to go through. Variables concerning for example the pressure of the cold water are omitted.

The problem is then conceptualized through defining the causal structure of the model. This is performed by studying the main relevant variables and discussing how they may influence each other through causal relationships. Whether the influence is positive or negative, that is, whether a change in one variable leads another variable to change in the same or opposite direction is also discussed. In the example of the polluted city there may be a causal relationship between the number of cars on the roads and the level of pollution in the air. There may also be a causal relationship between fares of public transportation and the number of people who choose to use public transportation instead of cars. The relationship between the number of cars on the roads and the level of the pollution may be positive in that the more cars there are on the roads, the higher the level of pollution there will be. If the number of cars decreases, the pollution level will also decrease (see figure 2-4).

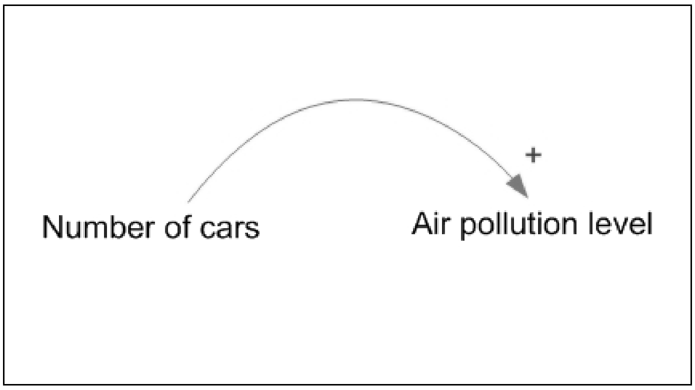


*Figure 2-4: Example of a positive causal relationship.*

On the contrary, the relationship between the fares of public transportation and the number of people who utilize public transportation is likely to be negative. This means that a change in fares has an opposite effect on the number of people who utilize public transportation. If the fares are raised it is likely that less people will choose to use public transportation as a means of travel. If the fares are lowered it is likely that more people will use public transportation than otherwise (see figure 2-5).

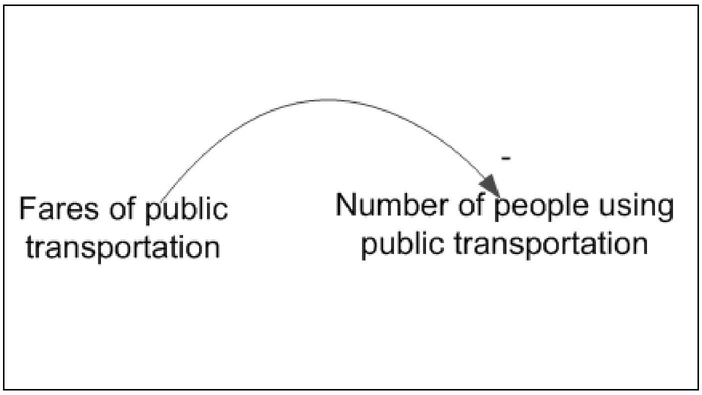


*Figure 2-5: Example of negative causal relationship.*

As for the original two-shower model there is a positive relationship between the temperature at the tap and the temperature at the showerhead. If the temperature at the tap decreases the temperature at the showerhead will also decrease after some time. Similarly, if the temperature at the tap increases the temperature at the showerhead will also increase after some time. There is a negative relationship between the tap setting of one shower and the portion of hot water that is available to the other shower. This means that the higher the tap setting of one shower is, the smaller the portion of hot water there is available to the other shower. Likewise, the lower the tap setting of one shower, the larger the portion of hot water there is available to the person in the other shower (see figure 2-6).

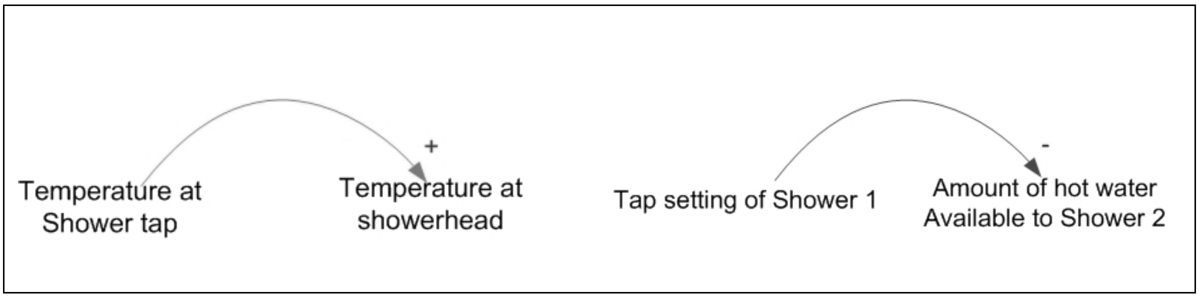


*Figure 2-6: Examples of positive and negative causal relationships in the original two-shower model*

These variables and their connections through causal relationships can be seen as the basic building blocks of a complex dynamic system that is portrayed by use of the system dynamic method. They are used to form causal loop diagrams (CLDs) and stock and flow diagrams (see figure 2-11 p. 65 and figure 2-14 p. 68).

The following section describes the development of causal loop diagrams while Section 2.5.5 describes how the conceptual model is transformed into a formal simulation model portrayed by a stock and flow diagram.

### 2.5.4 End of Conceptualization Phase: Causal Loop Diagram

Part of the conceptualization phase is to develop a *causal loop diagram (CLD)* of the system. In system dynamics a causal loop diagram is used to obtain an overview of the problem and the causal relationships that exist between variables in the system. The causal loop diagram was developed in order to communicate models better to a wider audience without a background in engineering or system dynamics (preface by Sterman in Richardson, 1986). In addition to the causal relationships between all the important variables in the system the causal loop diagram also includes the polarity of these relationships. When representing the causal relationships in such a diagram one can see that they sometimes form loops, or what is referred to as *feedback loops*. The feedback loops are marked with a circular arrow with an *R* or a *B* placed inside. The *R* stands for *reinforcing* which means that the causal links between the variables cause the variables to behave in the same direction as they are already heading and that they reinforce the behavior of each other. Returning to the example of air pollution and public transportation there could be a reinforcing loop between public transportation fares and the number of people who use public transportation (see figure 2-7). If the fares of public transportation for some reason are raised it may lead to a lowering in the number of people who use public transportation. A lowering in the number of people who use public transportation results in fewer passengers to divide the costs among so the public transportation company may raise the fares to cover the expenses. Another raise may lead to even fewer passengers and it may be necessary to raise the fares even further. Note that even though the causal links between the variables are negative, that is, they influence each other in the opposite direction, the loop is still reinforcing. A feedback loop is reinforcing if there are an equal number of opposite/negative causal relationships (minus signs). The typical behavior of a reinforcing loop is exponential (see figure 2-8).

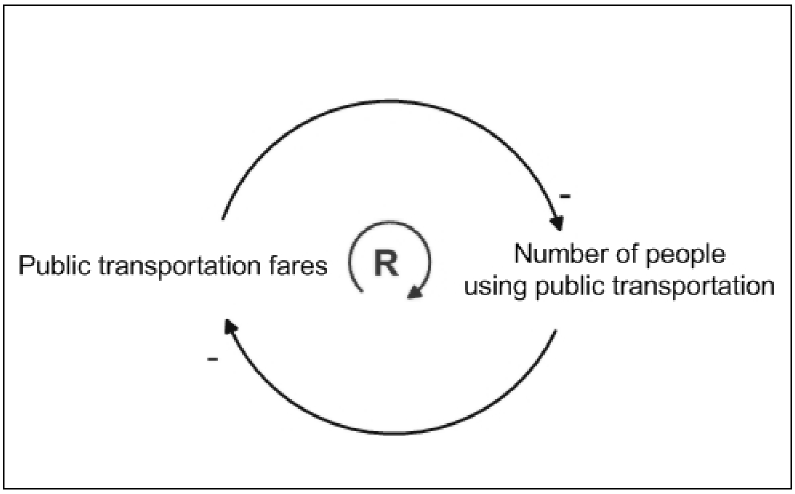


*Figure 2-7: Example of a reinforcing loop.*

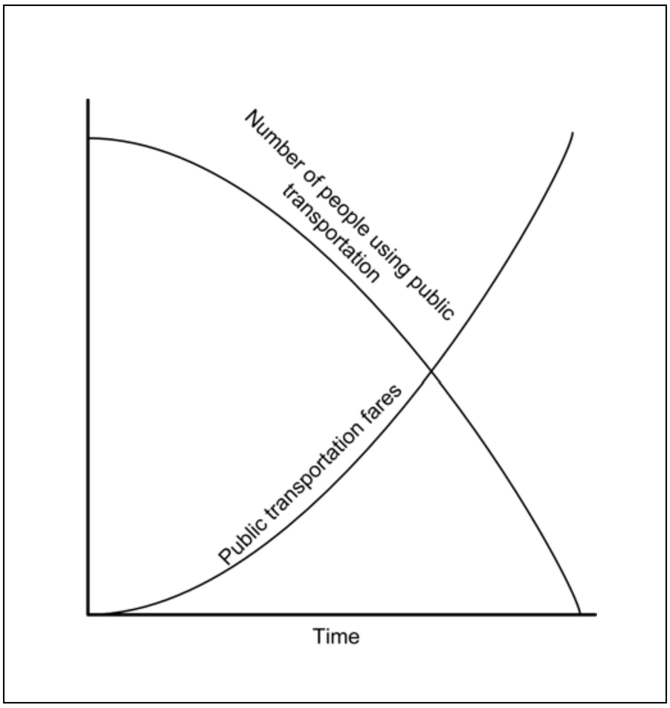


*Figure 2-8: Example of the behavior of an exponential feedback loop.*

Feedback loops marked with a *B* are called *balancing* loops. In contrast to reinforcing loops the causal relationships between the variables balance the behavior and prevent exponential behavior. In the pollution example there may be a loop that influences the number of cars on the roads where the traffic jam, or time spent in traffic jam, may influence people's desire to use their cars (if for example public transportation is faster) and there will be less cars on the roads (see figure 2-9). However, as there are less cars on the roads people may start to use their cars again because they will now spend less time in traffic jam and it may take less time to use private cars than public transportation. As the name indicates, the behavior of a balancing loop (if there are no other interferences) will find a balancing point and the behavior will be stable when this point is reached. For an example of the typical behavior of a balancing loop see figure 2-10.

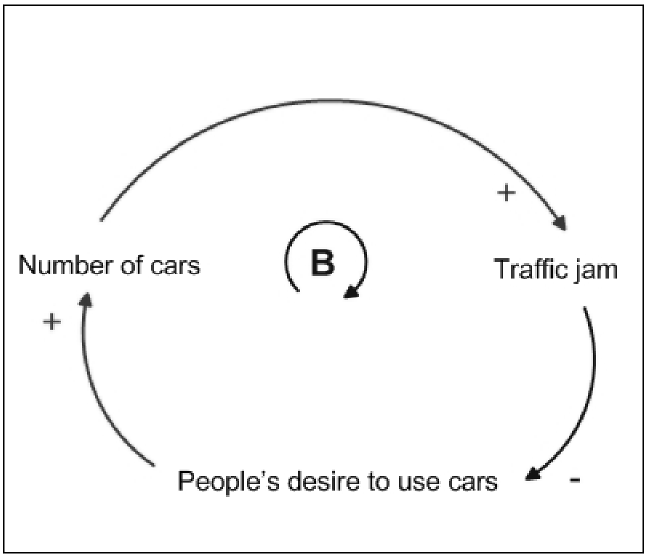


*Figure 2-9: Example of a balancing loop.*

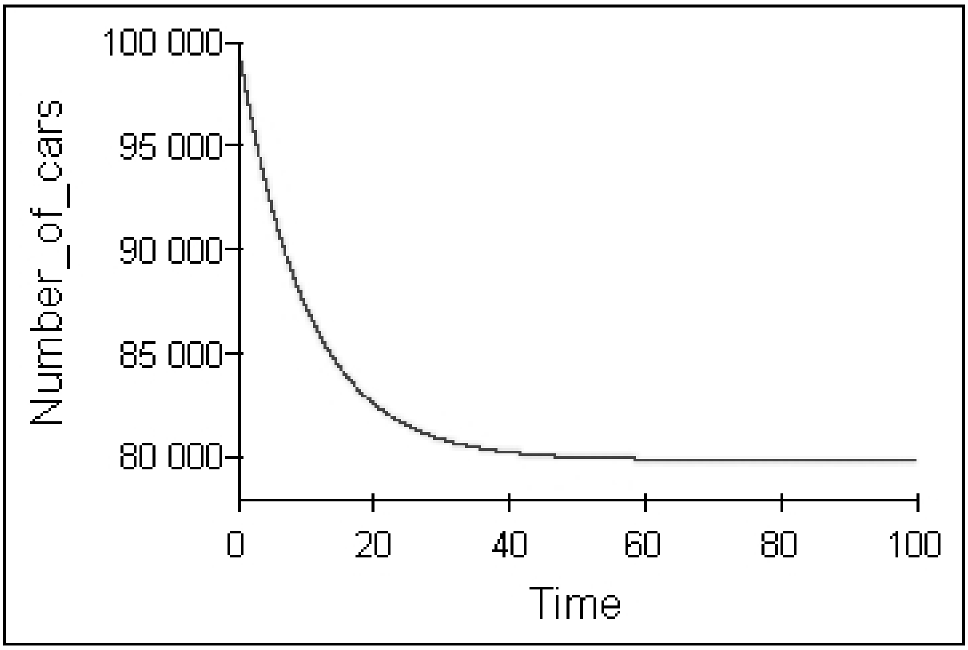


*Figure 2-10: Example of the typical behavior of a balancing loop.*

The causal loop diagram of the original two-shower model consists of three loops (see figure 2-11). There are two balancing loops that describe the adjustment of the tap settings of each shower and one reinforcing loop that describes how the two people in the showers both demand access to the hot water resource.

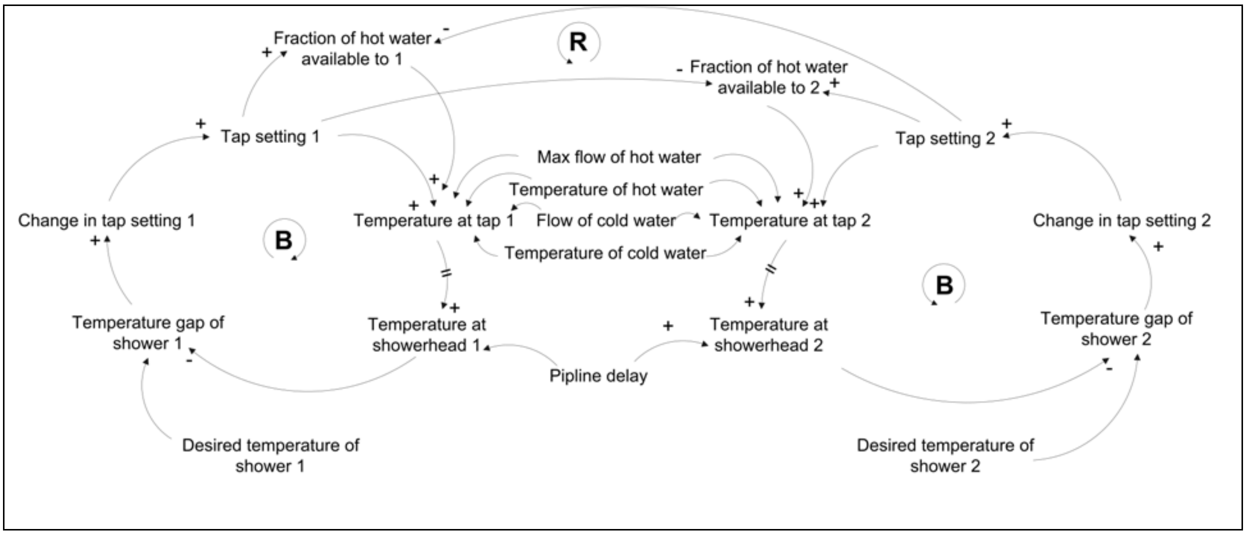


*Figure 2-11: Simplified Causal Loop Diagram (CLD) of the original two-shower model.*

*Balancing loop 1* describes how changes in the tap setting of *Shower 1* changes according to the desired water temperature. If the tap setting of *Shower 1* increases, the *Temperature at tap* of *Shower 1* will also increase. There is a delay between the *Temperature at tap 1* and the *Temperature at showerhead 1* because it takes time for the water to flow through the pipes. After some time the increase in the *Temperature at tap 1* will cause the *Temperature at showerhead 1* to increase. The simulated person in *Shower 1* compares the temperature at *Showerhead 1* to the *Desired temperature of shower 1* for each time step. When the *Temperature at showerhead 1* increases and there was previously a gap between the desired and actual temperature, the *Temperature gap of shower 1* will decrease. The *Change in tap setting 1* will then be lower than it would otherwise be (that is, if the gap had not decreased) and *Tap setting 1* will be lower than otherwise, as will the *Temperature at tap 1*. This is a balancing loop where a change in one variable causes the variable to change in the opposite direction and results in balancing behavior. *Balancing loop 2* is identical to *balancing loop 1*, but it regards the control of the temperature for *Shower 2*.

The *reinforcing loop* is the loop that connects the two showers and drives them to compete for the hot water resource. The main problem is that the temperature of each shower is determined by the tap settings of both showers. If the tap setting of one shower is changed, the temperature of the other shower also changes. For example, if *Tap setting 1* increases, the *Fraction of hot water available to 2* decreases. This again causes the *Temperature at tap 2* to decrease. A decrease in the *Temperature at tap 2* will after delay cause the *Temperature at showerhead 2* also to decrease. A decrease in the *Temperature at showerhead 2* again causes the *Temperature gap of shower 2* to increase. An increase in the temperature gap leads to an increase in the *Change in tap setting 2* which again causes *Tap setting 2* to increase. An increase in *Tap setting 2* causes the *Fraction of hot water available to 1* to decrease, and thereby decreases the *Temperature at tap 1*. Traversing the balancing loop of *Shower 1* with a decreased temperature will cause an increase in *Tap setting 1* and thereby an increase in the *Temperature at tap 1*. This completes the reinforcing loop and *Shower 2* will again try to increase the tap setting to obtain a larger portion of the hot water resource.

The reinforcing loop is controlled by the balancing loops. If there was no desired temperature the two simulated people would always struggle to obtain the largest portion of the hot water resource. The balancing loops result in lowering the tap settings as the temperature of the showers become higher than desired. The result is oscillatory behavior where none of the simulated persons in the showers are able to obtain a stable, comfortable temperature.

### 2.5.5 Model Formulation and the Development of Stock and Flow Diagrams

In a causal loop diagram there are usually a number of interconnected feedback loops and it is therefore not possible to infer the behavior of a system. The next step of the system dynamic method is to develop a formal, computerized model. This process is initiated by the development of *a stock and flow diagram*. Unlike the causal loop diagram the stock and flow diagram distinguishes between the different types of variables in the system. The causal loop diagram is not necessarily computerized but the stock and flow diagram is usually made in some kind of simulation software and the equations for each variable can be defined directly through the diagram.

As mentioned previously, one of the main structural components of a complex dynamic system are *stocks*, that is, reserves of some kind. In the pollution example, potential stocks in the system would be cars, buses, citizens, different pollution levels, and money available to spend on public transportation. As also discussed previously, the level of a stock changes through its inflows and outflows. The number of cars in a city is controlled by how many new cars are sold on the market (its inflow rate) and how many are wrecked (its outflow rate, see figure 2-12).

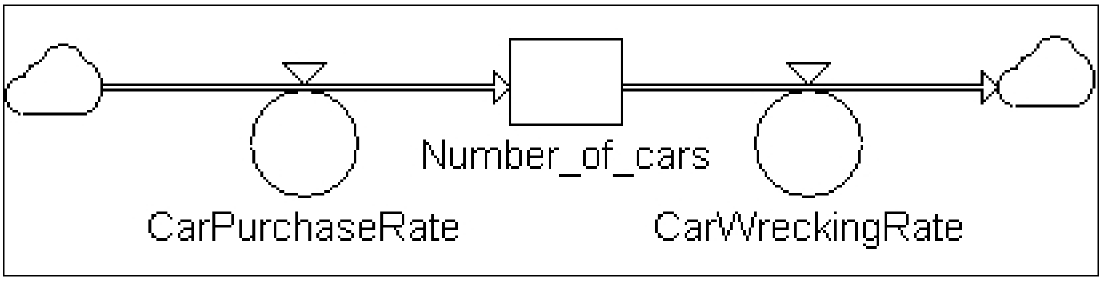


*Figure 2-12: Example of the representation of a stock and its inflow and outflow in a stock and flow diagram.*

The stock and flow diagram also shows variables that are called *auxiliaries*. These variables change instantly and unlike stocks do not accumulate over time. Auxiliaries are used to calculate the rates that change the levels of the stocks. Examples related to air pollution and public transportation may be the effect of public transportation fares on the population's willingness to use public transportation. Another example may be the gap between the actual and optimal number of cars on the roads (see figure 2-15). Auxiliaries are not necessarily variables that exist as such but are used in order to define the equations and should be based on data gathered about the system. In addition, some variables are modeled as *constants*. These do not change over time, or are not modeled to change over time, and thus represent the system boundary. In the air pollution example the modelers could decide that for modeling purposes the capacity of the roads would be constant and no new roads would be built. Constants are denoted by a diamond shape (see figure 2-13).

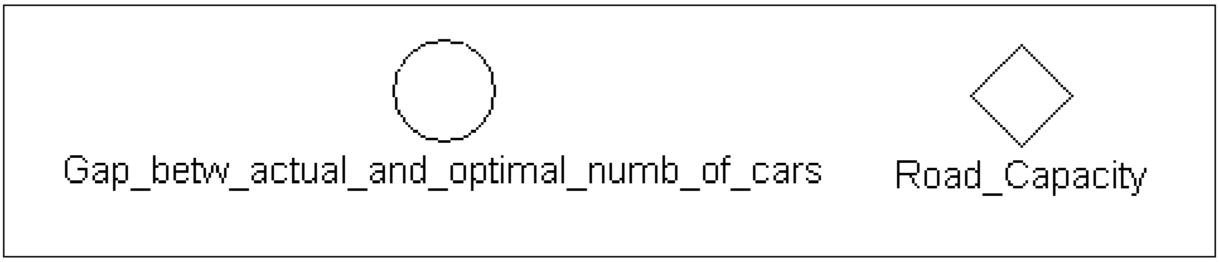


*Figure 2-13: Example of an auxiliary and a constant.*

Figure 2-14 represents the complete stock and flow diagram of the original two-shower model. The tap settings of both showers are modeled as stocks and their net flows are *Change in tap setting 1* and *Change in tap setting 2*. It is not considered good modeling practice to use net flows rather than distinguishing between inflows and outflows. This is because in most cases different variables control the outflows and the inflows such as, for example, the birth and death rate of a population. In this case, however, the inflow and outflow, i.e. turning the tap setting towards cold or hot, are controlled by the same processes.

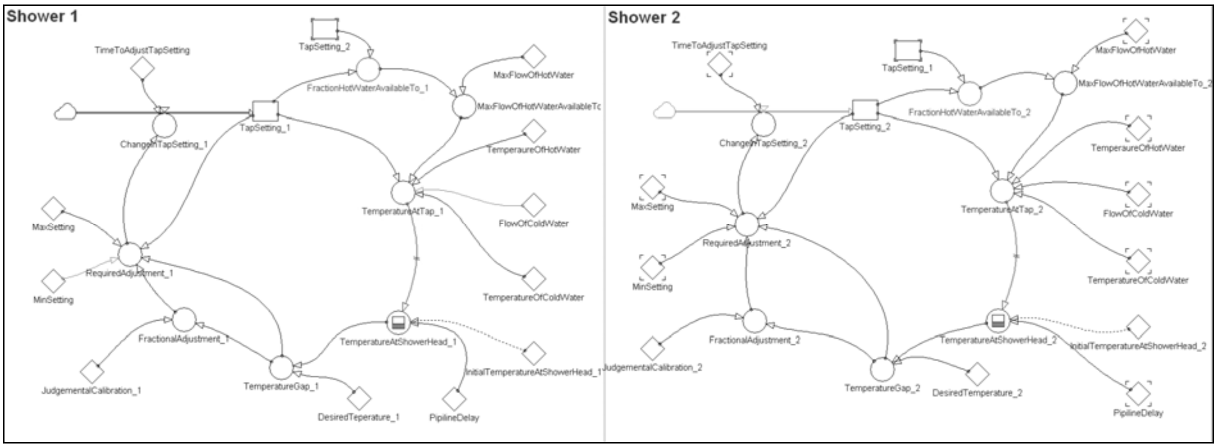


*Figure 2-14: Stock and flow diagram of the original two-shower model.*

An example of an auxiliary in the original two-shower model is *Fraction of hot water available to 1*, which is a variable used to calculate how much hot water *Shower 1* gets compared to *Shower 2*. An example of a constant is *Flow of cold water*. This represents part of the model boundary as the modelers have decided to omit changes in the cold water pressure from the model.

All system dynamic models consist also of a set of model equations. Determining the equations or parameters of the system may often be a difficult and time-consuming process. *"Whatever the parameter type, the system dynamics approach insists that any parameter (and any variable) in a system dynamics model should have a clear correspondence to a real quantity or concept"* (Richardson & Pugh, 1981, p. 232). By this they mean that there should be an *observable* quantity although it in some cases could be *intuitable*. They claim that striving to define the units for the variables can secure realism and observability. According to Richardson & Pugh parameters can be estimated based on three types of information: *"from firsthand knowledge of a process, from data on individual relationships in the model, and from data on the overall system behavior"* (ibid.). Parameters are largely based on data that is gathered about the system or about similar systems, the experience of the modelers, or through participation of other people or experts. The parameters are used to calculate how the values of the variables change over time.

### 2.5.6 Analysis of Model Behavior

Once all variable equations are defined the model can be run. The model behavior is analyzed in order to understand the relationship between the structure and behavior of the model and to get an understanding of which changes may be done in order to improve or solve the problem. *"By exploring the behavior generated by individual feedback loops and by various combinations of loops in the model, the modeler learns about structure and behavior….Simulation experiments isolating and combining the effects of suspected factors can precisely pinpoint the structure responsible"* (Richardson & Pugh, 1981, p. 268). By this they mean that the modelers run the model with and without certain feedback loops to seek to understand the effect that a particular feedback structure has on the behavior of the model and how the relative effect of the feedback structures change as time progresses.

### 2.5.7 Model Evaluation

Model evaluation is performed both at this stage and throughout the development process (Richardson & Pugh, 1981). Several formal tests may be performed during and after the modeling process in order to increase the confidence in the model. The model is changed based on the findings. These formal tests are important as more and more confidence in the model may be accumulated as the tests are passed (Forrester & Senge, 1980). The tests can be classified based on whether they intend to evaluate the structure of the model, the relationships between the structure and behavior in the model, or whether the behavior patterns correspond to data that is gathered about the problem previously (Barlas, 1996).

The simulation results may for example be compared to data that is gathered about the real system. The model may also be tested with respect to how well it responds to extreme initial conditions or input values. The birth rate of a population, for example, must be *zero* when the level of the population is set to *zero*.

The issues of model evaluation and validation are extensively discussed in the system dynamic literature. From a system dynamic perspective it is impossible to declare complete validity of a model. The validity of a model can only be evaluated with respect to how well it serves its purpose (Richardson & Pugh, 1981). The goal of developing a system dynamic model is to provide insight about a *problem* not to perfectly imitate a system. The *problem* is believed to be part of an open system with no set boundary; it is a segment of a vast reality of interconnected variables. Establishing the truth of an open system is unfeasible (Oreskes, Schrader-Frechette, & Belitz, 1994). There will always be unexpected input variables that affect the behavior of the real system. For example, Ruud (2000) developed a system dynamic model of a hydroelectric power plant in Colombia. One of the power stations had to be removed from the model after the real power station was destroyed by the guerilla.[3] If the model previously had been used to predict the future behavior of the

[3] This incident is not referenced in Ruud's article but was told through personal communication with Ruud in 2000.

power plant, the actual behavior of the system would have turned out different from that of the model. Unexpected input variables caused by the political or criminal situation in the country influenced the real system while these variables were not included in the model. Because the problem is part of an open system and no boundaries exist in reality the modeler must decide where to set the model boundaries - what to include and what to omit. In determining the model boundary and selecting which variables to include, the model is affected by the subjectivity of the modeler. System dynamicists therefore do not believe that it is possible to declare complete validity of a model, but that it should be evaluated based on how well it serves its purpose, that is, how well it answers the problem on which the development of the model is based.

Klabbers (2000) question the system dynamic notion of *validity with respect to purpose*. He makes the following comment regarding validation of system dynamic models of social systems:

*"…who are the judges or owners of such a purpose?"* (Klabbers, 2000, p. 384). He further states that:

> *"Because structure of social systems is a social construct, knowledge about it is not a product of merely mirroring reality. Moreover, structure evolves over time on the basis of systems of interactions, generating multiple reality based on various perspectives and interests. So, who can put themselves best in the role of model builders, who can judge the validity of structure? Who are the owners of a particular structure?" (Klabbers, 2000, p. 384).*

A further discussion of this would lead us too far away from the purpose of this chapter so I leave this debate and continue with a representation of the system dynamic method. The next step regards policy analysis and use of the system dynamic model.

### 2.5.8 Policy Analysis, Model Use, and Implementation

After validation different policies are tested on the model. Richardson & Pugh (1981) provide the following explanation of policy analysis:

> *"Model-based policy analyses involve the use of the model to help investigate why particular policies have the effects they do and to identify policies that can be implemented to improve the problematic behavior of the real system. The goal is an understanding of what policies work and why"* (p. 321).

The various policies should be realistic alternative decisions that could be made for the real system. These are tested on the system dynamic model to study how they affect the system and which policies should be implemented into the real system.

Policies in the example of the polluted city could be to double the toll fees on the roads and reduce the public transportation fares by 50 % on days when the weather conditions causes the pollution level to be particularly high. Another example could be to let cars whose registration number ended with an even number drive one day while the cars that ended with uneven numbers drive on others.

As for the original two-shower model one policy could be to always wait four seconds before changing the tap setting or to always ask the person in the other shower (if possible) what their temperature is and what their tap setting is before changing one's own tap setting. The policies that seem best to solve the problem are implemented into the real system.

With this description of the steps of the system dynamic method I move on to some problems both of the visualization methods in system dynamics (the following section).

## 2.6 Limitations with the Traditional System Dynamic Visualization Methods

This section considers some limitations of the traditional visualization methods within system dynamics for audiences without a system dynamic background. As presented in Section 2.5.4 causal loop diagrams (CLDs) are one of the main tools for representing and communicating the structure of a problem-related, complex dynamic system. It portrays how the variables are connected through causal links. The direction and polarity of the links are shown to display whether a variable has a positive/reinforcing or negative/attenuating effect on the other. This type of diagram is utilized to study the causal loops in the system and to obtain an overview of the system structure.

There are, however, some conceptual problems regarding causal loop diagrams (Richardson, 1997). For example, a positive or reinforcing link between two variables, such as an immigration rate and a population may, by a novice, be interpreted that if the immigration rate increases then so does the population immediately. However, an increase in the immigration rate does not lead to an immediate increase in the population; the integration process takes time. Causal loop diagrams do not distinguish between stocks and flows even though the integration of flows into stocks is one of the main characteristics of the system and very different from the instantaneous relationships. A diagram that does not represent stocks, therefore, does not represent the problem in a way that enables us to infer the system's behavior.

There are also some problems with stock and flow diagrams. Stock and flow diagrams represent the complete model structure and when implemented by simulation software the model equations can be accessed through the diagram. Traversing the structure of large models is, however, difficult. One of the main problems with stock and flow diagrams is that there are no embedded visualization techniques that relate model structure to model behavior, except for the ability to display the reference modes as the simulation runs. It is difficult to understand how the model will behave based on a study of its structure. The Quito prototype, which will be discussed further in Chapter 10, for example, is a relatively large model and inferring its behavior by simply studying the stock and flow diagram would be virtually impossible even by the help of simulation.

Studies of graphical interfaces have been performed that utilize other visualization techniques than the traditional ones applied in system dynamics. Howie, Sy, Ford, & Vicente (2000) studied how dynamic decision-making can be improved by making the feedback structures of the system more pronounced through the use of human-computer interface principles. They claim that the design of a graphical user interface can reduce the misperceptions of feedback when trying to understand and control a complex dynamic system. The new interface improved the decisions that were made and increased subjects' knowledge about the system. According to Howie et al. (ibid.) poor performance may be caused by a lack of information, - and not necessarily by a lack of knowledge or a fundamental psychological limitation.

In recent years, interactive learning environments have been utilized with the assumption that the opportunity to present information differently provide an advantage in controlling the amount of information that is given at a time, the details, and controlling the opportunities to make changes in the models.

The following chapters consider how the Two-Shower model, originally developed using system dynamics, may be enhanced and made into prototypes for interactive learning environments. Chapter 3 is a more detailed discussion of enhancing the model by adding narrative elements to its presentation. Chapter 4 is a description of how the model may be enhanced through particular graphic elements, while Chapter 5 is a further discussion of how the enhanced model was developed into a prototype drawing on theories of learning.

# 3. Narrative Design

## 3.1 Introduction

In the last chapter the original two-shower model was presented. In this chapter I consider how the original two-shower model was enhanced through the development of a user interface that included several narrative elements. Chapter 3 is a first run-through of the Two-Shower prototype and the focus is on the narrative elements of the interface. A second run-through, with a more thorough description of the graphical elements of the prototype, is provided in Chapter 4. Chapter 5 builds on both these chapters with a focus on design and learning theories. The issues of narrative design, graphical visualizations, human computer interaction-based (HCI) design, and learning design are somewhat overlapping. The following chapters may therefore contain elements that overlap to a certain degree.

The chapter starts with a discussion of the development process of enhancing the original two-shower model and the requirements that prototype should fulfill. Narrative theory, computer games, and system dynamics are discussed in Section 3.2. Section 3.3 deals with the major design decisions that were made related to the narrative aspects of the Two-Shower prototype, the links between narrative theory and system dynamics, and the more specific design decisions that were implemented in the enhanced model. Section 3.4 is a discussion about the story generator that is implemented as part of the enhanced model. Sections 3.4.1 and 3.4.2 are general descriptions of the intentions of the story generator. Sections 3.4.3-3.4.6 are more specific about how the story generator is implemented and may be too detailed for those who lack experience with complex systems, system dynamics, or programming. The chapter ends with a brief discussion of whether the story generator can actually be defined as a story generator, and if not, what could be done to further develop it.

### 3.1.1 The Process from a Traditional System Dynamics Model to a Prototype of an Interactive Learning Environment

As discussed in the last chapter the goal of the project was to revise the original two-shower model with the intention to create a prototype of a more viable learning environment. As mentioned in Chapter 1, the envisioned user group is novice system dynamic students. These are presumed to be in an age range of 18-30.

For the intended user group the original model may be too complicated and motivational hooks may be added (eg. problem-solving, narrative etc). An advantage of interactive learning environments also is that it is possible to present information gradually without exposing all the details at once. The aim is to reveal information gradually without presenting the participants with an overload of information.

With these starting points as a basis, it was decided that the prototype should include:

1. Elements of narrative
2. Graphics and visualizations
3. Support for different forms of understanding and learning e.g. collaboration and problem-solving

As will be explained in detail in this chapter, elements of narrative were included because the target audience may be acquainted with interpreting narratives through other media such as computer games. The user interface aims at supporting the participants in using their skills of analyzing cause and effect in stories to analyze cause and effect in a complex system. Graphics and visualizations were developed with the purpose of enhancing important aspects of the original two-shower model and present them in a manner that seemed suitable for the audience. Collaboration between participants was emphasized because it may bring about other qualities to the learning experience (Koschmann et al., 1996).

Enhancing the original two-shower model required a number of steps:

1. Programming
2. Production, taking photos etc.

3. Incorporating and debating literature on these three areas (narrative, visualizations, learning).

These steps are interrelated and needed to be considered at each stage and for each solution. Chapters 3 through 5 are divided to follow narrative, visualization, and learning while programming, production and literature are reviewed throughout all three chapters. I begin with narrative because it offers an overview of the major structural changes to the original model. In Chapter 4 I then move to graphics and visualizations as these carry forward our concern with reaching a novice audience and also offer an overview over the major visual changes to the original model. Chapter 5 on design for understanding, is the largest as it deals with numerous aspects of interactive learning environments and of perspectives on learning and knowledge in general.

Before going into a discussion about the narrative elements, I will provide a description of four main requirements that guided the development process of the Two-Shower prototype:

1. As discussed in the beginning of this chapter, the enhanced model should be suitable for a younger, inexperienced audience.

2. The enhanced model should be part of a package for learning about complex systems and system dynamics. In other words, the purpose of the enhanced model was not that it should function as a stand-alone product.

3. The Two-Shower prototype should support collaboration. The enhanced model should communicate the importance of communication and collaboration between the participants. If the participants communicate and agree on which decisions to make it may be possible to find an optimal solution where they all get the most out of the resource.

4. More specifically, certain structural characteristics of the underlying model should be presented in the interface of the prototype. These model characteristics are important clues as to how the model behaves. Although there is an assumption that an understanding of these characteristics is important for a further discussion between the participants about the model and complex systems in general, it may be possible to control the model without having an extensive understanding of the underlying simulation model.

In addition, the development of the Two-Shower prototype had to be guided by three important aspects related to the *structure* of the underlying model. These aspects of the model structure were also discussed in Chapter 2 but are briefly repeated here:

1. There is a *delay* (in this case a pipe that the water must flow through) between a change in temperature at the tap and a change in temperature at the showerhead.

2. The two-shower model is a *feedback* system.

3. The system is *nonlinear*. The implication of this is that the effect of a change in the tap setting depends on the current state of the system, in this case, the current tap settings of both showers. The same tap setting will thus not always result in the same temperature at the showerhead even when the delay is accounted for.

The delay, feedback, and nonlinearity make it difficult to use experience to control the system. The decisions of the other participant, for example, are often hidden until the effects of their decisions become evident. In addition, the same decision does not always lead to the same effect. This should be visualized in the Two-Shower prototype with the purpose of supporting the participants' understanding of the system.

With this brief discussion of some of the starting points for developing the Two-Shower prototype, I now present the major design decisions that were made regarding elements of narrativity in the model. I start with a discussion of narrative theory and computer games and how they may be linked to system dynamics and the study of complex systems.

## 3.2 Narrative Theory, Computer Games, and System Dynamics

This section regards narrative theory, computer games and system dynamics. Characteristics of narratives and computer games are compared to characteristics of system dynamics models. The purpose is to find new ways of representing complex systems.

### 3.2.1 Narrative Theory and Characteristics that May be Used to Enhance a Model of a Complex System

A traditional narrative generally follows a relatively fixed structure consisting of *"…a series of causally related events taking place in a specific time and place"* (Bordwell & Thompson, 1997, p. 480). In an interactive narrative, on the other hand, the chain of events changes as a result of user interaction. According to Meadows (2002) *"an 'Interactive Narrative' is a narrative that allows someone other than the author to affect, choose, or change the plot"* (p. 2). The ability to interact must be provided and the *principles of interactivity* must be integrated with the narrative (ibid.). Meadows lists three principles of interactivity that should be incorporated into the interactive narrative:

1. *Input / Output*
2. *Inside / Outside*
3. *Open / Closed*

*(ibid., p. 39).*

By *input / output* he means that input should result in an output and a user should be able to respond to the output by providing new input. The response time between input and output should be short enough for the user to understand that there is a relationship between output and input. As for the Two-Shower prototype this could refer to how the participant may be able to change the tap setting and receive timely feedback from the system on the effect of the input. The participant should then be able to respond to the feedback from the system by making additional changes to the tap setting.

Meadows distinguishes between *inside* and *outside* interactivity where inside interactivity is what goes on in the mind, or *"the world of the reader's imagination"* (ibid., p. 40). The outside is the more tangible, physical characteristics of the system, such as the frame rate or the design. He claims that there should be a balance between the focus on the *inside-the-skull* and *outside-the-skull* information (ibid.). In the Two-Shower prototype *outside* could refer to the system architecture that is used while the intended *inside* may refer to the atmosphere that we have tried to create through the design with the purpose of triggering and maintaining the participants' attention.

By *open* versus *closed* systems Meadows means that a closed system is predictable while an open system is unpredictable. He claims that closed systems are *boring* while open systems are *interesting*. As will be discussed in Chapter 5 this may not be quite compatible with systems that are intended for learning. Learning to predict how the system will react to input is part of the challenge. If the Two-Shower prototype, for example, would react differently to input every time it would be impossible to predict how it would respond to input and more difficult to understand the underlying simulation model. Trying to obtain control of the model would maybe also not be interesting if it was impossible to know how.

Meadows (2002) claims that interactivity leads to a blurring of the roles between the writer and the reader. Both the reader and the writer provide information to the story. In interactive narratives the role of the author changes compared to traditional narratives presented as for example written text with a fixed plot. The author becomes a facilitator that should aid the participant in making her own narratives. The way one narrates is thus changed. In interactive narratives the role of the reader is also changed. Instead of being a relatively passive spectator the user must herself participate in the construction and presentation of the narrative. Zapp (2002) presents a similar view and claims that *interactive narration* at first sight seems to be a contradiction. *Interaction* indicates that a user or reader should interact with the system while the traditional and narrow interpretation of narrative is related to storytelling.

### 3.2.2 Narration and Complex Systems

A focus on causal relationships that occur over time is central within both narrative theory and studies of complex systems (see for example Bordwell & Thompson, 1997; Forrester, 1961). The behavior of a simulation model of a complex system can be seen as the course of a story that is driven by causal relationships (the structure) that evolve over time and that are triggered by the decisions of a participant.

Edgar Allan Poe was interested in presenting narratives as what cause produced what effect (Meadows, 2002). In mystery stories, for example, the gradual solution of the problem served as the story itself. The readers take part in an investigation process. According to Meadows (ibid.) *"Poe turned the reader into an investigator"* (p. 23) and this can be seen as a first step towards connecting narratives and interactivity.

In a similar manner, the behavior patterns of models of complex systems may be presented as narratives where a participant should seek to find the causes that lead to the behavior. Narrative theory may be used to determine what to present, how, and in what order, with the purpose of presenting a complete and interesting story. Theories of how the participants' interest should be kept throughout the story may be based on narrative theory. Just as the information that is given in traditional narratives is given with the purpose of achieving special effects on the viewer, a complex systems story may present particular information, in a particular manner, at a particular time, in order to emphasize important issues concerning the dynamic story. The information that is provided should focus on important issues for understanding the relationship between structure and behavior while at the same time letting the participant be an *investigator* of what happened in the model.

In some cases, interactive narratives may involve that the participant purposely determines the structure of the story. In case of narratives based on the behavior of a simulation model, the plot of the story is determined by user actions that influence the model behavior and the responses from the participant to the behavior of the model. The participant may not be aware that the story is being constructed or may not know how her actions influence the story. The goal, however, is that the participant should learn how the system responds and thus be able to predict what the story will be based on her input.

A complex systems narrative may be presented as a story where the effect is known as an undesirable state of the system; however, the user does not know how and why this state occurred (the causes that lead to the effect). The challenge is to develop a story that explains how the effect occurred. This must be done by explaining the important parts of the system structure that lead to the undesirable state, the behavior that should have been monitored in order to avoid the undesirable state, as well as the relationship between the structure and the behavior that made the state undesirable. There should be a focus on previous decisions and how these have affected the undesirable behavior.

### 3.2.3 Previous Work on Presenting System Dynamic Models as Stories

Previous work on presenting system dynamic models as stories has been performed by Mojtahedzadeh & Andersen (2001) and Mojtahedzadeh, Andersen, & Richardson (2003). They have developed some software called *Digest* which is based on the Pathway Participation Metric (PPM). PPM is a mathematical method related to the eigenvalue analysis, which identifies the dominating structure of a complex system. Their work, however, is mainly concerned with technical and mathematical issues, not as much with the presentation and visualization of the results. Their intentions of how users should interact and learn from their software are not discussed. The presentation of the system stories do not seem to be deeply grounded in theoretical approaches concerned with for example presentation, perception, or interactivity.

The intention is for the software to aid users in identifying important links and variables in the system. The presentation of the stories is based on the causal loop diagram, which as discussed in Chapter 2, contains simplifications that may mislead the user. These diagrams may also not be suitable for an inexperienced audience.

Their target audience is, however, different from the young, inexperienced audience of the Two-Shower prototype. Their system is rather meant *"...to help managers better understand the systems which they manage and in which they live"* (Mojtahedzadeh et al., 2003, p. 1). Managers may have more experience with the system of which they are part than people who are going to learn with a simulation model of a system they may not have given much thought previously.

### 3.2.4 Computer Games and Characteristics that May be Used for Enhancing a Model of a Complex System

In addition to incorporating elements of narrativity in the Two-Shower prototype, we have borrowed some elements of a related field, *computer game design*. Computer games also often contain narrativity and narrative characteristics from computer games may be used to enhance a model of a complex system. It is, however, important to stress that a computer game is not a narrative, but may only contain narrative elements (Juul, 1998). Juul states that *"...in many cases the player may play to reach a narrative sequence"* (p. 9). By this he means that the actions or input made by the user and their consequences may unfold as a *narrative sequence*. Similarly, a simulation run of a complex system, although consisting of a series of causes and effects, is not a narrative, but its behavior may be presented as a narrative sequence.

A good or popular game has the advantage that it generally manages to keep the attention and interest of the user throughout a session. Juul (2003) presents three principles of what makes a computer game interesting:

1. *No single option can be consistently superior to the others.*
2. *The options should not be equally attractive.*
3. *The player must be able to make an informed choice.*

*(p. 1).*

If the first principle is not followed and one alternative is consistently better than another the predictability of the game could cause the user to lose interest. In the game *SimCity*[4], for example, if better results were constantly gained through lowering the income taxes rather than raising them the game would lose part of what makes it challenging. Applying this principle to a model of a complex system, the consequence of the nonlinearity of such a model is that it is not possible for one specific decision or action to constantly be better than another. In the Two-Shower prototype, for example, it is not consequently better to increase the tap setting.

If the second principle is not followed and equal results are obtained through alternative actions, a game will lose part of what makes it a game. The decisions of the user will be of no consequence because she will obtain the same results anyway. In the SimCit*y* example, if both raising taxes and reducing taxes provide equal results the game will lose part of what makes it exciting. Nothing will be at stake for the user. The user does not have to consider which decisions to make because it will not matter. In a model of a complex system it is usually not possible to obtain equal results from alternative actions because the variables in the system are connected through nonlinear causal relations and a change in one variable would affect the system differently from a change in another variable.

If the third principle of informed choices is not taken into account in the development of a computer game, it will not be possible to obtain or improve skills through playing the game and that may make the user lose interest relatively quickly. In SimCity, for example, there is information about the conditions in the city in the user interface. Hints about variables with critical values and access to graphs of previous behavior are provided. This forms a body of information which the user may try to interpret and use in her decisions. If this information is not available, there will be no challenge to the game or it will be so challenging that the user may lose interest. It is also possible that when the user has enough information and experience with the game it will no longer be interesting to continue playing the game. The principle of

[4] For those unfamiliar with *SimCity,* it is a simulation game, developed by Maxis, where the player builds a city and takes the role of the mayor who controls the city through decisions regarding taxes, road construction, land allocation etc. (http://simcity3000unlimited.ea.com/us/guide/ (accessed May 12, 2006).

informed choices may also apply to the Two-Shower prototype. Once the participants know how to obtain a stable, comfortable temperature there may seem to be no challenge left of controlling the model.

Giving participants the ability to make informed choices is an important part of the development of complex systems interactive learning environments. When dealing with complex systems it is difficult to know and predict how the system will respond to a decision and the decisions must always be evaluated based on its believed consequences. These decisions can only be made based on information about the system. The objective of, for example, a system dynamicist is to provide the decision maker with this information.

Because of this commonality between computer game theory and complex systems elements of computer games may be utilized in the development of interactive learning environments that are developed with the purpose of supporting the understanding of complex systems. The following is a description of the design decisions regarding the general elements of narrativity in the user interface of the Two-Shower prototype along with the motivation for these design decisions. Some details of the user interface are not explained fully until Chapter 4 on particular graphic elements of the user interface and Chapter 5 on design for understanding. The screenshots in the next section therefore only represents the part of the user interface that is directly related to the narrative characteristics of the model. The prototype contains additional characteristics that will be described in the next two chapters.

## 3.3 Major Design Decisions – Narrative

In this section I consider the major design decisions in the Two-Shower prototype which draw on elements from interactive narratives and computer games. Narrativity is incorporated into the Two-Shower prototype on two levels: there is a general focus on narrative in the user interface and a *story generator* is created in order to let the participant go back and analyze what happened during the simulation run. This section is divided into a discussion about the setting (Section 3.3.1) and characterization (Section 3.3.2). Following the major design decisions related to narrative is Section 3.4, which is a discussion of the story generator that is developed as part of the prototype. All the programming of the interface as described in Section 3.3 was performed by myself, with the exception of the inclusion of the graphics of the hotel (see figure 3-1)[5]. As mentioned in Chapter 1 Kristian Kastet developed a first version of the program that controlled the communication between this interface and Powersim SDK. I did all the work on the interface while Kastet worked on the communication between Flash and Powersim. Section 3.4 is primarily based on the work of Frotjold (2005).

### 3.3.1 The Setting

One of the first major design decisions that were made regarding the Two-Shower prototype was that the setting should be a hotel where the participants control a character representing a guest who is taking a shower. Embedded in this decision is the issue of narrativity and that the enhanced model should contain certain narrative elements.

Based on the idea of the hotel setting it was decided that the participant should first be presented with a simplified graphic representation of a hotel and the two showers (see figure 3-1). A silhouette of a hotel with a pink neon sign is intended to create the atmosphere of a simple hotel. Photos of two people who are each located in a shower and a graphical representation of their connection through the hot water tank is included. It may not be evident for the participants what is going on at this stage, but part of the intention is to trigger their curiosity. The hotel setting is partly inspired by the famous shower scene in Alfred Hitchcock's movie *Psycho* from 1960.

[5] This was, however, part of the initial plans for the prototype presented in (Skartveit et al., 2003).

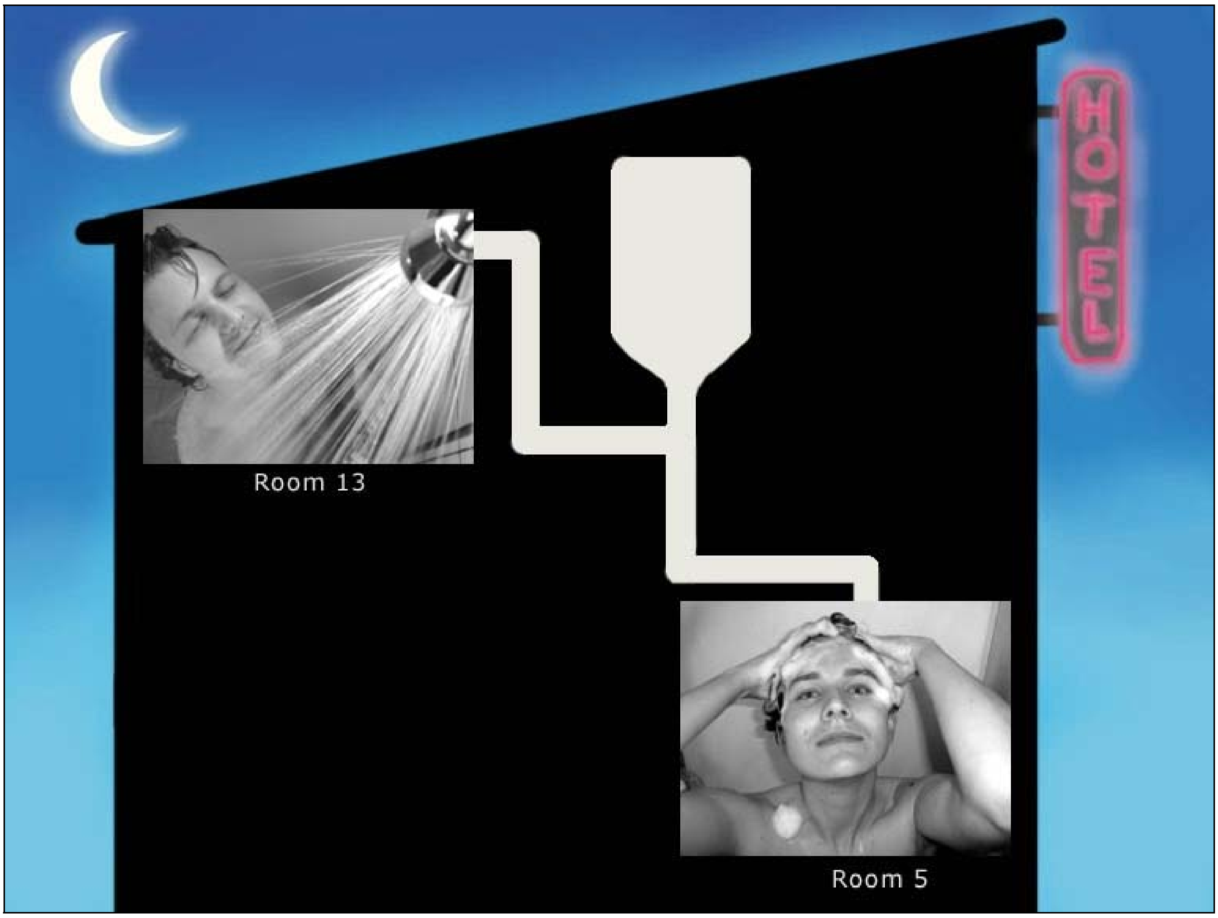


*Figure 3-1: The introduction to the hotel setting.*

The participant is then presented with an introductory story which forms the foundation for what is going to happen (see figure 3-2).

*Imagine the following scenario: You arrive at a hotel for the night, and tired from an exhausting journey you long to take a refreshing shower. The bathroom looks OK and the shower as well - until you turn on the water... You struggle to find the right temperature, turning the tap more and more desperately towards hot or cold with not immediate results. Finally, you realize that a little patience waiting for the temperature to change before you turn on even more hot or cold water helps more than cursing. Then you hear the shower in the next room being turned on, and it is the same story all over again. What happened? Remember that you can press the pause-button at any time during the simulation and receive comments about what has happened so far!*

*Figure 3-2: The introductory story of the Two-Shower prototype.*

Here the participants get additional hints as to what is going on and they are introduced to the fact that there is another shower. The intention of the introductory story is to trigger the participants' curiosity by enabling them to create an anticipation of a setting. It also frames a setting that may be familiar to many, as many people may know what it is like to be tired from traveling and the problems of obtaining a comfortable shower temperature in an unfamiliar shower.

After having read the introductory story the participants may enter the Shower Room view where they can control the model, *take a shower*, and try to obtain a comfortable temperature. Figure 3-3 is a representation of the part of the user interface that is particular for the *Shower Room view*.

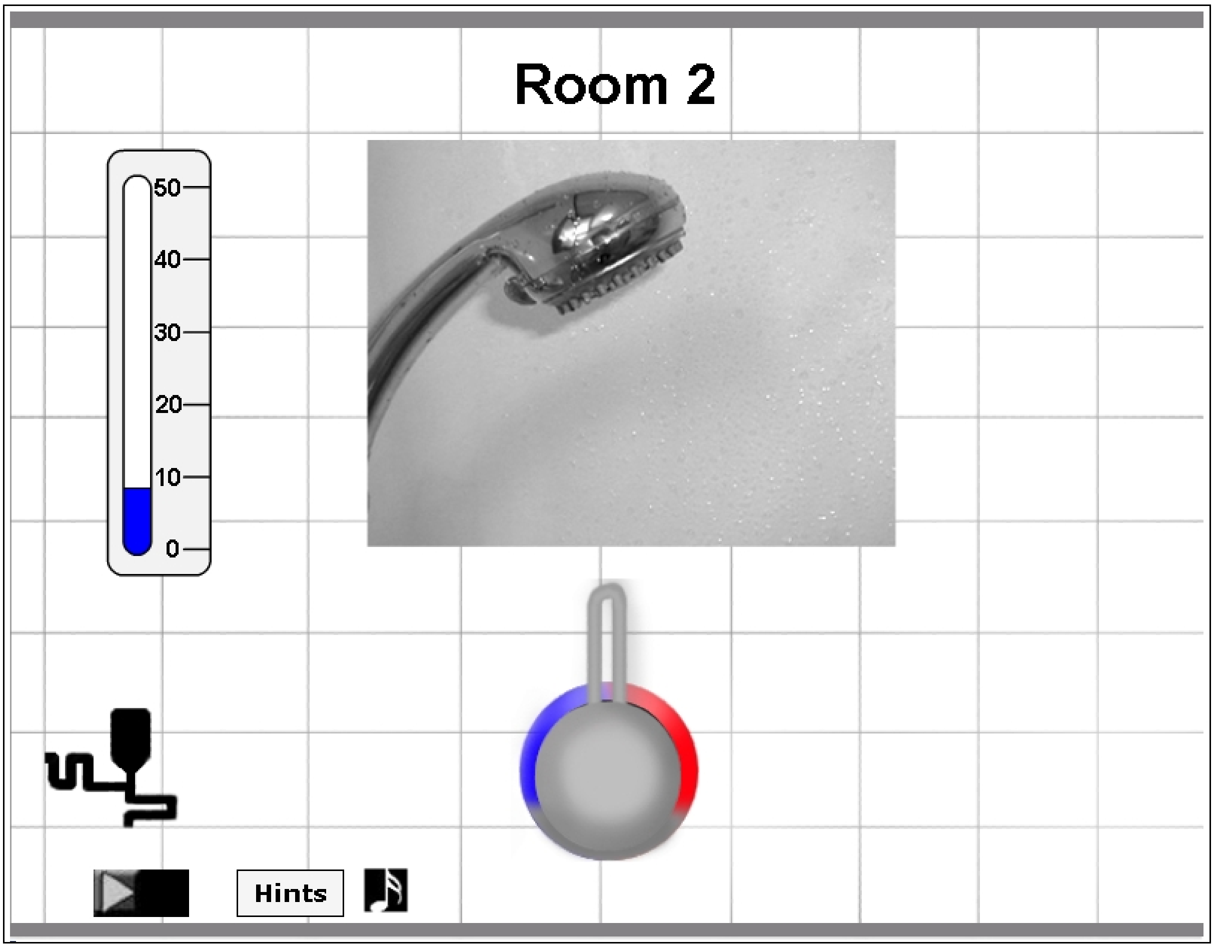


*Figure 3-3: Graphics portraying the Shower Room.*

The Shower Room view has the atmosphere of a bathroom with tiles on the walls, a mixing battery to control the tap setting, and a showerhead.

The room number is displayed on the top of the page and the two showers have room numbers *one* and *two*. One of the original ideas was to include a Phone button where the participant could call a receptionist and the receptionist would provide an explanation of why the temperature of the water would tend to oscillate, but this was not implemented in the final version of the prototype.

Another initial idea was to include a Keyhole button in the lower right-hand corner to let the participants *peep* through the keyhole to see the conditions of the other participant by clicking the button and opening a small window. This was implemented in the earlier versions of the prototype but omitted from the final version. Part of the intention was to trigger the participants' curiosity and encourage them to wonder about what was happening in the other room.

The participant can click the *Pipe System button*, which is shaped like a hot water tank with connected pipes, and enter the *Pipe System view* where a graphical explanation of the pipe system is given. Figure 3-4 is a screenshot of the Pipe System view.

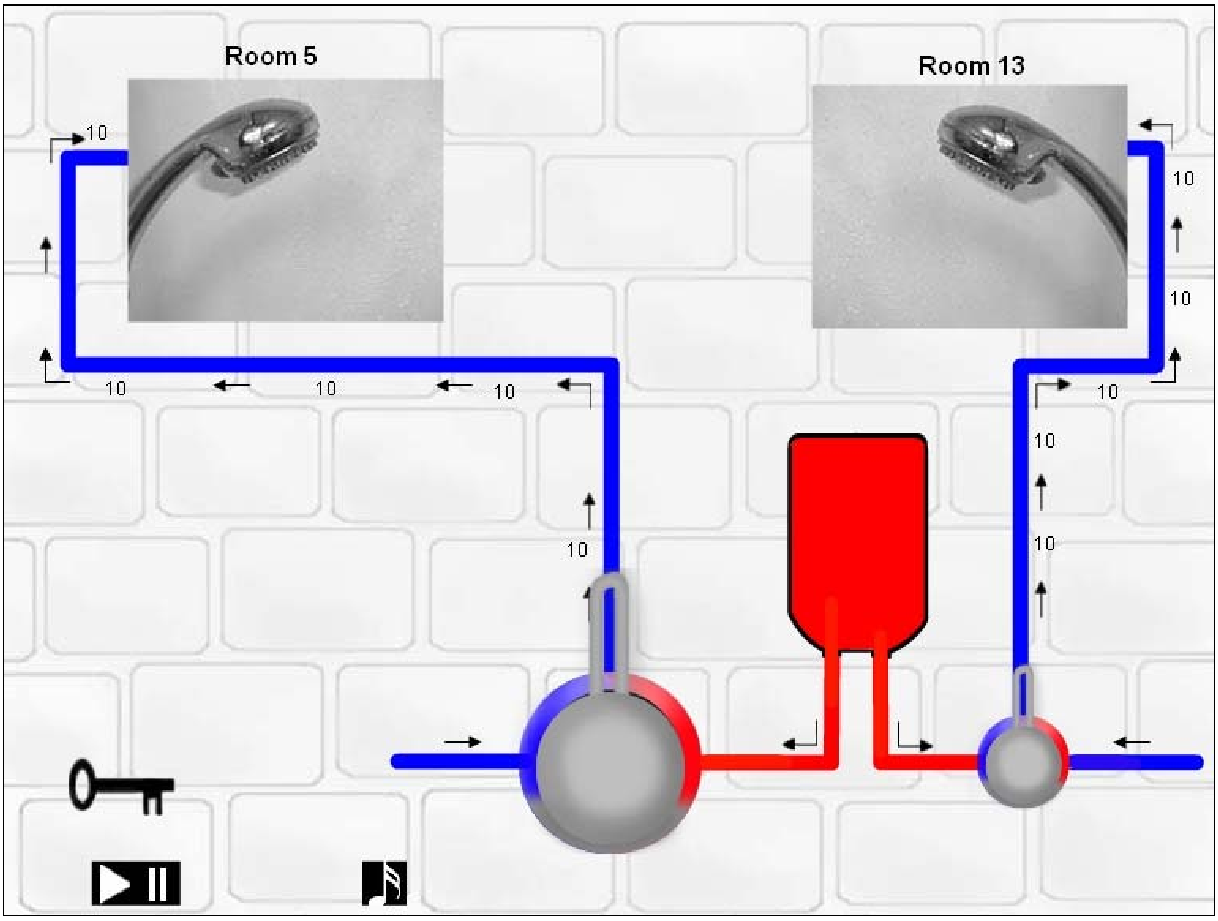


*Figure 3-4: Graphics portraying the Pipe System.*

The Pipe System view represents the inside of the walls and the participants may study the structure of the pipes and the hot water tank. A graphical representation of the hot water tank is displayed with pipes connecting it to the showerheads via the mixing batteries of the two showers. What makes this view different from the Shower Room view is that the part of the model that is controlled by the other participant is visible. A miniature version of the mixing battery of the other participant is displayed in the lower right hand corner. Changes in the tap setting, the temperature of the water in the pipes, and the temperature of the other participant's showerhead are also displayed. The participants are here presented with the *full story* of the system structure, with the exception of the details of the underlying equations.

Some of the main intentions behind the Shower Room and Pipe System views are to create a narrative setting to trigger and maintain the attention and curiosity of the participants. There are also some underlying intentions of how it should support understanding. These are discussed in detail in Chapter 5.

### 3.3.2 Characterization

Along with the decision to include elements of narrativity and the creation of a bathroom setting comes the issue of characters and who are the stakeholders in the system. In the Two-Shower prototype photo slideshows of changing facial expressions of the characters are displayed based on the temperature at the showerhead of each shower during the simulation run. The intention of using photos is to provide a means for the participants to become more involved in finding out how to give the characters a comfortable, stable water temperature.

Characters are important in both computer games and narratives. Most computer games progress through the players' control of characters. In the game the *Sims*[6], for example, the player creates the characters, determines his or her main characteristics, such as gender, appearance, main preferences, and attitudes. The game then progresses as the player controls the characters in their *everyday lives*.

Characters are also important in more traditional narratives. According to Meadows (2002) *"the human element of interpretation needs to be present for writing to become narration"* (p. 29). All narratives need an opinion or perspective. A simulation model with cause and effect relationships is thus not a narrative in itself. A perspective must be added. This may be obtained through characterization and in the Two-Shower prototype this is approached by letting the participants control one of the characters in the showers. Characters are important for personalizing a story and

[6] For those unfamiliar with the *Sims*, it is a simulation game developed by Maxis, where the player creates and controls characters, build homes for their characters, and form relationships with other characters in the game (see http://thesims.ea.com, accessed September 24, 2006).

creating a basis on which a user can identify with. According to Meadows (2002), *"characters, be they protagonists, antagonists, or narrators, offer perspective, deliver opinion, provide interpretation, and generate a kind of emotional foundation the story is built upon"* (p. 28). Characters bring the narrative life and allow users to identify with the story. Meadows states that *"a character that is present in an environment, someone who cares about something, someone who has some form of opinion, perspective, or passion, is something that gives a narrative life"* (ibid.). In the development of the user interface of the Two-Shower prototype we found the addition of perspective by including characters adequate. The intention was to enable the participants to identify with the perspective of the characters in the system.

Characters are the agents who react to events and often trigger the cause and effect relationships (Bordwell & Thompson, 1997). In the case of the Two-Shower prototype it is the character in the shower (which is controlled by a participant) who *reacts* to the temperature of the water by changing the tap setting. The change in tap setting triggers a chain of cause and effect relationships, which again changes the water temperature for the character in the shower.

As described previously, the interface of the Two-Shower prototype is based on the concept of a hotel where the participant is a guest who wants to take a shower. The participant is first introduced to the story by being told that she has just arrived at the hotel after an exhausting journey and long to take a refreshing shower. The introduction is presented in this way in order to let the participant control a character with a role in the session and in this manner provides the ability to identify with the problem.

*Figure 3-5: The characters of the Two-Shower prototype.*

A character is someone that cares. In the Two-Shower prototype the participant can see the facial expressions of the character in the shower. This is done with the intention of letting the participant sympathize with the character they are controlling.

The use of photos with changing facial expressions is intended to make the participant feel some sort of attachment or a feeling of responsibility regarding the situation of the characters. An intention is that photos of real persons rather than graphical animations create a stronger expression (see figure 3-5). Part of the goal is to make the participants feel empathy towards the characters when they are scalded or freezing.

I have now introduced some of the general narrative elements in the Two-Shower prototype and move on to a more specific narrative feature of the model, the *story generator*.

## 3.4 The Story Generator

As discussed in Chapter 2, people often have problems understanding how behavior occurs as a result of the underlying structure of a complex system. Our assumption was that if the participants could *pause* the simulation and *go back in time* to study how their input affected the behavior this could aid them in developing a better understanding of the system. A *story generator* that explains what has happened in the system during the simulation run through text and graphics was therefore developed. The decision to create a story generator was part of the initial discussions of the project group. The actual architecture and programming was carried out by Master student Laila Frotjold.

The background for the story generator and its main intentions are described in Section 3.4.1. Section 3.4.2 continues the discussion of intentions with a focus on the user interface of the story generator. Sections 3.4.3-3.4.6 describe the system architecture and the components of the story generator. The technical solutions at a highly detailed level, however, are not dealt with. For a more detailed description of the technical issues see Frotjold (2005). Section 3.4.7 is a discussion of whether this is really a *story generator* and whether the generated *stories* fulfill the requirements for being defined as *stories* or *narratives*.

### 3.4.1 Background and Main Intentions

The idea of the story generator was part of the early intentions and discussions in the VOCS project group (Skartveit et al., 2002) and it was decided to have a student do this as part of a Master's thesis. As mentioned in Chapter 1, Laila Frotjold became part of the project in 2003 and developed a text-based story generator . The description of the text generator described in this chapter is therefore partly based on her Master's thesis (see Frotjold, 2005). The story generator was further developed by external programmer Olve Sæther Hansen and myself in 2006. In this new version of the story generator graphics and animations were included. This means that when the participant activates the story generator graphics or animations will appear in addition to the text (see 3.4.2). The story generator was reengineered as an agent that gives the participants an indication of important time steps where important events have occurred. Some network issues were also solved.

The story generator is embedded as part of the Two-Shower prototype and is integrated into the user interface. The participant may press a pause button that triggers an agent that constructs a narrative sequence, presented as a series causes and effects. The participants may place their mouse cursor over the timeline, go back, and examine points in time where important events occurred. In this manner they can investigate the causes of undesirable behavior. Textual information about changes in the model are provided at the same time as the graphics and animation change according to the state of the model at specific time steps. The intention of the story generator is to provide the ability to go back in time in the simulation, look at previous actions and their consequences, and preferably learn from the mistakes that were made.

It may often be difficult to obtain an overview and to discuss what is happening during a simulation run because one is busy trying to control the model. The assumption is that the opportunity to pause the model will give the participants some time to discuss how the underlying model causes the behavior and which actions may be taken in order to control the model.

The narrative sequences are generated from information that is gathered about the model structure, the model behavior, and input from the participants. The story generator is composed of a *model analyzing component* that analyzes the structure of the underlying simulation model in order to find critical characteristics that are important for the understanding of the model (Frotjold, 2005). It further consists of an *explanation generator component* where textual explanations are generated along with multimedia visualizations of the conditions in the system at given time steps.

The participant constructs the *plot* of the narrative sequence or *story* as the simulation runs. The story of the impact of the decisions may be accessed afterwards. The Two-Shower prototype has an open plot structure in that the participant influences the plot through the decisions that are made. Decisions to change the tap setting can be made throughout the simulation. However, the plot is not completely open in that there is an underlying structure of the simulation model that responds to the changes in tap setting made by the participants.

### 3.4.2 The Story Generator Interface

Indications of the existence of the story generator are given early in each simulation session. The participants are told in the introductory story (figure 3-2, p. 88) that they may pause the simulation at any time to receive an explanation about what has happened. There is a text box in the user interface where the text *"If you think the shower temperature behaves strangely you can pause the system and receive comments about what has happened so far"* is displayed at all times when the simulation is running. If an enduring too cold or too hot temperature occurs, that is, if the participant has not been able to stabilize her shower after 20 seconds, a text box with a yellow *warning pop-up* appears (see figure 3-6).[7] It suggests that the participants pause the simulation and get an explanation of what has happened. The text says *"Oh, too hot! Remember that you may pause the simulation and get an explanation"* or *"Oh, too cold! Remember that you may pause the simulation and get an explanation"*. The pop-up disappears when the participant is able to obtain a comfortable temperature.

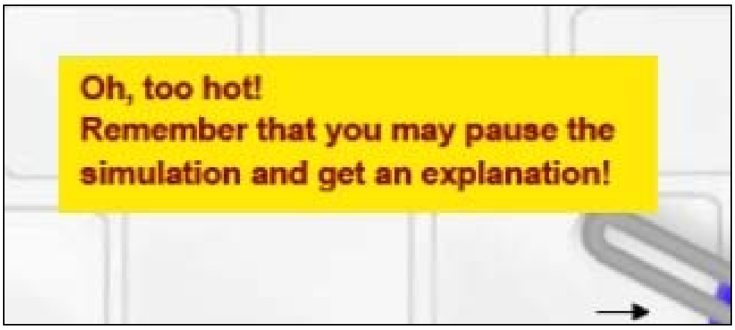


*Figure 3-6: The yellow warning pop-up that reminds the participants that they may pause the simulation and get an explanation of what has happened if an enduring too hot or too cold temperature occurs.*

Another element in the user interface related to the story generator is the *timeline* that shows how many seconds the simulation has run. When the participant presses the pause button, an agent is triggered. The agent traverses all *events* that have happened during the simulation run and places several yellow markers on the time line to indicate when the important events occurred and which time steps it may be useful to

[7] The warning pop-up was discussed as part of the initial plans for the prototype and was implemented by Frotjold (2005).

study more closely[8] (see figure 3-7). In this case the monitored events are changes in tap settings and temperatures at the showerheads of both showers. The participant can then place the mouse cursor over the yellow markers on the timeline. The text in the textbox refers to changes in a variable at specific points in time, the direction of the changes, and the cause of the change.[9] An example of such a text may be:

> *"An increase in the temperature of your showerhead occurs at this time, a while after the cause for the increase happened. In this case the cause was a reduction in the other user's tap setting"* (see figure 3-7).

In addition, the graphics, animations, and slideshows show the conditions at the current time step.[10] As the participant places the mouse cursor over the various yellow markers on the time line, the graphics and animations change according to the conditions that were at the particular point in time. The participant can thus go back and forth on the timeline and study the *story* of what happened during the simulation run.

[8] The implementation of this feature as an agent and the yellow markers on the timeline was developed by Olve Sæther Hansen..

[9] The text generator was, as mentioned previously, designed and implemented by Frotjold.

[10] The changes in graphics and animations for the story generator was implemented by Olve Sæther Hansen and myself.

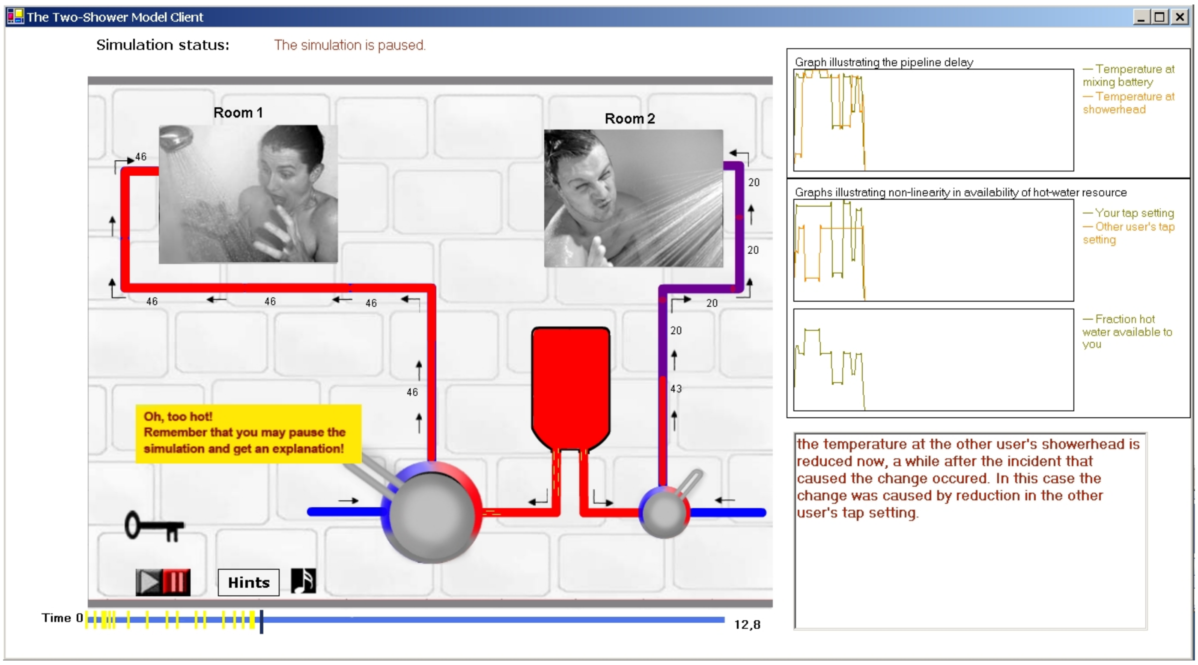


*Figure 3-7: The story generator. When the model is paused the yellow markers on the timeline in the lower part of the interface appear. When the mouse cursor is placed over the yellow markers the graphics change according to the state of the system at the particular point in time and a textual explanation of the change that occurred in the model is displayed in the text box in the lower right hand corner.*

## 3.4.3 System Architecture

Figure 3-8 represents an overview of the system architecture used for the Two-Shower prototype. When the participant starts the model and provides input through the graphical user interface an agent is triggered. The agent sends user input data and retrieves simulation data from the server. When the participant pauses the model, the agent sends a pause signal to the server, retrieves textual explanations, update the graphics, and show the markers on the timeline. Based on data from the client, the server proceeds or pauses the simulation model, triggers the model analyzing component and the text generator.

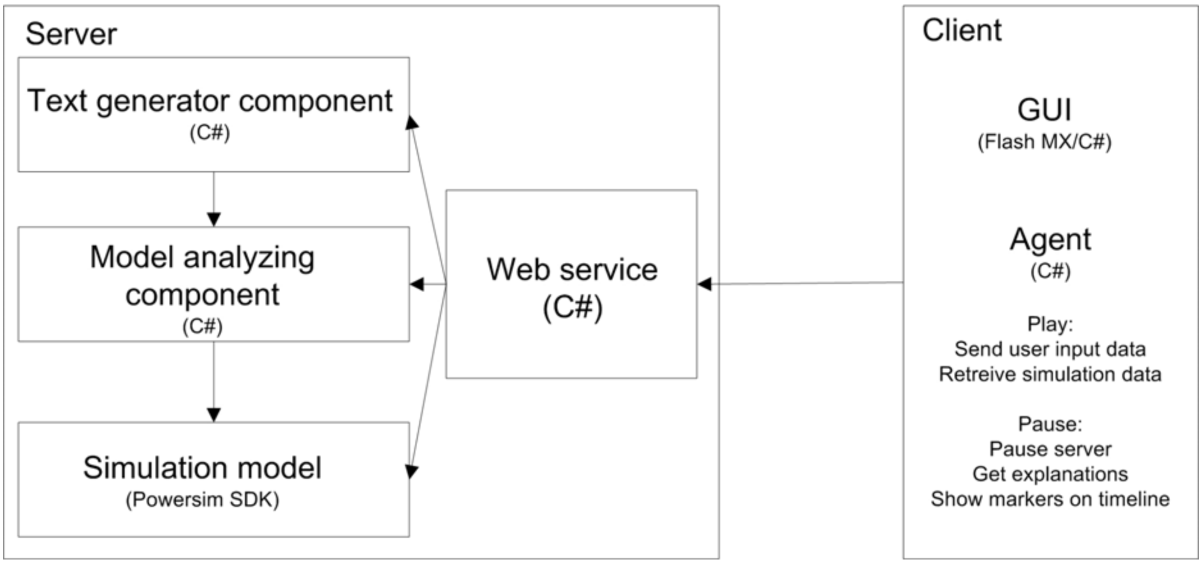


*Figure 3-8: The system architecture of the Two-Shower prototype with the story generator.*

### 3.4.4 The Model Analyzing Component[11]

The story generator has two main components: the *model analyzing component* which is used to traverse the model and find important characteristics about the structure and an *explanation generator component* where the textual explanations of the important characteristics and events in the model are created (Frotjold, 2005). The explanation generator uses the results from the analysis performed by the model analyzing component to create explanatory texts of what has happened during a simulation run (see figure 3-9).

[11] The following subsection is based on Frotjold (2005).

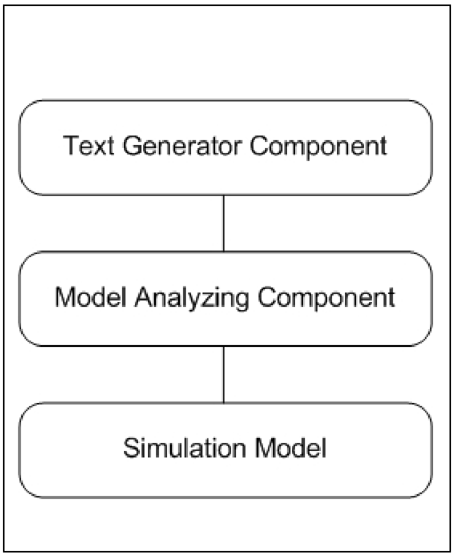


*Figure 3-9: The components of the text generator.*

As explained in Chapter 2 all system dynamics models consist of a set of equations. The equations are connected through a causal structure where the value of each variable is defined as a function of other variables and constants. An important task for the story generator is to find important characteristics of the model structure on which the explanations can be based. In order to do this a component that performs an analysis of the structure of the underlying simulation model was made (ibid.). As discussed previously part of what makes a complex system difficult to understand and control are delays, feedback loops, and nonlinearity. The model analyzing component loads all the equations of the underlying simulation model at the start of each session, traverse the structure, and finds interesting paths in the model used to create readable and intendedly more understandable explanations. It retrieves the equations from Powersim and uses them to find feedback loops, nonlinearities, and delays in the model.

The *loop-finding unit* starts at a stock variable and traverses through the model structure until it returns to the initial stock variable (ibid.). The path of connected variables is defined and stored as a feedback loop. In the original two-shower model there were two stock variables, the tap settings of Shower 1 and Shower 2. These were used as starting points for the loop-finding unit.

A *delay-finding unit* was created in order to identify the delays in the model structure (ibid.). Delays are defined as a set of functions in Powersim and the equations can thus be checked to identify the delays. Stock variables also represent delays (it takes time to accumulate and reduce the stocks) and their delay time can be found by analyzing their inflows and outflows and their linked variables.

The *nonlinearity-finding unit* is more complex than the previous units because nonlinearity can take various forms (ibid.). The equations are divided into components and checked for variables, constants (linked variables or constants that were used to calculate the value of the current variable), and plus, minus, division, multiplication signs, and potentials. All these characteristics are important for determining the nonlinearity of the equation components. The components are then analyzed in order to find out whether they are nonlinear and the variables are analyzed to register whether there are nonlinear relationships between them.

A problem was encountered in the analysis of the model because the actual simulation model that is used in the Two-Shower prototype is modified compared to the original model presented by Morecroft, Larsen, Lomi, & Ginsberg (1995). As previously mentioned, the original model contains two *simulated* people whose decisions to make changes in the tap setting are simulated based on the temperature at the showerhead, the temperature gap, simulated judgments of the change that is needed, and the time it takes to change the tap setting. In the Two-Shower prototype the *participants* take the role of the simulated people and decisions to change the tap setting are made outside the simulation model. In a way, one may say that the participant becomes part of the feedback loops in the model, but the simulation model itself does no longer contain any loops. The problem of how to analyze the model, when it no longer contained any feedback loops, was therefore encountered (ibid.). It was solved by copying some of the structural elements from the original two-shower model and using these in the structural analysis of the model. The analysis of the behavior, however, was performed on the actual underlying simulation model during the simulation run.

### 3.4.5 The Text Generator – The Explanation Generator Component[12]

The information that is gathered in the analyzing component is turned into readable text by the *explanation generator component*. The explanation generator is based on the works of White & Caldwell (1998). White & Caldwell describe a similar text generation tool where a hierarchy of description classes is used. A description hierarchy is a taxonomy of different description elements correlating to model elements and characteristics such as delays, nonlinearity, or variables which are controlled by participants.

Through the description class hierarchy it is possible to let the subclasses handle different states that are to be described to the participant. Different characteristics of the underlying simulation model, such as nonlinearities and delays are also described by different subclasses in the description class hierarchy.

The text generator described in Frotjold (2005) generated both a *basic explanation* that was shown as an introduction to the system and *continuous explanations* that were generated when the participant pressed the pause button. The basic explanation was generated before running the model and was therefore only a description of the structural characteristics of the model. The generation of the basic explanation was used as an example of how a textual explanation of the structure of a model can be generated. The generation of the basic explanation is not used in the final version of the Two-Shower prototype as this made something that was relatively basic more complex but was used by Frotjold as an example of how this could be done for larger models. Generated basic explanations may be more useful, however, for larger models. The next section is a further discussion of the continuous explanations that are generated as part of the story generator.

[12] The following subsection is based on Frotjold (2005)

### 3.4.6 Designing the Content of the Generated Explanations[13]

Through the continuous story generator the participant is presented with information about specific changes in variables at specific points in time during their simulation run and the cause that triggered the change. The continuous explanations thus focus on causes and effects in and between variables. The generated text also states the resulting direction that the change in one variable has on another, that is, whether the cause was an increase or decrease in a variable, and whether it resulted in an increase or decrease in the affected variable. In addition, it refers to changes that will occur in an affected variable after a delay (forward directional explanation) and to changes that has caused a change at a earlier point in time (backward directional explanation). The story generator in this way links the structure and the behavior of the underlying model of the complex system. Because the story generator indicates how a change in one variable affects another it would be able to intercept a shift in loop dominance, this is, however not studied here.

The following text is generated by the story generator if, for example, there is a change in the temperature at the showerhead:

> *"An increase in the temperature of your showerhead occurs at* this time, a while after the cause of *the increase happened.* **In this case** the cause was *a reduction* in *the other user's tap setting."*

This text refers to the delay that there is in the system from a change in the tap setting of one of the showers is made until the water with the new temperature reaches the showerhead.

The words in bold are taken from a list of predefined, semantically identical phrases. These phrases vary in order to make the language more dynamic and less monotonous. These words may be replaced by, for example, *additionally*, *but this is not all*, *further*, *as well as*, etc. The words in italic refer to variables in the system (for

[13] The following subsection is based on Frotjold (2005)

example *showerhead*), their properties such as the direction of their change (for example *a reduction*), and their ownership (for example *the other user*). The ownership is used to define which participant the text should refer to, such as *your showerhead* and *the other user's hot water setting*.

Another example of a generated text is:

> *"A decrease* in *your tap setting* at this time leads to an immediate *reduction* in *the temperature at your mixing battery,* **as well as** *a reduction* in *the fraction of hot water available to you.* **Additionally** it will lead to an immediate *increase* in *the temperature of the other user's mixing battery,* and *an increase in the fraction of hot water available to the other user."*

This text explains the nonlinearity in the system where the tap settings of both showers determine the amount of hot water that is available to each shower. The tap settings of both showers influence the temperature at each showerhead and the participant is told that in this case *her* reduction in the tap setting causes her water temperature to decline.

A decision regarding the level of accuracy in the representation of a change in variable values was made. The options were to state that there was a change in a variable, to state the direction of the change, that is, whether there was an increase or a decrease, or to present the exact measure of the decrease or increase. The middle solution was chosen because a too high level of accuracy could be confusing for the participant and would not be necessary for the intention of the model, that is, to understand the major underlying mechanisms of how the model works.

It was also necessary to decide which variables to monitor. In the case of the Two-Shower prototype it was decided to monitor the variables that were related to user input, because this would give the explanation a natural flow. In this case the input that is provided by a participant is particularly important because part of the participants' task is to understand how they can control the system through their input, and how their input influences the system behavior. In addition, monitoring changes in other variables, such as nonlinear variables, would create several explanation elements because a change in one of the input variables would affect several nonlinear variables.

The delays in the system are also monitored as they represent important events for the participant, that is, changes in the shower temperature. Representing the delays in the underlying model is important to support understanding of the two-shower model, its characteristics, and how it can be controlled.

### 3.4.7 Can the Story Generator be Defined as a Story Generator?

It is appropriate to ask whether the story generator in the Two-Shower prototype can actually be defined as a *story* generator. If the conclusion is that it cannot be defined as a story generator, one needs to discuss what could be done to develop it into a more advanced story generator.

A story or narrative has a plot which consists of a number of events driven by a series of cause and effect relationships. The narratives generated by the Two-Shower prototype consist of a series of cause and effect relationships and it is the participants' input and the simulation model that drives the story. As mentioned before, characters are also important elements of a story and the Two-Shower prototype has characters that compete or cooperate on sharing the same hot water resource. It does lack, however, something that is essential for being defined as a story; a story must have a tension curve in order to make it interesting. The Two-Shower prototype has an introduction (the introductory story in figure 3-2, p. 88), but there is, for example, no closure. It is also questionable whether it contains a tension curve.

In a future version of the model this can be solved by implementing a more advanced text generator and incorporating more sounds. For example; the sound of the two people in the showers shouting at each other, calling the receptionist to complain, the sound of the characters apologizing after having yelled at each other when they have been able to share the hot water resource equally. The language in the story generator could also be made more exciting in order to create more tension in the story.

Another problem in using the term *story generator* is that the participant may jump back and forth on the timeline. In a traditional narrative the plot is fixed according to how the author wishes to tell the story. In letting the participants jump back and forth on the timeline it may be difficult to create a tension curve in the story. It could be better to only let the participants study what has happened chronologically but this would also have to be discussed regarding how it is intended to support understanding.

The stories developed by the story generator cannot clearly be defined as narratives as they lack a clear narrative structure. To conclude, although the story generator clearly lacks some elements of what a story generator should include, it may be a starting point for further development of a more advanced story generator.

It is important to note that all of the suggested changes would have to be discussed based on how they were intended to support the participants' understanding of the underlying model. It may be tempting for designers to include as many advanced features as possible, forgetting that this may not necessarily support understanding.

This chapter has provided a first run-through of the model with a description and discussion of the narrative characteristics of the Two-Shower prototype. I run through the model and the various views once again in the next chapter (Chapter 4) with a focus on further visual elements of the revised model and some of the theories of visualization on which it is based. I add also a first discussion of structure and behavior in this chapter. This is expanded upon with a focus on transparency in Chapter 5 which deals specifically with design for understanding.

# 4. Enhancing the Two-Shower Model by a Multimedia User Interface

## 4.1 Introduction

The previous chapter was a description of how the original two-shower model was enhanced by including narrative characteristics in the user interface. This chapter deals with other multimedia elements (graphics, color, slide-shows, sound) that are incorporated into the user interface with the purpose of communicating particular characteristics of the underlying simulation model. Multimedia is used to visualize the structure of the original model, its behavior, and the relationship between structure and behavior. Because the decisions regarding the incorporation of narrative elements in the prototype, discussed in the previous chapter, are closely related to the decisions regarding user interface design and graphical representations there may be some repetition from the previous chapter.

Section 4.2 is a discussion of theory that supports the use of graphics and multimedia to enhance information. Further theoretical perspectives are introduced throughout the chapter as different issues are raised. Section 4.3 considers how the original two-shower model is enhanced by use of multimedia. There is a description of each view in the prototype with a focus on how structure, behavior, and the relationship between structure and behavior are represented.

## 4.2 User Interface Design

This section concerns some of the reasons for using multimedia representations. Different multimedia representation forms such as graphics, animation, and sound and their use in the Two-Shower prototype are considered.

### 4.2.1 Multimedia

According to the cognitive psychologist, Neisser (1976) more than one of the senses are utilized in most everyday perception experiences (for example taste, vision, touch, hearing). He argues that humans utilize the multiplicity in information in order to interpret what is being perceived. Information through different senses helps people interpret and understand what is perceived. "*Having heard something, we look to see it, and what we see then determines how we locate and interpret what we hear*" (Neisser, 1976, p. 29-30).

Multimedia is the use of different modes of representation such as graphics, text, audio, video, and animation to present information. Different senses are used to perceive, for example, sound and graphics. Multimedia may thus be used in a system dynamic interactive learning environment to represent aspects of the underlying model in a manner that may be interpreted by the participants through both vision and sound.

Alessi (2002) considers how multimedia may be utilized in interactive learning environments and how information may be provided using a combination of pictorial, textual, and auditory symbols. He discusses this in relation to how individuals may have different learning styles and the theory that some people prefer learning through visual material while others prefer information that is presented verbally.

According to the views of cognitive psychology three main principles related to perception and attention are relevant when incorporating multimedia in interactive learning environments:

*(1) Information (visual and aural) must be easy to receive.*
*(2) The position (spatial or temporal) of information affects our attention to and perception of it.*
*(3) Differences and changes attract and maintain attention.*

*(Alessi & Trollip, 2001, p. 21).*

Alessi and Trollip (ibid.) state that the concept of ease of perception may be used as a guide to screen design considerations. They refer to how ease of perception should guide decisions regarding text fonts, colors, or the size and level of detail in pictures. The ease of perception also affects decisions regarding the mode of expression that should be used, such as whether something should be presented graphically or by audio.

Regarding the position of information on the screen Alessi and Trollip note that important information is often placed close to the center of the screen while supporting information, such as for example menus, are placed closer to the edges. They further state:

> *"Differences and changes attract and maintain attention and underlie the use of various text sizes, colors, and fonts; changing backgrounds and music; and dynamic techniques, such as animation and motion video. Attention is drawn to change, whether it is dynamic (such as animation) or periodic (such as background color changing from one lesson segment to another)" (ibid., p. 21).*

In this way multimedia may be used to attract and maintain the attention of the participants in an interactive learning environment.

Mayer and Andersen (Mayer, 2001, 2005; Mayer & Anderson, 1991) also study the use of multimedia for learning and argue that using different modes of expression may add something to the learning process. *"There is reason to believe that – under certain circumstances –people learn more deeply from words and pictures than from words alone"* (Mayer, 2005, p. 3). What Mayer (2001) finds most interesting is that *"…understanding occurs when learners are able to build meaningful connections between visual and verbal representations…"* (p. 5). He emphasizes that graphical and verbal representations together may enhance what is being presented. *"In the process of trying to build connections between words and pictures, learners are able to create a deeper understanding from words or pictures alone"* (ibid., p. 5).

Mayer and Anderson (1991) discuss how graphical and textual explanations function together in learning of scientific material. They argue that:

> *"...effective understanding of scientific explanations requires a mapping between words and pictures. Presenting verbal and visual explanations together in a coordinated way was found to be more effective in promoting creative problem solving than was giving separate verbal explanations and visual explanations" (p. 484).*

By this they mean that textual and visual explanations together may provide additional value when learning than when a problem is presented either through visualizations or as written text. This is emphasized further in Mayer (2001) where he states that *"...words and pictures, although qualitatively different, can complement one another and that human understanding occurs when learners are able to mentally integrate visual and verbal representations"* (p. 5). Mayer further states that the characteristics of the material must be considered when deciding how it should be represented. He continues by arguing that:

> *"...words are useful for presenting certain kinds of material – perhaps representations that are more formal and require more effort to translate – whereas pictures are more useful for presenting other kinds of material – perhaps more intuitive, more natural representations" (ibid.).*

The implication of this is that the type of material or information that is being represented to a certain degree influence the mode of presentation.

In the Two-Shower prototype photos, as discussed previously, are used to represent the *feelings* of the characters in the shower, while a number with the exact temperature is displayed next to the photos. The photos are intended to have an intuitive effect on the participant, while the numbers may be used to observe and determine the exact shower temperature that the character *finds* comfortable.

The following is a consideration of different multimedia elements and how they are utilized in the Two-Shower prototype.

### 4.2.2 Graphics

Tufte (1997) investigates the arrangements of images into narratives with the purpose of forming designs that enhance richness, complexity, resolution, dimensionality, and clarity of what the images should mediate. A number of visual representations of data with emphasis on how cause and effect may be represented are studied. Tufte (ibid.) states that *"assessments of change, dynamics, and cause and effect are at the heart of thinking and explanation. To understand is to know what cause provokes what effect, by what means, at what rate"* (p. 9, emphasized in original). He asks the question of how knowledge about cause and effect should be represented graphically and analyzes various visual representations of dynamics and cause - effect relationships. He emphasizes the importance of making a close link between cause and effect as well as time in visual representations of data in order to be able to communicate graphically. The importance of clear visualizations of causal relationships is exemplified through a discussion of how graphics that lacked a link between cause and effect may have been part of the reason for the Challenger accident where a space shuttle exploded during take off in Florida in 1986.

A visual display where the order of events is unclear also distorts the links between the causes and effects. It is impossible to determine the direction of a causal relationship unless it can be placed in time. Tufte (1997) uses an example of a graph showing deaths during a cholera outbreak in London in 1854. In the original graphs made by John Snow it seemed that the removal of a water pump ended the epidemic. A closer study of the data, however, disclosed that the epidemic was already in decline before the pump was removed. The original graphs did not link the incidents in time properly thus displaying a causal relationship that did not exist.

One of the differences between the data that was presented by Tufte and the Two-Shower prototype is that the Two-Shower prototype is computerized. It is therefore possible to exploit the opportunities of animation. This may provide a better base for representing cause and effect and the placement of events in time. Before I go into a discussion about animation I provide a short consideration of the use of color.

### 4.2.3 Color

Both, Tufte (1997) and Spence (2001) that caution should be shown in the inclusion of colors in a visual display. *"Use of colour should be undertaken with caution, however, and not with the hysterical abandon of a child discovering its first paint box"* (Spence, 2001, p. 66). Tufte (1997) uses an example of two different versions of a map of the Japan Sea. In the first map differences in depths of the sea are illustrated by subdued, similar colors while in the second map depths are illustrated by highly contrasting colors. Tufte argues that strong contrasts in color are often used in scientific publications and states the following regarding the most colorful map: *"These aggressive colors, so unnatural and unquantitative, render the map incoherent, with some of the original data now lost in the soup"* (ibid., p. 77). The danger is that too much use of highly contrasting colors may suppress the data that one is trying to visualize. Tufte argues that *"minimal distinctions reduce visual clutter. Small contrasts work to enrich the overall visual signal by increasing the number of distinctions that can be made within a single image"* (ibid.).

The use of color was also considered for the Two-Shower prototype. This will be discussed further later in this chapter. To note now, however, bright blue and red were used to indicate cold and hot water, while the rest of the interface has relatively subdued colors.

### 4.2.4 Animation

There are a number of cognitive studies of how movement and animation affects perception. According to Neisser (1976) object movement may aid the perception of an object because the movement of an object may provide more information than an object that is idle. The movement of an object occurs over time and provides more information. A person may look at the moving object and view it from different angles. Different patterns of movement may provide additional information than that when observing a non-moving object.

Alessi (2002) argues that animation can be used to provide model transparency in system dynamic interactive learning environments because the system can be represented as simplified animations that makes it easier to understand, both for less experienced learners, and for more experienced learners who deal with very complex models. Animations of changing colors may also be useful for visualizing the behavior of the underlying simulation model. Animations with changing colors can be used to emphasize parts of the structure or to emphasize the state of the system and in this way make certain elements of the underlying model more transparent (ibid.). A further discussion about model transparency is provided in Chapter 5.

As discussed previously, Tufte (1997) emphasizes the importance of a clear placement in time in visual displays of information. *Animation* may be used to overcome some of the difficulties of placing information in time and visualizing the relationship between cause and effect. *"Animation is the term given to the addition of motion to images, making them move, alter and change in time"* (Dix et al., 1998, p. 598). Dix et al. (ibid.) argue that *"also, in the area of information visualization the most exciting developments are all where users can interact with a visualization in real time, changing parameters and seeing the effect"* (p. 136). They further state that:

> *"Animation can be used to great effect to show the changes in datasets, where slowly fluctuating changes can be visualized by the help of rippling three-dimensional coloured surfaces, or abrupt changes shown by sudden discontinuities in an otherwise regular motion" (ibid., p. 598).*

They argue that modern computers with greater computational power and high-quality graphics are used in data visualization to better represent data and allow for movement and manipulation of the objects in real time. They further state that:

> *"New ways of representing data, especially changing data, allow users to gain new insights into the behaviour of the systems they are trying to understand and make the computer an invaluable tool for understanding and discovery as well as for interpretation and mundane calculation (ibid.).*

One of the advantages of utilizing animation in a system dynamic interactive learning environment is that it may provide a means for better linking of the relationship between cause and effect. As the animation occurs in time this also provides a better means for placing the cause and effect relationships in time. The participant may provide input to the system and follow the resulting changes through the systems.

### 4.2.5 Video

Video may be used to make models more transparent (Alessi, 2002). Alessi argues that video may especially be beneficial for novice learners or learners who lack motivation. Video as a means to provide model transparency has so far not been used much in system dynamics (ibid.). Alessi argues that this may partly be because of the costs of producing the video material. There could also be a problem of using large video files although this is becoming a smaller problem as the capacity of personal computers increase.

As discussed previously, in the Two-Shower prototype, photo slide shows of a person that is showering is utilized rather than video recordings. In our second prototype, the Quito prototype, which will be discussed in Chapter 10, video clips serve as indications of the states of variables, explanations of some causal relationships, as well as hints about the influence they may have on the behavior (see also Djupvik, 2006; Skartveit, 2007, forthcoming; Skartveit & Viste, 2003).

### 4.2.6 Sound

According to Alessi and Trollip (2001) sound may be used to attract the attention of users, even when they are not looking at the screen. They argue that sound should not be the only alternative manner of representing information but that there should also be a textual alternative for the users who prefer text rather than audio.

Experiments of using digital sound to represent characteristics of system dynamics models have previously been performed. Pfeiffer and Lossius (2001) present a prototype for the use of sonification in system dynamic models. Their assumption was that use of different digital sounds could help a user learn more about the underlying structure of the system and obtain control of the system behavior. They demonstrated that the use of sound could support the analysis of the relationship between structure and behavior in two models. Their first case was a simple predator-prey model of the relationship between hares and lynx while the other described the dynamics of alcohol and yeast cells. They used pitch, timbre, and amplitude to denote changes in variable values, as well as to give warnings of shifts in dominance between the variables in the system. They also suggest combining sound and graphics as an extension of the project in order to utilize both hearing and vision.

This section has briefly touched the subject of representation of information through graphics and multimedia with the purpose of supporting understanding and learning. In the case of multimedia the theories were based on cognitive psychology. Learning theories that represent alternatives or supplements to cognitive psychology will be discussed in Chapter 5. The next section is a presentation of the particular multimedia elements that are developed with the intention of supporting understanding in the Two-Shower prototype.

## 4.3 Multimedia and the Representation of Structure and Behavior in the Two-Shower Prototype

The previous section considered a number of issues related to the use of multimedia in design of graphic displays, user interfaces, and interactive learning environments. In this section I describe how multimedia is used in the Two-Shower prototype to visualize the structure and behavior of the underlying model. The presentation here expands on the previous chapter but some repetition will occur, this time emphasizing different elements.

This section is divided into four subsections, the first three considering a different view of the Two-Shower prototype. As described previously, the Hotel view is the introductory view of the prototype, followed by the Shower Room view and the Pipe System view. The Shower Room view mainly portrays the behavior of the underlying simulation model, while the Pipe System represents both structure and behavior and to a certain extent the links between them. The assumptions underlying this organization are strongly related to the issue of model transparency which will be discussed at greater depth in Chapter 5. The section ends with a consideration of how structure and behavior is represented in the story generator.

### 4.3.1 The Hotel View

According to Alessi and Trollip (2001) the first view of an interactive learning environment should be a title page whose main function is to attract the attention of the participants. It should inform what the learning environment is about and there should be a set of instructions of how to use the interactive learning environment. The title page should not include issues that are part of what the participant should use for learning.

The Hotel view consists of a text with information about what the participants are about to take part in (see figure 4-1).[14] The task is introduced through the following text: *You will now participate in a simulation of two showers that share a hot water tank. You control one shower while another participant controls the other. Your goal is to obtain a stable comfortable temperature.*

The participants are asked to enter some login details. The *session id* may be used by the facilitators to log the session so that simulation data may be recorded and identified. This information must be provided to the participants by a facilitator. After entering the correct login details the participants must press the Ok button to start the session.

[14] The Hotel view was implemented by Frotjold (2005). The graphics in the view is designed by Skartveit and was part of the initial plans for the user interface in the spring and fall of 2002 (see Skartveit et al., 2003).


*Figure 4-1: The introductory view of the Two-Shower prototype: the Hotel view.*

The simulation model is then loaded. An *Introduction* text box appears over the Shower Room view (see figure 4-2 and 4-3). The *introductory story* that is displayed in the Introduction text box was described in Section 3.3 and displayed in figure 3-2, p. 88. The text is divided into chunks and the participant must click the arrows in order to make the text proceed. This both has the purpose of including some dramaturgy to the presentation of the text and makes it easier for the participants to read the text at their own speed. The participants may use the arrows to go back and forth in the text if they want something repeated.

As discussed in Section 4.2 several modes of multimedia representation may be used to enrich the information that is provided. In the Hotel view both text and graphics are used to introduce structural elements of the two-shower model. There is some reference to the structure in the text of the Hotel view. The participants are told that there are two showers that share one hot water tank. As mentioned previously, in the lower end of the view there are graphics representing a simplified hot water tank with pipes connected to two photos of people who are showering, each representing a shower in the system.

There is also some reference to structure in the Introduction text box (see figures 4-2 and 4-3).[15] The words in bold in the following excerpt represent respectively the shower, the tap, and the delay from a change in tap setting is made until the water with the new temperature reaches the showerhead:

> *The bathroom looks OK and the **shower** as well - until you turn on the water... You struggle to find the right temperature, turning **the tap** more and more desperately towards hot or cold with **no immediate results**.*

The bold words in the following text excerpt refer to the other shower:

> *Then you **hear the shower in the next room** being turned on, and it is the same story all over again. What happened? Remember that you can press the pause-button at any time during the simulation and receive comments about what has happened so far!"*

This shows that some reference to the structure is presented already in the Hotel view and the Introduction text box, but according to the narrative aspects (discussed in Chapter 3) and issues regarding model transparency (to be discussed in Chapter 5) not all information should be revealed concurrently. As mentioned previously in this section, the first view should also not contain information that the participants should use for learning. In this case, the references to the structural elements of the underlying model are mainly intended to provide a setting and an introduction to the task in which the participants are about to take part.

In the Hotel view and the Introduction text box there is no representation of actual behavior of the model because the simulation has not yet started. There is, however, some reference to the desired behavior in the text saying: *"Your goal is to obtain a stable, comfortable temperature"* (see figure 4-1).

There is also some reference to possible model behavior in the Introduction text box. The oscillating temperature is not mentioned explicitly but how the character in the

---

[15] The text in the introduction text box was changed for this version of the prototype. The version tested by Frotjold (2005) included an explanation of the system structure, while the latest version of the prototype described here provides this story in order to include more narrative elements to the interface.

story struggles to find the right temperature by turning the tap towards cold and hot is described. This provides the participants with a hint about how the temperature of the shower may oscillate.

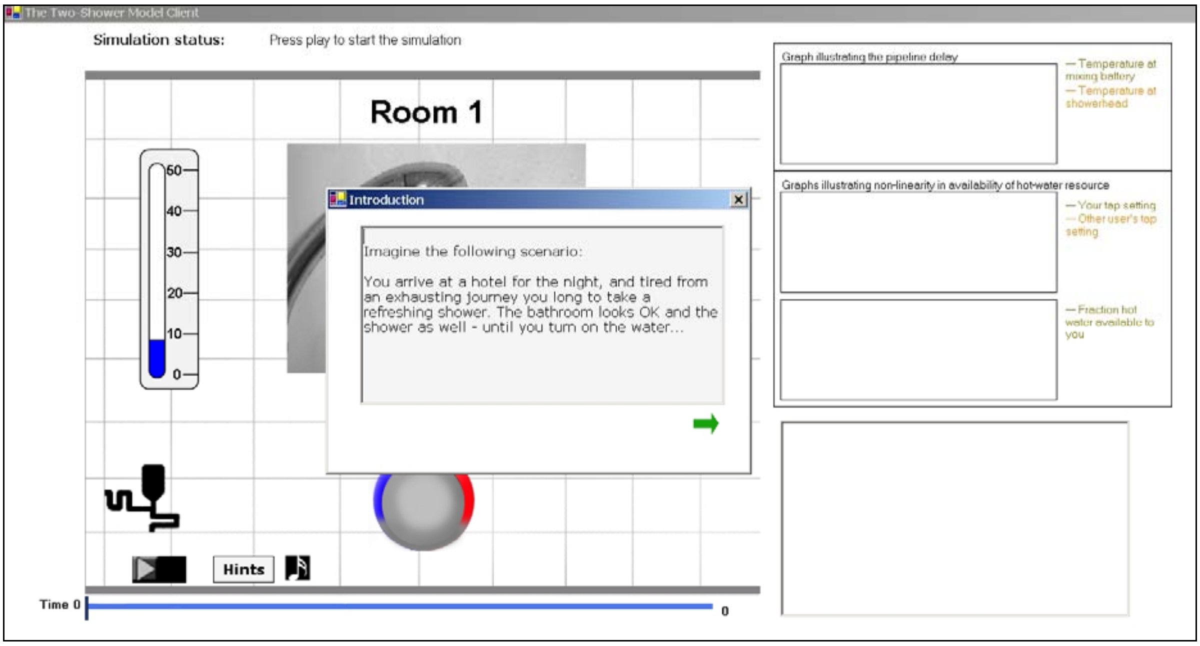


*Figure 4-2: Representation of the introduction story in the Shower Room view before the simulation is started.*

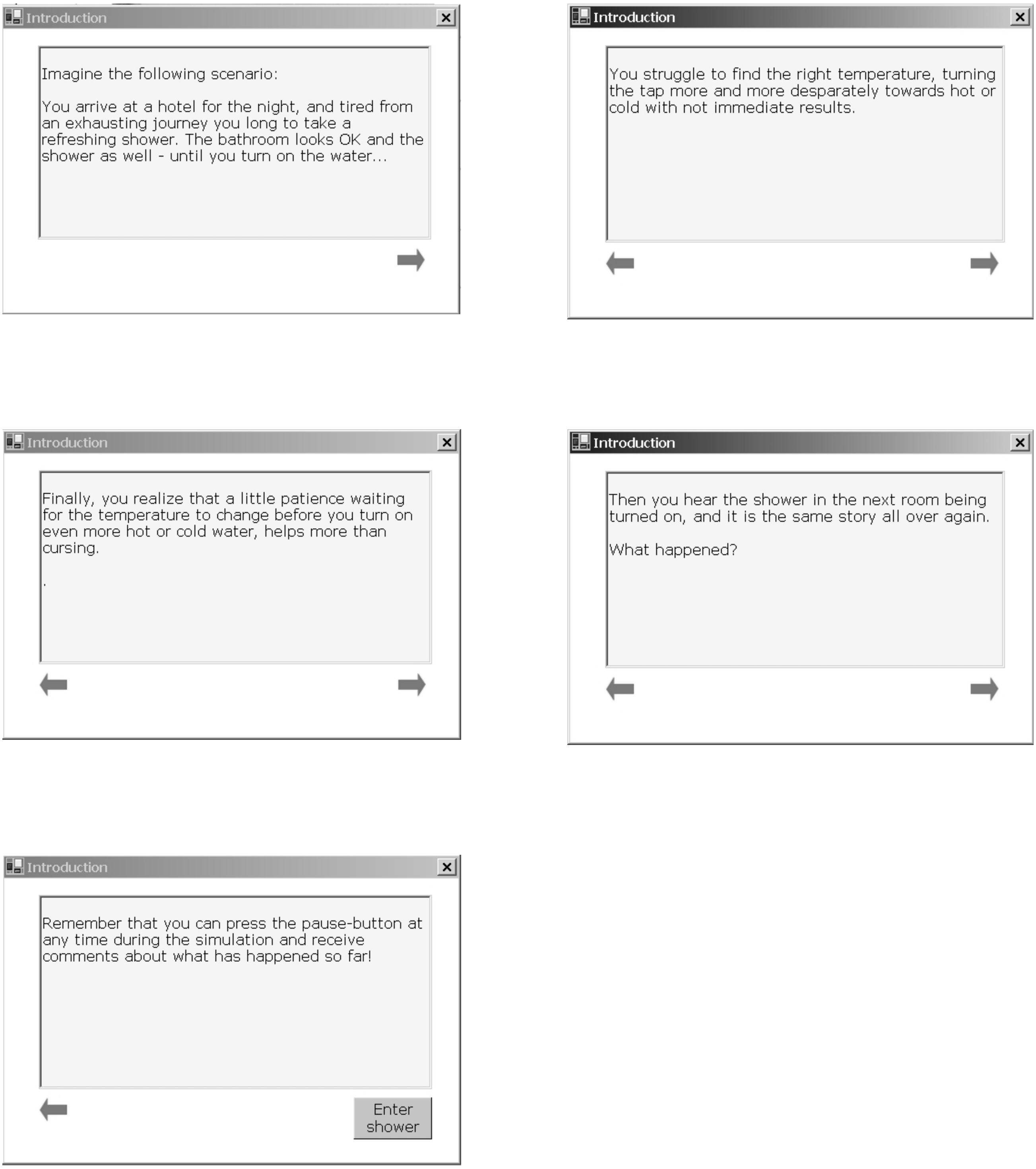


*Figure 4-3: The Introduction text box.*

### 4.3.2 The Shower Room View[16]

When the participant presses the Enter Shower button the Introduction text box disappears and the Shower Room becomes present (see figure 4-4). The Shower Room view and the Pipe System view (which will be explained in the next section) have some common elements that are displayed in both views. The text box in the lower right hand corner is used to display text that is generated by the story generator and an explanatory text that the participants may choose to read by pressing the Hints button. During the simulation run the text box displays a message saying that the participants can pause the simulation if they want to learn more about what is happening.

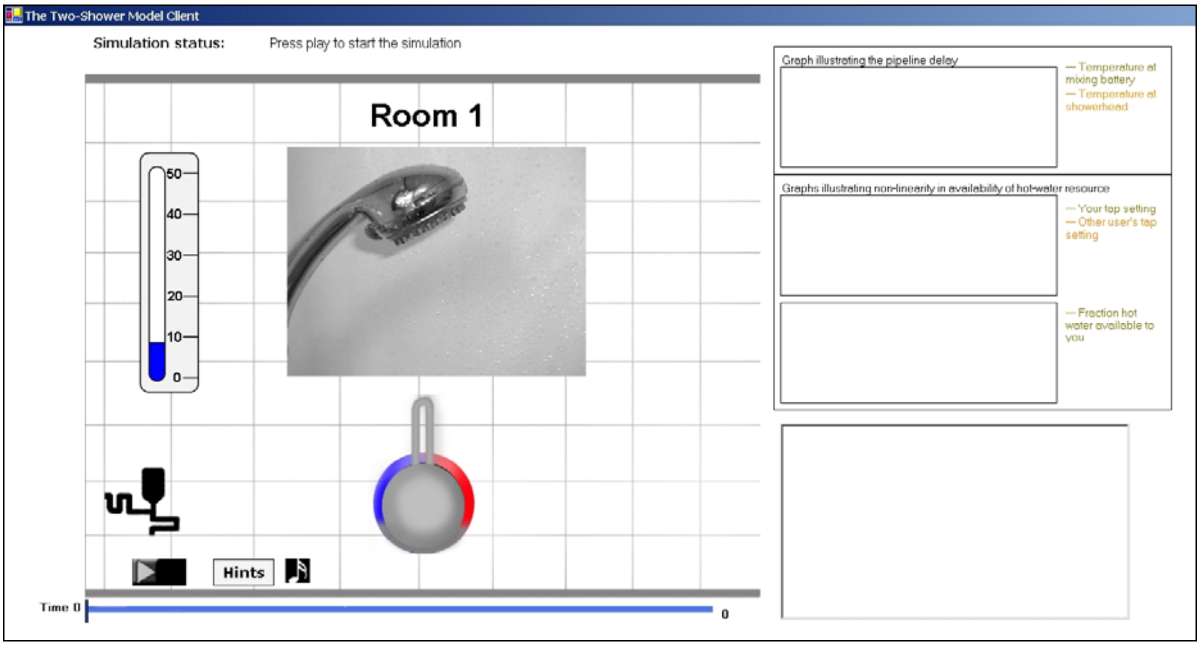


*Figure 4-4: The Shower Room view.*

By pressing the Hints button a relatively detailed explanation of the system structure is displayed in the text box. The text is as follows:[17]

---

[16] As mentioned before the graphics representing the bathroom for the Shower Room view were designed by Skartveit and myself. I implemented in the design in Flash.

[17] The text was formulated by Frotjold (2005).

*You are involved in a system where your decisions will have consequences for you in various ways, a certain time after you have made them. You must also realize that the same action may lead to different consequences, at different degrees, at different times. This may confuse your understanding of the system.*

*We are going to take a closer look at parts of the system, so that you can obtain a better understanding of it.*

*When your tap setting changes, it takes time before the temperature at your showerhead changes accordingly. This delay occurs because of the time it takes for the water to move through the pipes connecting the tap to the showerhead.*

*You can therefore not expect that a change in your tap setting will give an immediate change in the temperature of your shower.*

*When you change your tap setting, you change your demand on the hot water pressure. This changes the hot water pressure for the person in the other shower. It will not have an immediate effect on the temperature of the other shower because it takes time for the water with the changed temperature to move through the pipes.*

The intention of the text is to provide a more detailed description of the system that the participants can access if they are unable to understand the structure of the underlying simulation model.

As discussed in Section 4.2 there should be a clear indication of time. Below the shower room graphics there is a blue timeline that indicates the time the simulation is currently at. The yellow markers that are displayed as part of the story generator to indicate important events are also displayed on this timeline (see Section 3.4). In the upper right hand corner there are behavior graphs where the participants can study the behavior of some of the variables in the system over time.[18] Although some of the participants in *Test 2*, which will be described in Chapters 7 and 8 found these graphs too complicated, an indication of how the underlying simulation model has behaved over time is necessary. They were implemented in the last minute before the prototype was tested and there was unfortunately no time for refinement of their

[18] The behavior graphs were implemented in the latest version of the Two-Shower prototype by Olve Sæther Hansen and myself.

visual appearance. Such a refinement with the inclusion of units and, for example, a line that indicated the current time could improve the usefulness of the behavior graphs. This is further discussed in Chapter 9.

In the Shower Room view a number of multimedia elements are used to represent behavior, such as graphics, animation, color, audio, and photo slide shows. In the center of the view there is graphical representation of a mixing battery. This is where the participant controls the tap setting and it is the only access point that the participant has to controlling the behavior of the underlying simulation model.

Above the mixing battery there is a relatively large photo of a showerhead which represents the shower of the participant. On the left hand side is a thermometer. The simulation control buttons are placed above the timeline.

As discussed previously, the intention of the Shower Room view is mainly to represent the behavior of the underlying two-shower model with the purpose of triggering the curiosity of the participants and gradually introducing the underlying model.

In the Shower Room view animation and color are used in several ways to visualize the behavior of the underlying simulation model. The Two-Shower prototype is illustrated by relatively subdued colors. The hot and cold temperatures are visualized by red and blue colors because the changes in the temperature of the water are most important for the participants to monitor. The background of the Shower Room view changes color according to the temperature of the shower. If the temperature is cold it is blue, if it is too hot it is pink, and if it is comfortable it is just slightly pink (see figure 4-5). The blue and pink colors change gradually and become brighter pink or brighter blue according to the temperature of the water.

Changes in the shower temperature are also visualized through an animation of the thermometer in the left hand corner. The thermometer goes up and down and changes color in a similar manner to that of the background (see figure 4-5).

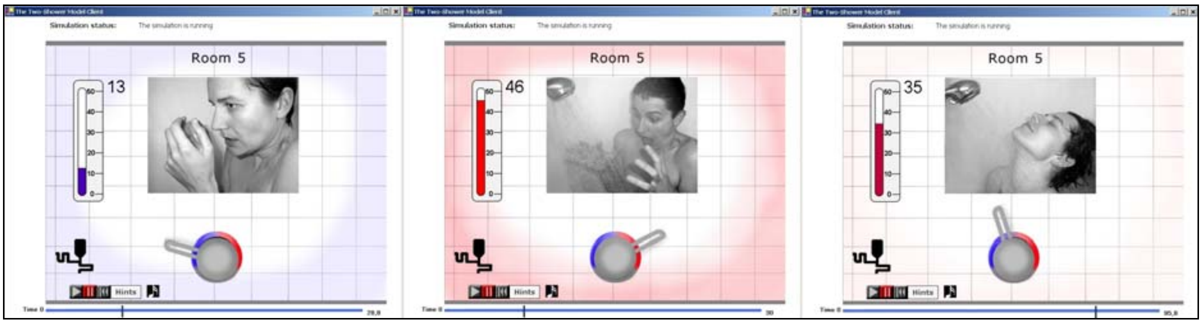


*Figure 4-5: Representation of behavior in the Shower Room view.*

As discussed in Section 4.2 caution should be shown when using colors to display information. Dix et al. argue that the color red is frequently associated with danger, stop, alarm or emergency and designers should be careful when using this color to represent something other than this. In the Two-Shower prototype red is used to represent hot water and bright red represents too hot water.

Many people may have difficulties perceiving the color blue (ibid.). Important information should therefore not only be displayed in blue without also providing other cues. There are, however, other cues about the temperature in the user interface of the Two-Shower prototype, such as a number indicating the exact temperature of the water.

The yellow warning pop-up that was discussed in Section 3.4 warns the participants about the extreme temperatures and suggests that they pause the simulation. It has a bright, contrasting color, in order for it to be noticed by the participants.

As discussed previously, in the Two-Shower prototype slideshows are used to show the behavior of the temperature of the water at the showerhead.[19] The photos are edited as four series of photos for each shower. One series represents the introduction where the character in the shower turns on the water (see figure 4-6). This slideshow is presented when the participant presses play. There are also series that represent cold, comfortable, and hot temperatures. The slideshows that represent the temperature of the water in Room 1 are presented in figure 4-7. Figure 4-8 shows the

[19] As mentioned in Chapter One the photos are taken by Skartveit and the use of photo slideshows for the prototype was discussed in Skartveit, Goodnow and Viste (2003), Skartveit and Viste (2003), Viste and Skartveit (2003; 2004).

slideshows that are used to indicate the temperature of the water in Room 2. Each series of pictures starts with a neutral photo, such as of the showerhead with water running from it, in order to make the transition from one state to another smoother. The behavior is here represented as an aggregation and divided into three states; one that represents cold, one that represents too hot, and one that represents comfortable temperatures. When the temperature of the water changes between the different states the slideshow may look like the one presented in figure 4-9.

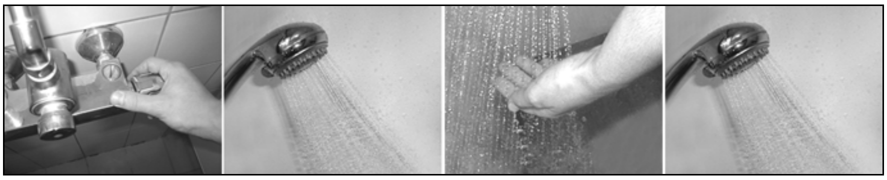

*Figure 4-6: The introductory slideshow displayed for both rooms.*

*Figure 4-7: The photo slideshows for Room 1, indicating cold, comfortable, and hot temperatures.*

Figure 4-8: The photo slideshows for Room 2, indicating cold, comfortable, and hot temperatures.

Figure 4-9: Example of slideshow in the Two-Shower prototype.

In addition to animation, changing colors and photo slideshows, sound is incorporated into the prototype. When the temperature reaches a too hot level the sound of a person who is screaming is triggered. The sound of a shower is played when the underlying simulation model is running. This was implemented to contribute to the shower atmosphere. After some time both these sounds may be annoying and may therefore be switched off by pressing a *mute* button. The sound was not used in Test 2 of the prototype (which will be discussed in Chapters 7 and 8), however, some participants asked to test the prototype with sound.

As discussed in this subsection color, photo slideshows, and audio have been used in the Shower Room view to represent the behavior of the underlying simulation model. In the Pipe System view some of these characteristics have been used in a similar manner, but here there is to a larger degree also emphasis on the structure.

On the left hand side of the user interface there is a button that is shaped like a hot water tank with connected pipes. If the participant presses the Pipe System button she enters the Pipe System view.

### 4.3.3 The Pipe System View[20]

The previous section discussed how the Shower Room view mainly represents the behavior of the underlying simulation model. The Pipe System view represents to a greater extent both structure and behavior (see Figure 4-10). Through this view the participants are taken inside the walls of the hotel to explore the structure of the plumbing, how the two showers are connected, and the consequences for the temperature of the water.

In addition to representing structure and behavior, this view represents to some extent the relationship between structure and behavior. There is a link between structure and behavior in that the participants may follow the consequences of an action by tracing the action through the structure and behavior.

In the Pipe System view, graphics are used to represent the hot water tank, how the mixing batteries of both showers are connected through the tank, and how the two showerheads are each connected to their respective mixing batteries.

[20] As mentioned before the graphics for the Pipe System view were designed and implemented in Flash by myself.

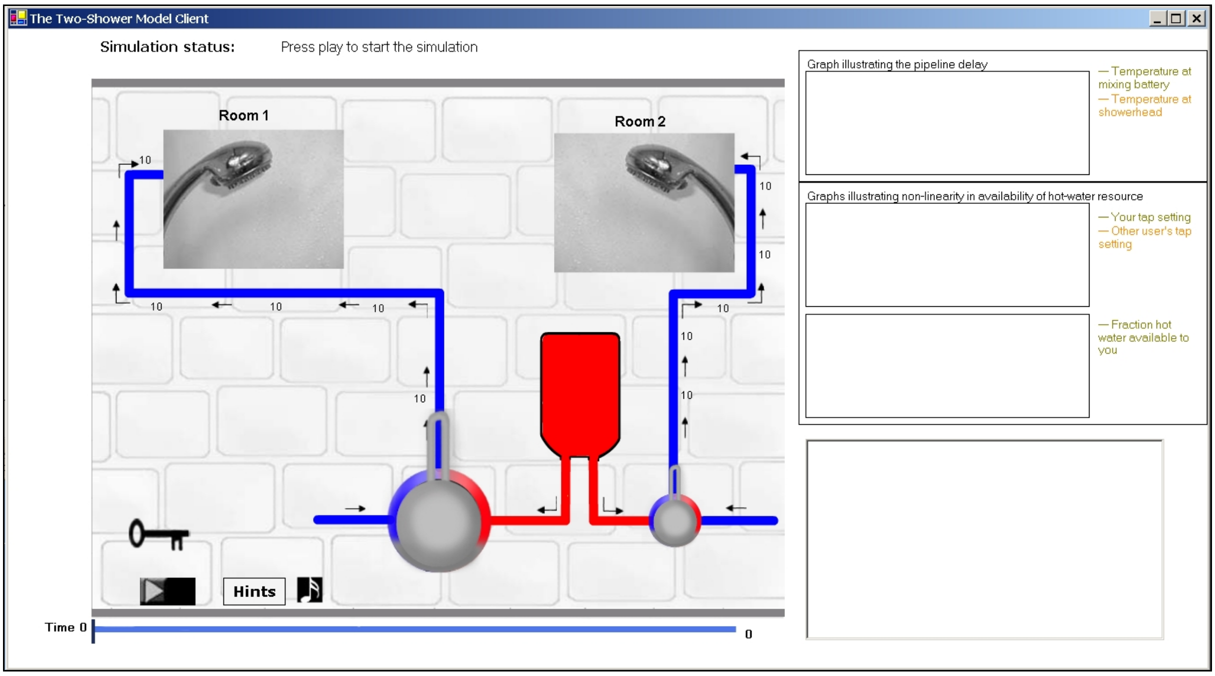


*Figure 4-10: The Pipe System view.*

As explained previously there is a delay from a change is made in the tap setting until the water with the new temperature reaches the showerhead. Animation and changes in color are used to represent this pipeline delay. Between the mixing battery and the photo slideshow, which represents the temperature of the water at the showerhead, there are animations of water that is flowing through the pipes. The representation of the pipe is divided into the number of seconds it takes for the water to flow through the pipes, and each section of the pipes have blue or red colors according to the temperature of the specific section. If the tap setting is changed, the flow of the water with the new temperature through the pipes is animated.

There could be some discussion as to whether the division of the pipeline into seconds is a realistic representation. Water flows continuously, not in steps. It is implemented in this way, however, to emphasize the number of seconds it takes from a change is made in the tap setting until the water with the new temperature reaches the showerhead.

In the same manner as in the Shower Room view the colors of the pipes are supplemented by numbers indicating the exact temperature of the water. There are also arrows indicating the direction in which the water is flowing.

As discussed in Chapter 1 the interface went through several changes during the development process. Figure 4-11 Represents an initial version of the Shower Room view, as it was presented at CSCL2003 (Viste & Skartveit, 2003). After discussions at the conference the interface was changed with the purpose of displaying a clearer relationship between structure and behavior. The reason was that it was difficult to see how decisions to change the tap setting affected the temperature of the water. It was also found that the structure of the pipe system with the mixing batteries was not represented sufficiently.

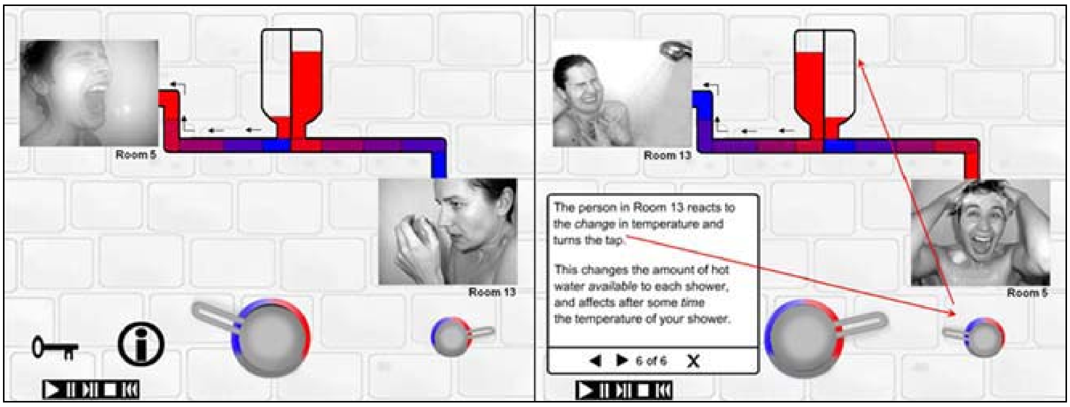


*Figure 4-11: The Pipe System view in the initial version of the prototype.*

The nonlinearity in the access to the hot water resource was difficult to visualize. In the earlier versions of the prototype this was represented by a hot water tank that was divided vertically by a line (see figure 4-12). An animation of the hot water moving up and down according to the change in the access to the hot water was implemented. Frotjold (2005), however, found that the participants in her test had problems understanding this visualization (Frotjold's test, *Test 1* of the prototype will be discussed in Chapter 6).

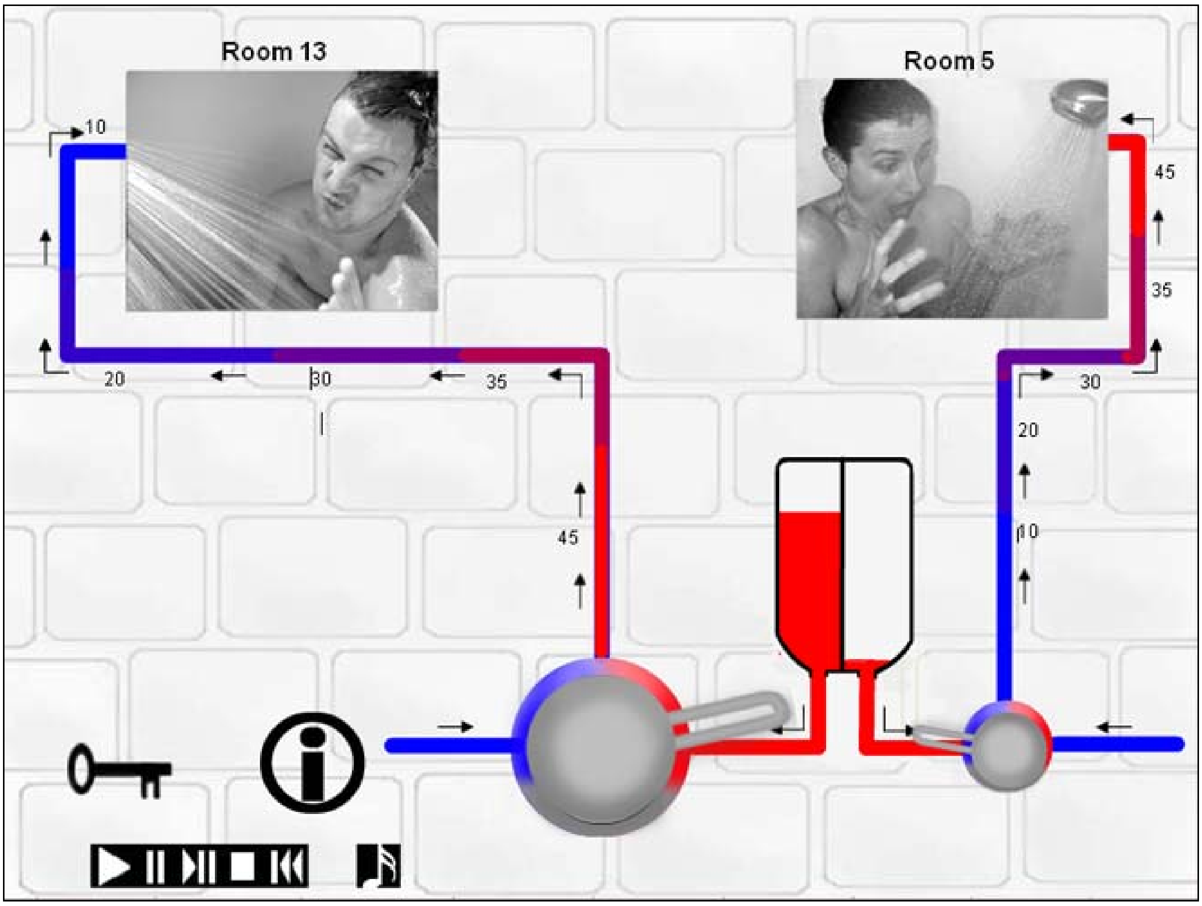


*Figure 4-12: The Pipe System view with an older version of the representation of nonlinearity in access to the hot water resource.*

The representation of the hot water tank was therefore changed in the latest version of the Two-Shower prototype. The nonlinearity is now represented through small yellow lines that change speed according to how fast the water flows through the pipes (see figure 4-13). When a participant has access to a large portion of the hot water, that is, if this participant has her tap setting turned towards hot while the other has his tap setting turned towards cold, the yellow lines move quickly. As the access to the hot water decrease, the speed of the yellow lines also decreases. The second test of the prototype, which is presented in Chapter 8, however, shows that also this solution may need some reconsideration.

In the Shower Room view only the tap setting of the current participant was visible. In the Pipe System view, however, there is an animation of the changes in the tap setting of the other participant. The decisions of the other participant to change the tap setting are in this way shown as an animation on the current participant's screen. The participants can follow each other's decisions and their effects throughout the simulation run. This is intended to support *grounding* between the participants and will be further discussed in Chapter 5.

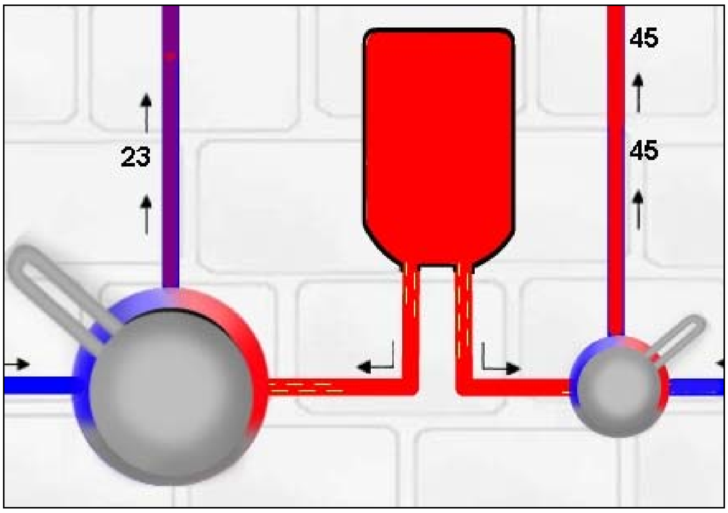


*Figure 4-13: The current representation of the nonlinearity in the Pipe System view.*

In Section 4.2 it was discussed how there should be a close visual link between cause and effect. When a change in tap setting is made in the Pipe System view, the results are seen immediately in the pipeline and it is possible to follow the water with the new temperature to the showerhead. At the same time, as mentioned previously, the hot water that flows from the hot water tank changes speeds according to the tap settings of both showers. The participants can trace in this way the causes and effects through the interface.

If the participant wishes to return to the Shower Room view she can press a Key button on the left hand side of the view.

### 4.3.4 The Story Generator

The story generator was discussed in detail in Chapter 3. As shown then it has the same graphics as the other views in the prototype but includes an additional textual element with descriptions about changes in the underlying simulation model over time (see figure 3-7, p 98). There is a stronger representation of the relationship between structure and behavior than there is in the other views. In the story generator the text directly refers to changes in the model, and the direction of these changes, based on input by the participants. The animations change according to the state of the system at the particular point in time. I will not repeat the description of the story generator here, but emphasize that it as an example of how text, graphics, and animation are used to represent the relationship between structure and behavior of the underlying simulation model.

This chapter has been a further description of issues considered in the development of the Two-Shower prototype. The focus has mainly been on the use of a variety of multimedia elements to represent characteristics of the system. The user interface is organized as a gradual introduction to the structure, the behavior, and the relationship between structure and behavior of the underlying simulation model. This gradual presentation is closely linked to the issue of transparency, which will be discussed in the following chapter.

# 5. Designing for Understanding

## 5.1 Introduction

The previous two chapters dealt with the use of narrative and graphic design in development of the Two-Shower prototype. This chapter deals with further design decisions that were made, based on learning theory, with the purpose of enhancing the original two-shower model. Learning theory affected design decisions in all versions of the prototype (both single and multiple users). Although the tests of the interfaces did not focus on learning but on design-functionality, concepts from learning theory that were useful in the design process included transparency, interactivity, and collaboration.

A number of decisions regarding how the underlying model is to be represented had to be made particular in relation to the intended audience of early learners of system dynamics. Some of the decisions related to interactivity, as in this case, how the user should be able to interact with or control the underlying simulation model through the user interface. In the case of the Two-Shower prototype, two participants are intended to collaborate. Decisions regarding how the user interface should support their communication had to be made.

In Section 5.2 the cognitive perspective on knowledge and learning is discussed, followed by a discussion of the constructivist perspective in Section 5.3. The concept of transparency, (how information about the underlying simulation model is portrayed through the user interface), is addressed in Section 5.5. Section 5.5 is a discussion of interactivity and how theories of interactivity have influenced the development of the Two-Shower prototype. The chapter ends with Section 5.6, in which collaborative features of the Two-Shower prototype are described based on the theories of Computer Support for Collaboration and Learning (CSCL).

## 5.2 The Cognitive Perspectives on Context and Learning and the Implications for Design

In this section the focus is on theories and design methods of learning. As with the later parts of this chapter I, will tie learning theory with the design of the prototype. I open with a brief discussion of the cognitive perspective and the view of the nature of learning and understanding within the cognitive perspective.

There are currently two main positions of how learning occurs and thereby how educational environments should be constructed. In the following I will give a short introduction to the cognitive approach and briefly discuss the implications for the development of our prototypes. This section is followed by a section on the constructivist perspective.

Cognitive psychology emerged as a response to behaviorism in the 1950s (Anderson, 2005). The focus of cognitive theory is to understand how the intelligent human mind functions (Anderson, 1995, p. 1). There is a focus on the individual, individual thought, and knowledge. In cognitive theory internal processes such as perception, memory, attention, pattern recognition, problem solving, the psychology of language, and cognitive development are of concern (Neisser, 1976). Anderson emphasizes thought processes and the organization of knowledge: *"Cognitive psychology is the science of how the mind is organized to produce intelligent thought and how it is realized in the brain"* (Anderson, 2005, p. 1).

In cognitive theory, learning is seen as acquisition of knowledge and knowledge structures (Shuell, 1986). Learning is internalizations of perceptions, experiences, and reasoning. An interactive learning environment based purely on cognitive theory would be developed with the purpose of supporting the participants in internalizing perceptions and reasoning.

### 5.2.1 Mental Models in Cognitive Theory

Within cognitive theory there have been theories dealing with *mental models* – the way people organize knowledge internally. The cognitive approach embraces a number of theories and models of how reasoning and decision-making are organized. The concept of mental models is addressed by only one such theory. I chose to focus on that theory as it sheds light on how mental models are viewed in system dynamics and, therefore, have been part of our deliberations regarding the design of the Two-Shower prototype.

There is debate about how people organize knowledge and mental models - whether people approach a problem with pre-existing models in their mind or create them at the moment they meet the problem. In system dynamics the leading view is that mental models are pre-existing models that are challenged and changed through new perceptions (Doyle & Ford, 1998). Cognitive psychologists like Johnson-Laird (1983; 1989) represent the view that mental models are developed rapidly to make inferences during the course of perception or thinking.

The concept of reasoning based on models was first introduced by Craik (1966) in the 1940's.[21] He notes that *"...models seem to play an important part in facilitating thought…"* and that *"the process of modeling does not stop outside the body"* (ibid., p. 68). From Craik's perspective people create symbolic representations or models in their minds which represent external events (Doyle & Ford, 1998, p. 8).

According to Johnson-Laird (1989) mental models are used to make rapid and automatic inferences, to fill out the missing details of perceptions, discourse or knowledge. People are seldom aware that they use their mental models. Johnson-Laird refers to an earlier publication (Miller & Johnson-Laird, 1976) and argues that people are able to understand three principles which make it possible for them to

[21] The work on *The Mechanism of Human Action* was begun by Craik in 1943, but he died in a car accident in 1945. A collection of his works was published by Sherwood (1966).

construct mental models: *"first, in a deterministic domain all events have causes; second, causes precede their effects; and third, an action on an object is the likely cause of any change that occurs in it"* (Johnson-Laird, 1989, p.483). People have to understand the premises before they can construct the model. There are three semantic procedures when constructing a mental model. The first is to construct the mental model based on the premises. The second is to make the conclusion based on the premises (ibid.). The third step is to seek alternative mental models that may refute the conclusion.

Seel (2003) discusses how instruction may be organized in order to support the forming of mental models. He discusses how mental models are internalizations of external representations and that the construction of external representations could be done with the objective of supporting the forming of mental models. Seel (ibid.) suggests presenting a conceptual model at the beginning of a learning session with the purpose of supporting the formation of mental models. Seel refers to Mayer (1989) who also asserts the view that conceptual models may support mental model formation. Mayer explains that a conceptual model can be represented by words or diagrams for supporting learners in forming mental models of a system. The conceptual model would emphasize important objects, actions and their causal relationships. According to Mayer (1989) presenting a conceptual model may especially be adequate for novice learners arguing that conceptual model may be most beneficial for those who have little or no experience and who have not formed their own mental model of the problem.

Goldvarg and Johnson-Laird (2001) argue that humans make causal deductions based on mental models. They have performed a number of studies of mental models (see for example Goldvarg & Johnson-Laird, 2001; Johnson-Laird, 1989). They further argue that their studies show that people use mental models rather than *schemas* when making causal deductions. Schema theory is another theory of how deductions are made within cognitive theory. The schema theory has some commonalities with the theories of mental models in system dynamics. The following section therefore offers a brief overview of schema theory in cognitive theory.

### 5.2.2 Schema Theory

In cognitive theory the concept *schema* relates to how humans deal with categorical knowledge: *“Schemas are abstractions from specific instances that can be used to make inferences about instances of the concepts they represent”* (Anderson, 2005, p. 158). Kelley (1973) argue that causal inferences are based on schemas. He represents what he calls a *causal schema*: *“A causal schema refers to the way a person thinks about plausible causes in relation to a given effect”* (ibid., p. 114). A causal schema is formed based on observations, experience with causal relationships, and from being taught about causal relationships. A set of abstract conceptions of causal relationships are used when making inferences or decisions. These conceptions form a framework that is used to fit new information (ibid.).

Some of the main criticism against schema theory is that it is vague and that there are no descriptions of how schemas are formed (Sadoski, Paivio, & Goetz, 1991). A second criticism concerns how it is possible to claim the existence of an abstraction. Schemas are derived through experience: *“The epistemological question is how conceptual or schematic knowledge can exist in the abstract, isolated from any of the examples that gave rise to it”* (ibid., p. 467). Sadoski et al. (ibid.) also reference Alba and Hasher (1983) who argue that the abstractions in schema theory fail to explain the details observed in memory research. According to Reynolds, Sinatra, and Jetton (1996) schema theory can also be criticized for involving a strong focus on what is going on in the mind, leaving out the impact of the social and cultural aspects of the environment.

### 5.2.3 Mental Models in System Dynamics

The concept of *mental models* plays an important role in system dynamics as it is used to represent a person’s knowledge about a complex system. From a system dynamic point of view, mental models must be studied in order to understand decisions that are made by actors in a system, to see what a person understands about a system and how his or her understanding changes through modeling and operating simulation models.

As indicated earlier in this section, the concept of mental models in system dynamics differs from the use of the concept in cognitive theory. Rather than being formed rapidly through perception or discourse, mental models are most often viewed as relatively stable cognitive representation in system dynamics. Although they are relatively stable they are subject to change based on new perceptions and inferences.

Several authors have written about the concept of mental models in system dynamics (see for example Doyle & Ford, 1998; Doyle & Ford, 1999; Forrester, 1961; 1971a; 1980; 1994a; 1994b; 1997; Lane, 1999; Morecroft, 1994; Richardson & Pugh, 1981; Vennix, 1990). I will not go into the details of each view here. Doyle and Ford (1998), however, have made a review of the system dynamic literature on mental models and recognized that there was no unified definition of the term mental model within system dynamics. They attempted to make a unified definition. They emphasize that a mental model is a *cognitive phenomenon* that only exists in the mind. They point out the lack of distinction between mental models and their representation by, for example, causal loop diagrams and stock and flow diagrams within system dynamics. Representations are products of attempts to elicit mental models (ibid.). Doyle and Ford (ibid.) presented the following definition:

> *"A mental model of a dynamic system is a relatively enduring and accessible, but limited, internal conceptual representation of an external system whose structure maintains the perceived structure of that system" (ibid., pp. 17-21).*

This definition was criticized by Lane (1999) for not taking into account that mental models can represent a *planned* or *imagined* system that has not yet come into being. Doyle and Ford (1999) answer by including the term *projected* instead of planned – projected meaning *"planned, figured, or estimated for the future"* (p. 413). They also add the term *historical* and point to systems that have existed, but that no longer exist:

> *"A mental model of a dynamic system is a relatively enduring and accessible, but limited, internal conceptual representation of an external system* (historical, existing, projected) *whose structure is analogous to the perceived structure of that system" (p. 414, emphasis in original).*

Sterman (2000) also comments on the concept of mental model:

> *"In system dynamics, the term 'mental model' includes our beliefs about networks of causes and effects that describe how a system operates, along with the boundary of the model (which variables are included and which are excluded) and the time horizon we consider relevant – our framing or articulation of the problem*" *(p. 16).*

According to this view, mental models contain quite specific characteristics such as variables, model boundaries, and a time horizon. Sterman further argues that we can be both consciously aware and unaware of when we use our mental models. Sterman (2000) describes a mental model as something that is actively constructed by our senses and the brain.

Sterman (ibid.) argues that our previous mental models influence our perception and that we use what we have perceived to elaborate on our mental models. According to Sterman, mental models are changed through feedback containing information about the real world. The changes to the mental model will then alter the basis for our decisions.

According to Doyle & Ford (1998) the majority of system dynamics researchers see mental models as relatively enduring cognitive structures, not as temporary mental structures that are formed on the spot. The view of mental models as relatively enduring cognitive structures has commonalities with schema theory, discussed in the previous subsection.

### 5.2.4 Design Implications for the Two-Shower Prototype

Whether people approach a problem with pre-existing models in their mind or create them at the moment of recognizing a problem, the design of the user interface of the Two-Shower prototype should support the participants in making inferences about the underlying simulation model.

The theories of mental models may be linked to the design of the Two-Shower prototype in a number of ways. I do not borrow directly the idea of mental models as pre-existing in participants' minds. In the design, however, we tried to take into account that people have a pre-conceived idea about how things work. The user interface is, as discussed previously, based on a situation that people are expected to have experienced in real life. The intention is that the participants would try to use what they already know: their previous experiences with showers. For the Quito prototype, to be discussed in Chapter 10, the intention is that the participants would use their experience with the organization of urban spaces. These perceptions are harnessed to get participants started in trying to understanding the model, but are also challenged by revealing gradually how the system is more complicated than what people tend to believe. If asked, the participants would not necessarily be able to present their pre-conceived ideas as a model but they may try to grate a model in their mind. .

If one sees mental models as formed rapidly based on perception and discourse (as Johnson-Laird, 1989), then the Two-Shower prototype should support the participants in making inferences about the underlying model based on the representations in the user interface. In the Shower Room view the goal was, for example, that the participants should be able to infer that a change in tap setting would lead to a change in the temperature. Changes in temperature were indicated, as described in Chapter 4, by changes in the photo slideshow, the animated thermometer, and the color of the background. Even though the relationship between the change in the tap setting and the resulting temperature is not represented by, for example, a conceptual model, the intention was that the participants should use the information in the user interface to infer that relationship.

Similarly, in the Shower Room view (see figure 4-4, p. 121), an animation representing the decision of the other user to change the tap setting, the representation of the pipes and the hot water tank were intended to support the participant in inferring that the change in tap setting would also affect the current participant. All representations of the underlying model in the user interface are developed with the purpose of supporting the participants in making inferences about the underlying model although they cannot see the stock and flow diagram or equations. This also relates to transparency, which will be discussed in Section 5.4.

So far, the focus of this chapter has been on cognitive theories of understanding and learning. Sfard (1998) discusses two well-established metaphors of learning, namely *learning as acquisition* and *learning as participation*. The theories of mental models discussed in this section generally deals with the acquisitions metaphor of learning, seeing learning as internalizations of knowledge. The following section contains a discussion of the constructivist perspective that can be related more closely to the participation metaphor of learning.

## 5.3 The Constructivist Perspective on Context and Learning and the Implications for Design

The view of focusing on learning as an internal process is criticized because the context in which learning occurs is not taken into account (Lave & Wenger, 1991). When learning is viewed as an internal process, only assumptions about how learning occurs can be made. Lave and Wenger made the following comment regarding the focus on learning as internalizations: *"It establishes a sharp dichotomy between inside and outside, suggests that knowledge is largely cerebral, and takes the individual as the nonproblematic unit of analysis"* (ibid., p. 47).

In later years constructivism has been a growing theoretical foundation and there has been a focus on learning as a socially situated activity. Constructivism has its background in Piaget's developmental psychology (Koschmann, 1996b) and can be seen as a fusion of socially shared cognition, situated learning, activity theory, and Deweyan pragmatism (Jonassen et al., 2000). One of the fundamental differences between constructivism and previous theories such as behaviorism is that humans are seen as natural learners who attempt to make sense of their surroundings:

> *"Humans interact with and experience their environment and naturally seek to understand those interactions by developing their own theories in action and sharing them with others" (ibid., p. 107).*

Jonassen et al. (ibid.) list three fundamental differences between the theories on which constructivism is based and previous theories of learning related to *meaning making*, *social aspects of learning* and *distributed cognition* (pp. 108-109). The first one is that rather than seeing learning as transfer of knowledge, learning is an active process in which the learner ascribes meaning. Knowledge must be constructed by the individual and can therefore not be transferred from one person to another. Further, knowledge is not a stable entity but constantly changing. Important implications of the view of knowledge as meaning making are that it must be interpreted in a context and in relation to the interactions within this context (ibid.).

Salomon and Perkins (1998) argue that there is no *solo learning* because what one learns is deeply grounded in the social cultural context. The tools and concepts that are used for learning are part of this cultural context and influence learning even if the physical presence of others are lacking. Even though an individual may be located in a room by herself, learning is more than her internalization of concepts; learning is influenced by the context.

Second, this *meaning making* process is social in nature. Learning is the process of negotiating meaning among participants in an activity. *"Learning in this perspective is dialogue, a process of internal as well as social negotiation"* (Jonassen et al., 2000, p. 109).

The third theory that characterizes constructivism is *distributed cognitions*:

> *"Not only does knowledge exist in individual minds and in socially negotiating minds, but it also exists in the discourse among individuals, the social relationships that bind them, the physical artefacts that they use and produce, and the theories, models, and methods they use to produce them" (ibid., p. 109).*

Knowledge is in this manner distributed among members in the society. Salomon and Perkins (1998) who refers to Pea (1993) argue:

> *"knowledge, rather than being transmitted or internalized, becomes jointly constructed ('appropriated') in the sense that it is neither handed down ready-made, nor something one constructs on his or her own. Rather, knowledge, understandings, and meanings gradually emerge through interaction and become distributed among those interacting rather than individually constructed or possessed" (Salomon & Perkins, 1998, p. 6).*

From a constructivist perspective, learning can only take place intentionally. It is a conscious action of reflection (Jonassen et al., 2000). If there is a dissonance between previous understanding and what is being perceived, the learner will seek to solve the problem and remove the dissonance. The will to solve the problem and the solving of it can only be conscious actions.

According to Jonassen et al. (2000), constructivist learning environments should be case-, project- and problem-based and one should seek to support the learners in articulation, problem-solving, and reflection. These learning environments should also support collaboration and include cognitive tools for scaffolding. Support for collaboration in Two-Shower prototype will be discussed further in Section 5.6. Section 5.4 deals with transparency that may be related to support for scaffolding.

According to, Milrad, Spector and Davidsen (2003) theories such as situated theory and cognitive flexibility theory, which are both constructivist views of learning, are not to a significant degree embraced by the system dynamic community. This is the case both in the case of constructing system dynamic based interactive learning environments and studying learning with models. Most studies of learning in the system dynamic literature are based on a cognitive approach, with a focus on learning particular models and on changing of the subjects' mental models.

Milrad, Spector and Davidsen have made substantial efforts in incorporating constructivist theories into system dynamic interactive learning environments (see for example Milrad, 2006; Milrad et al., 2003; Spector, 2000). I agree with them that incorporating views of constructivist theory may contribute to the design of system dynamic learning environments. The Two-Shower prototype was developed in light of constructivist theory and the belief that learning is a social activity. Support for collaboration, drawing on previous experience, and scaffolding of information and was an important part of the design ideas. This will be discussed further throughout this chapter. The prototype, as previously discussed, is not intended to be used as a standalone product, but as part of an educational setting where the students and the facilitator or teacher also play important parts. The Two-Shower prototype is not valuable in itself – it is how it is being used for learning purposes, that is essential (although this is not tested in this thesis).

### 5.3.1 System Dynamics-Based Interactive Learning Environments

Davidsen, Spector and Milrad (Davidsen, 2000; Davidsen et al., 1999; Milrad, 2006; Milrad et al., 2003; Spector & Davidsen, 1997b, 1998), however, offer a more specific focus on constructivist system dynamic-based learning environments and present a method which they call *model facilitated learning* (Milrad et al., 2003). Their theoretical perspective is grounded in situated and problem-based learning (Lave & Wenger, 1991) and cognitive flexibility theory (Spiro, Coulson, Feltovich, & Anderson, 1988). Their instructional design methods are based on elaboration theory (Reigeluth & Stein, 1983). Their design suggestions, however, relate particularly to the development of system dynamic learning environments.

Four principles that should guide the development are presented (Milrad et al., 2003). The problems and complexity of the domain should be introduced in order to situate the learning experience. Complexity should then be gradually introduced before the learners are challenged by dealing with increasingly complex problems. Finally, the learners should be challenged to develop policies.

Davidsen (2000) notes several issues, more specifically related to the representation of the underlying model, that should be considered in the development of a system dynamic learning environment. The user should be supported in relating the behavior to the underlying structure. The learners must be provided with tools that aid their understanding of how the behavior arises based on the underlying structure. The integration process and the parts of the structure that significantly contribute to the behavior over time must also be represented. Support must be provided to understand how to control the behavior and how to use this understanding to develop policies. The user should be able to identify nonlinear relationships, map the systems' operating points with the purpose of identifying variations in dynamic sensitivity, and, as a result of her understanding, obtain the desired behavior of the system. All these components must be integrated into the learning environment.

In an earlier article, Davidsen, Spector and Milrad (1999) present a methodology for the design of system dynamic-based interactive learning environments. They use a *simple* dynamic system as an example, but argue that their method may also be used for the development of interactive learning environments for complex dynamic systems. They argue that the design should support multiple representations of the system, the understanding of the underlying complexity of the system, and the interrelations between the components of the system.

As part of model facilitated learning Davidsen et al. (1999) present six steps that should support the design of the learning environment. These steps are based on the principle of *graduated complexity* which, they argue, provide increasing transparency in the underlying model: *"in the development of ILEs, implementation of this principle leads to graduated transparency and support for learner-directed evaluation"* (Davidsen et al., 1999, p. 3, see also Spector & Davidsen, 1998; Spector & Davidsen, 1997b). (The issue of transparency will be discussed further in Section 5-5). Davidsen et al. particularly emphasize the need to divide the system into components that may be studied separately: *"We assume that it is not possible to understand a model as a whole unless we understand each of its components"* (Davidsen et al., 1999, p. 8). They also argue that synthesis is important.

Step 1 of graduated complexity is to *"challenge the learner to identify and characterize the reference mode of behavior…"* (p. 9). They argue that the learner should be supported to identify the target variables of the system and their behavior over time. Step 2 is to *"challenge the learner to identify the preference variables, each associated with a target variable and uncover the underlying preference structure"* (p. 10). By *preference* they refer to the preferred situation, that is, the situation existing if the problem was solved. Preference variables are linked to target variables and represent the preferred state of the target variable or the preferred behavior of this variable. Step 3 is to: *"challenge the learner to identify the structure underlying each the target variable and the associated preference structure"* (p. 11). The learner may in this way find the discrepancies between the preferred state of the system and the actual state of the system and the origins of these discrepancies.

Step 4 is to *"challenge the learner to halt at each stock, to investigate its dynamic characteristics, to infer the associated preference and to develop a management policy"* (p. 13). The learners should trace back through the structure, identify, and study the stocks of the system. Davidsen et al. (ibid) put particular emphasis on the integration process in the development of interactive learning environments and argue that this is the most important feature that contributes to the complexity of a system: *"The accumulation process is the core of the dynamic system and, at the same time, the most difficult process to understand"* (ibid. p. 13). The learner is challenged to develop policies with the purpose of providing a better understanding of the integration process. Step 5 to *"challenge the learner to encapsulate the unit of instruction and incorporate it into his body of knowledge"* (p. 15) means that *"the learner is assumed to recognize the response of this system to any typical input patterns of behavior"* (ibid. p. 15). It will thus no longer be necessary to analyze the structure in detail in order to understand the behavior of the model. Step 6 is to *"challenge the learner to diversify and generalize"* (p. 16). This means that the learner should be challenged in using the existing model for other problems or expanding on the existing problem.

Some of the steps above are reflected in the design of the prototypes (particularly in my use of multiple views and transparency). The steps, however, were not followed strictly in the design process, and we did not test learning outcomes (step 5) or generalizability (step 6).

With these broader issues of designing for interactive learning environments for complex dynamic systems I now turn to a number of specific design issues such as transparency (Section 5.4), interactivity (Section 5.5), and collaboration (Section 5.6). Methods for providing transparency is closely related to the development of the user interface. I discussed transparency in previous sections of this chapter and in Chapter 4. For the Two-Shower prototype, this was partly discussed in Chapter 4. The following section is a further discussion of transparency in the Two-Shower prototype, with more detail regarding the background to particular design decisions.

## 5.4 Transparency

*Transparency* is traditionally the degree to which the underlying simulation model is visible and accessible to the participants of the interactive learning environment. High opacity is the opposite of a high level of transparency and means that the underlying simulation model is less accessible to the participants. Davidsen et al. (1999) present the following definition of transparency:

> *"By transparency we refer to the notion that learners need to be able to see trough an interface to a high level representation (e.g., a causal loop diagram) through to deeper structures and causal mechanisms (eg., stock and flow diagrams)" (p. 6).*

I begin this section with a short discussion of the difference between correlational and causal descriptive models, as this is important as a basis for the subsequent discussion of transparency in system dynamic interactive learning environments. A discussion about various levels or degrees of transparency, and how transparency must be related to the purpose of the interactive learning environment, is followed by a short presentation of various methods for providing transparency. I end the section with a discussion of a different way of defining transparency and how it is utilized in the Two-Shower prototype.

### 5.4.1 The User Interface as a Filter to Glass Box Models

Modeling approaches may be classified in two categories based on whether they explicitly reveal the underlying structure that causes the behavior of the system or merely reveal the behavior (Barlas, 1996). Purely co-relational models are often called *black box* models because they only produce output in response to input and do not describe the causal structure between the variables in the system explicitly, i.e. the mechanism that produces the output. It is not the aim of these models to explain the underlying assumptions embedded in the structure. Because the structures of these models are not shown explicitly, it may be difficult to understand how input results in output.

In contrast, *causal descriptive* models exhibit the relationships between the variables explicitly and the purpose of such models is precisely to demonstrate how interactions between variables cause behavior. The concept *causal descriptive* implies that for each link between two variables an actual causal relationship exists in the real system, causing the value of one variable to influence the value of the other. These models are called *white box* or *glass box* models (ibid.). The purpose of such models is for the user to learn about the relationships between the variables and the behavior that they generate and thus develop a better understanding of the complex system. A system dynamic model such as, for example, the original Two-Shower model or the underlying model of the Quito prototype, is intended to be a *glass box* model. In many previous learning environments, however, the model has been hidden from the participants thus presenting it as a black box model (Spector & Davidsen, 1999).

As discussed in Chapter 2 it may be difficult to understand system dynamic models and it may be difficult to understand the relationship between structure and behavior even though all the information required to do so is available. In an interactive learning environment that is based on a system dynamic model a user interface functions as a filter of information made accessible to the participants. One of the fundamental aims of system dynamics is to aid the understanding of how the structure of a complex dynamic system generates the behavior. It is therefore necessary to have some kind of presentation of the underlying structure in order to link this to the resulting behavior. When developing an interactive learning environment for learning about complex dynamic systems, one must decide what information should be made accessible to the participants and at what level of detail. One of the purposes of creating the interactive learning environment is to provide a tool for understanding the system through controlling the information that is provided. In a system dynamic-based interactive learning environment, the degree to which the underlying simulation model is revealed to the participants is called *model transparency*.

### 5.4.2 Some Problems Regarding Transparency

There may be some problems regarding model transparency and the filtering of information through a user interface. Important information may be concealed. Größler, Maier and Milling (2000) ask the question of how it is possible for users to evaluate a model without seeing the model structure. Model evaluation is important in order for the users to be able to critique the model and to develop confidence or lack of confidence in the model. According to Größler et. al it is therefore necessary to provide access to the underlying structure of the model. For the VOCS project the problem of concealing information through lack of access to the underlying model particularly applies to the Quito prototype where it would be important for the users to assess the underlying assumptions of the model by comparing with their own assumptions about the problems in Quito. Transparency of the Quito prototype will be discussed in Chapter 10.

An additional objection against filtering out information about a model through a user interface is presented by Machuca (2000). He argues that the participants in a system dynamic interactive learning environment should be able to investigate the causes and effects of their decisions. Otherwise the participants will make decisions based on the symptoms of the systems and operate the system as they would if they were controlling a black box model. The result could be that the users learn about a system through trial and error.

### 5.4.3 Degrees of Transparency

In contrast to Machuca (2000), Alessi (2002) does not believe that transparent models necessarily are more effective tools for learning than opaque models. According to Alessi, transparency may be beneficial in some cases and in other cases not. He argues that models may be classified on a continuum from transparent to opaque and that the degree of transparency should depend on the model and the intended users.

Alessi argues against Machuca (2000) stating that learning with black-box learning environments is not necessarily based on trial and error. He points out that there are many learning environments where learning is intended to be based on discovery and where trial and error learning is not aspired. Alessi point to de Jong and van Joolingen (1998) who study what they call *scientific discovery learning* where in spite of using so called black box simulation models, the learner is encouraged and aided in a *"highly self-directed and constructivistic form of learning"* (de Jong & van Joolingen, 1998, p. 179). In their study, the main task of the participants was to infer the characteristics of the underlying simulation model by giving the model input and studying the resulting output.

Alessi (2002) agrees that transparency may foster a general understanding of the relationship between structure and behavior, but does not have an equal confidence that it will directly improve the decisions that are made in real organizations. Being able to make good decisions through an understanding of particular models would then indicate that the models had to be correct representations of the real system. As discussed in Chapter 2, according to system dynamics theory, a model is only a theory or simplification of the real system and not a correct representation of reality.

Learning a particular model therefore does not imply good decision-making. Alessi argues that the real system is not transparent and that *"...good decision-making will follow from learning to interact with and learn models in general, not from learning* particular *models"* (ibid., p. 4, emphasis in original).

### 5.4.4 Transparency Related to the Purpose of the Interactive Learning Environment

According to Alessi (2002) there are several factors that could guide the process of deciding the degree of transparency in a system dynamic interactive learning environment. He argues that the type of learning environment is fundamental in determining the level of transparency that it should have. He distinguishes between *expository* and *discovery* learning environments (ibid., p. 6). In expository learning environments the objective for the participants is to learn the model itself. These types of models are mostly used in higher education or by professionals such as in business or military education. The purpose of discovery learning environments is to learn problem solving and how to interact, use, and think in terms of models in general. This type of environment should give the participants the opportunity to investigate and explore the system in order to construct their own understanding of the system at hand and similar systems.

Alessi (ibid.) argues that there are several characteristics of the educational environment that should be identified in order to make decisions regarding transparency in the design process. He points out that the goal, what the participants should learn, is a critical factor for determining the degree of transparency. He lists several possible learning goals:

1. *The specific subject areas*
2. *Learning a process (such as research and exploration)*
3. *Learning specific content (such as principles of economic theory)*
4. *Learning skills (such as competitive behavior, cooperative behavior, or individual diagnostic skills)*
5. *Initial learning*
6. *Transfer of learning to a real-world work or other environment*
7. *Learning a particular system or problem structure and behavior*

8. *Learning generic skills such as thinking and problem solving*
9. *And in some cases non-learning goals, such as using a learning environment to facilitate model validation*

*(ibid., p. 6).*

Alessi (ibid.) refers to Repman, Lan, and Rooze (1995) and points out that also some characteristics of the learner should be considered when determining the degree of model transparency. These characteristics include, among others, age, previous knowledge and experience, cognitive load, and motivation.

Alessi (2002) further refers to Kashihara, Kinshuk, Oppermann, Rashev, and Simm (2000) and Kirschner (2002) and argues that according to cognitive load theory there may be a problem if there is too much information that should be handled by the participants. Alessi states that because system dynamics deal with complex dynamic systems a system dynamic interactive learning environment is likely to contain a high level of information that may be difficult for the participants to handle and interpret. Incorporating different levels of transparency in the interactive learning environment may aid the learners in handling the large amount of complex information.

According to Alessi (2002) it is also necessary to consider the educational philosophy underlying the interactive learning environment when deciding on the degree of transparency. An interactive learning environment that is based on an objectivist or behaviorist educational philosophy may indicate a different degree of transparency than a constructivist one. A more objectivist learning environment would emphasize expository learning while a constructivist environment would probably emphasize more exploratory and productive learning. Alessi further argues that interactive learning environments where the developers or educators have a more constructivist approach may be less transparent, because they want the learners themselves to construct ideas about the relationships of the underlying model and the problem it deals with. Educators with a more behaviorist or instructivist philosophy may therefore prefer learning environments with greater transparency.

The degree to which information should be available to the users is also important in deciding on the level of transparency. Alessi argues that *"the extent to which unknown or incomplete information is an essential part of the learning environment, and the extent to which risk or risk-taking is part of the learning environment"* (ibid., p. 8) is important for deciding the degree of transparency that should be implemented. Risk may indicate that some information is incomplete or inaccessible. This includes human behavior, which is of course impossible to predict.

Alessi (ibid.) presents several hypotheses of when greater or lesser transparency may be appropriate in a system dynamic interactive learning environment. (Not all details are discussed here as I have chosen to focus on those relevant for the Two-Shower prototype.) Important for our prototype development are comments on differences between users.

Expert or advanced learners may benefit from greater transparency because they may not be confused by the high level of detail as easily as inexperienced learners (ibid.). Experienced learners may be more used to handling a large amount of information and may also be more used to reading and interpreting diagrams and models. Learners with less experience may benefit from greater opacity because there may be too much information to handle simultaneously. This also involves the order and complexity in which the information is presented to the learner. This would be the case for the novice system dynamics students for which the Two-Shower prototype is intended. If the prototype was used later in the course it might need to be modified to allow greater transparency (a facilitator could also set specific tasks related to part of the interactive learning environment).

Alessi (ibid.) also suggests that learning environments that are designed to be used as part of a course in system dynamics should contain a higher degree of transparency since equations, stock and flow diagrams, and behavior graphs are part of what system dynamics students should learn. System dynamic skills involve being able to read, interpret, and understand such representational forms. Equations and stock and flow diagrams, however, are not part of the present version of the Two-Shower prototype. They are considered, however, for a future version of the prototype which will be discussed in Chapter 10.

### 5.4.5 Methods for Providing Transparency in System Dynamic Interactive Learning Environments

Alessi (2002) presents several methods for constructing interactive learning environments with varying degrees of transparency. He notes that *"...when discussing transparency we are referring to a design continuum with high transparency on one end and low transparency (or high opacity) on the opposite end"* (ibid., p. 14).

Davidsen and Spector argue that various parts of the learning environment may have greater or lesser transparency. Transparency of the underlying model should be provided gradually, allowing learners to spend time mastering parts of the model before moving on to other part of the model (Davidsen, 1994). Learners should also be provided an overview of the model before going into the specific details of parts of the model (Spector & Davidsen, 1997b).

Alessi (ibid.) distinguishes between three main categories of methods for providing model transparency: visual methods, verbal methods, and auditory methods. By visual methods he refers to pictorial representations such as stock and flow diagrams. Verbal methods primarily refer to textual descriptions of the system, while auditory methods may include speech and sound effects.

As discussed in Section 5-2 Spector, Davidsen and Milrad. (Davidsen et al., 1999; Milrad, 2006; Milrad et al., 2003; Spector & Davidsen, 1997b, 1998) present a methodology for developing system dynamic-based interactive learning environments, *model facilitated learning*. This methodology outlines a way of providing transparency to the interactive learning environment by gradually providing information about the underlying model. The learners are first provided with general information about the problem. They suggest that this is followed by information on three levels of increasing degrees of complexity: The learners are first offered information about the variables that they influence through their decisions in the system. The second level provides information about causal relationships of the underlying model, and the third level is a simplified representation of the structure in a stock and flow diagram.

Thus far, transparency has been discussed in the context of how underlying models may be represented in a user interface. The following two sections deal with additional concerns regarding transparency associated with the problem that the model should represent. The discussion also involves a consideration of transparency in the Two-Shower prototype.

### 5.4.6 Transparency and the Levels of Aggregation and Abstraction

In the previous subsections, definitions and discussions about transparency in the system dynamics literature have been considered. Transparency has generally been defined as the way in which the underlying simulation model is represented and made accessible through the user interface. In the system dynamics literature on transparency, there is generally a focus on representing the model in a number of different ways to support understanding. In my view, there is also a need for greater focus on the problem that the model represents. Greater transparency could also be related to the transparency of the problem. I will argue that presenting the details of the underlying model as part of the user interface is only one side of transparency and that a representation of details may be combined with other ways to make the *problem* that the model represents more transparent. The level of *transparency* is not necessarily equal to the level of *detail*. So far, the discussion of transparency has mainly been concerned with the level of detail versus the level of *aggregation*.

The concept of transparency could be extended and related to how developers seek to communicate about the problem through the user interface (and how they succeed at communicating about the problem through the user interface, which is not studied here). This communication does not necessarily involve only the discussion of the level of detail versus the level of aggregation. There could also be a discussion of *abstract* versus *concrete* representations of the problem that the underlying model represents. Concrete visualizations of the underlying model would involve relatively concrete visualizations of the *reality* that the model is intended to represent. The photos with changing facial expressions in the Two-Shower prototype may be an example of that. The aim would be to represent the model in a manner that the user, for example, could recognize from previous experience and, thereby, aid their understanding of the model. Stock and flow diagrams are much more abstract representations of what the model is intended to represent. It may seem like there is a tendency to believe that the abstract representations contribute to represent the model as it is. A stock and flow diagram is abstract, however, and is a particular way of representing the model. The representation of the details of the model as abstractions in a stock and flow diagram is of course important. I argue, however, that more concrete representations may be used to communicate other aspects of the underlying model and, thereby, the problem. Transparency can be implemented, in this way, on two different levels, but with different aspects of transparency (detail vs. aggregation and abstract vs. concrete). Figures 5-1 and 5-2 are examples of two ways of representing the Two-Shower model; one is relatively concrete but aggregate, while the other is abstract but detailed.


*Figure 5-1: Representation of the Two-Shower model at a relatively concrete but aggregate level of transparency.*

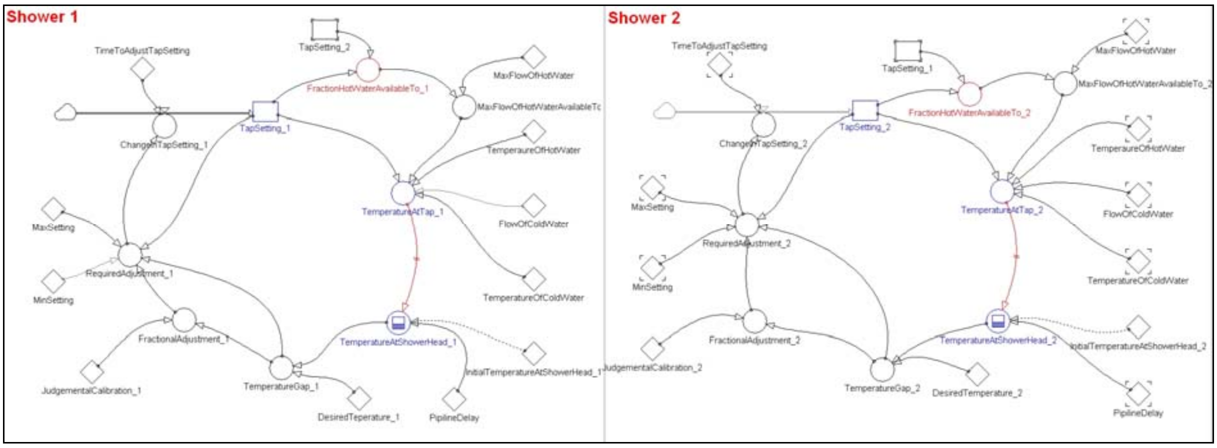


*Figure 5-2: Representation of the Two-Shower model at a relatively abstract but detailed level of transparency.*

There may be a difference in the adequate levels of abstract versus concrete and aggregation versus detail for different user groups, depending upon their previous experience. People with significant experience with stock and flow diagrams or equations may consider these representational forms as less abstract and may even prefer them to other forms of visualizations. System dynamicists may be more comfortable in dealing with stock and flow diagrams, even if more concrete representations of the problem that the model represents in the form of, for example, video or photos may remind them of the *reality* that the model is intended to represent. In the Quito prototype, which will be discussed in Chapter 10, video recordings are used as concrete representations of stakeholders' various points of view in Quito.

### 5.4.7 Transparency in the Two-Shower Prototype

In the current version of the Two-Shower prototype, we intended to implement a user interface with views that gradually went from a high level of concrete transparency with less focus on details, to views with increasingly abstract representations and a higher level of details regarding the underlying model. The goal was to provide a user interface that allowed for associations with what one would actually experience when dealing with the problem of two showers that shared a hot water tank and gradually introduce the complexity of the underlying model to explain the problem.

The details of how structure, behavior, and the relationship between structure and behavior were represented in the Two-Shower prototype were discussed thoroughly in Chapters 3 and 4. The prototype starts with a story with reference to the structure, moving on to representations of behavior in the Shower Room view. The participants can then go on to the Pipe System view where there is more emphasis on the underlying structure. When the participant triggers the story generator, explanations of the relationship between the structure and behavior are made. The intention behind this is in line with Spector and Davidsen's notion of graduated complexity that was discussed in Section 5.2 (see Spector & Davidsen, 1997a).

According to the discussion about transparency related to concrete versus abstract and detail versus aggregation, the Shower Room view contains a high level of concrete transparency and a relatively high level of aggregation. The goal was to provide visualizations and feedback that the participants could recognize from their own shower experience. The photos of a person with changing facial expressions, depending upon the temperature of the water, are relatively concrete visualizations.

The use of red and blue color to indicate hot and cold water are abstract representations, but build upon well-established cultural codes and is intended to make the underlying model more transparent. There is little information about the underlying structure and the causal relationships in the Shower Room view. The visualizations mainly refer to information one would normally have available when taking a shower (maybe with the exception of the thermometer), - that is a mixing battery and the changing temperatures of the water.

The Pipe System view contains a higher level of detail than the Shower Room view. Here there is a greater focus on representations of the structure. The representations are still more concrete than in a stock and flow diagram because they are simplified visualizations of real objects, such as pipes, mixing batteries, and again, photos or real persons who are taking a shower. The animations of the water that flows through the pipes are also intended to contribute to the level of concrete transparency. The Pipe System view is more abstract than the Shower Room view because it contains structural and behavioral information that would not be visible when taking a shower. It is, however, more transparent when it comes to the details about the structure of the underlying model and the behavior of the other shower.

The story generator with text that links structure and behavior represent more details of the underlying model through text explaining the effect of decisions that were previously made. This view is more abstract in that it facilitates jumps back and forth in time to enable the user to study the relationship between structure and behavior. In the two previous views, time is represented strictly chronologically.

The behavior graphs are visible in all the views discussed in this section. They represent a high level of detail regarding the behavior of the underlying model. In the current version of the Two-Shower prototype, there is, however, no access to the actual underlying system dynamic model as represented, for example, in Powersim, and the participants do not have access to the causal loop diagram or the stock and flow diagram. The Two-Shower prototype in this sense contains less transparency in terms of detail. Alessi (ibid.) argued, again as mentioned previously, how a model to be used for learning about system dynamics should also include causal loop diagrams and stock and flow diagrams as learning these representation forms is an important part of what a course in system dynamics is about. Spector and Davidsen have also done so in their learning environments (see for example Spector & Davidsen, 1997a). Because of time constraints, I have not included stock and flow diagrams in this version of the Two-Shower prototype. But I have suggested their inclusion in future versions of the prototype to allow access to representations of the model at an abstract, but highly detailed level (described in Chapter 9).

This section on transparency has dealt with how the underlying model has been represented in the user interface of the Two-Shower prototype. In the following section, I discuss the concept of interactivity and how it is possible for the participants to interact with the underlying model through the user interface.

## 5.5 Designing for Interactivity

In this section I discuss interactivity theories and how they have affected design decisions for the Two-Shower prototype. I begin the section with a short consideration of some of Manovich's (2001) critique of the use of the term *interactive* followed by a classification of interactive systems by Jensen (2000). I then turn to Alessi and Trollip (2001) and issues of control, input, and feedback, before I end the section with a note on interactivity and game structures.

### 5.5.1 Some Questions Regarding the Concept of *Interactivity*

According to Manovich (2001), *interactivity* is too general a term to use when describing a feature of new media. All computers require the user to interact with it in some manner in order to make it perform. As Manovich argues: *"once an object is represented in a computer, it automatically becomes interactive. Therefore, to call computer media 'interactive' is meaningless – it simply means stating the most basic fact about computers"* (ibid., p.55).

Rather than using the term *interactivity,* Manovich classifies it and uses other concepts such as "*menu-based interactivity, scalability*, *simulation*, *image-interface*, *and image instrument to describe different kinds of interactive structures and operations"* (ibid., p. 56). He also argues that all kinds of media and art, classical as well as new, are interactive in that the audience must interact and fill in missing information themselves. According to Manovich, there is a danger of confusing *physical* interaction between the user and the media with the *psychological* interaction that occurs when interpreting, filling in, and creating an internal mode of the media and the process. For Manovich, this is what interactivity is about, not the physical movements in interacting with the machine e.g. the physical process or option of pressing a button on a computer to produce some kind of response.

Designing for interactivity, how the participants are able to control the underlying model, and what feedback should be provided was an important issue in the development of the Two-shower prototype. The physical aspects of interaction are related to issues already discussed such as changes in tap setting, and the ability to change views. The psychological aspects of interactivity related to the issues of narrative (the hotel setting) and to emotional issues such as heat and cold. The red and blue colors, for example, are intended to be interpreted based on previous experience with the colors blue and red as metaphors of cold and hot temperatures.

### 5.5.2 Some Characteristics of Interactive Systems

Jensen (2000) also attempts to break down the term *interactive*. He suggests that interactivity is a measure of a medium's or media production's potential to enable the user's influence on the medium's communicated content or form. The degree of interactivity must be defined on a continuum, not as some criteria a medium needs to fulfill in order to be defined as interactive or not. Jensen (ibid., pp. 66-67) lists four concepts that may be used to describe the degree of interactivity:

- *Transmitative interactivity*, which is the degree of the user's potential to choose from a continuous stream of information transmissions produced by the sender (such as text TV or multiple TV channels).

- *Consulting interactivity*, which is the degree to which it is possible for the user to choose on request among ready-made information flows with a channel of response (such as CD-ROM encyclopedias, digital television).

- *Conversational interactivity*, which is the degree to which the medium enables the user to produce her own information through input that is made available to others (such as e-mail or video conference systems).

- *Registering interactivity*, which is a measure for the medium's ability to register input from the user and adapt to and respond to the user's actions and needs (such as for example intelligent agents).

These four concepts are all used to describe characteristics of the medium and the opportunities that the users have in controlling the content and form.

Jensen (2000) places dynamic simulations for learning in the high end of the interactivity scale compared to other media for learning, such as simple programmed questions and answers or browsing through electronic material. He argues that dynamic simulations for learning involves a mix of *consulting interactivity* and *registering interactivity* because the system to some extent adapts to the input that is given by the user and the user usually is able to explore parts of the system by herself, searching for useful information.

The degree of interactivity will also vary within simulation-based interactive learning environments. The amount and complexity of information that is available to the user may differ, as well as the time at which it is available. This determines the degree of *consulting interactivity*.

The Two-Shower prototype has a high degree of *registering interactivity* in that user actions are used to control the underlying simulation model and, thereby, the behavior of the model. The behavior of the model is used as output and the feedback provided changes accordingly. The story generator is implemented as an agent that gives the participant feedback about which points in time to consider based on input previously provided by the participant. The Two-Shower prototype has a certain degree of *consulting interactivity* in that the participants can switch between the various views in the interactive learning environment and, thereby, choose the kind of feedback they want and which type of information they wish to examine.

With these first comments on differing types of interactivity, I turn now to other concepts and forms of user activity related to the design of interactive learning environments, specifically control, input, and feedback.

### 5.5.3 Control, Input, and Feedback

Alessi and Trollip (2001) use the term *locus of control* to distinguish between programs that are controlled by a user and programs that progress without user input. *"Three considerations concern the design of learner controls: what and how much the learner can control, the method of control, and the mode of control"* (ibid., p. 51, emphasized in original). They also present some general principles that must be made in relation to how much control should be given to the learners, but emphasize that there are no correct answers and that this must be evaluated in each case. Their advice is that developers should have the intended users in mind during the development process and determine the learner control based on this. Alessi and Trollip further discuss how there are different methods and modes that could be used in order to let the users have control over a program. (Examples of methods of control are different types of menus, buttons, and typed commands.) They again emphasize the importance of knowing the users of the program and using methods of control that are suitable to the task and the users. *Modes of control* refer to controlling the program through the mouse, the keyboard, or by speech.

As described in the previous chapters, input to the Two-Shower prototype is provided by the participants as they change the tap settings of their showers. This is done by turning the graphic, animated representation of the mixing battery. The simulation control buttons also represent *input* in that the participants may control whether the simulation should run, stop, or pause. The participants may also change the view that should be present. This was described in detail in Chapters 3 and 4.

Part of what makes a system interactive is that when a user makes a change in the system through input, *feedback* is provided by the system. According to Spector and Davidsen (1999) informative feedback is a necessary prerequisite for engaging the participants. Alessi and Trollip (ibid.) distinguish between *artificial* and *natural* feedback of educational simulations. Natural feedback is feedback that would also be provided by the real system in a real situation. Artificial feedback is feedback that would not be present in the real situation such as for example text boxes with textual information about the current situation. According to Alessi and Trollip (ibid.) artificial feedback could be used in the beginning when the users may need some additional feedback in order to develop an understanding of the system. There could then be a gradual transition to presenting more and more natural feedback so that the user can experience a situation that is closer to dealing with the real system.

In the Two-Shower prototype, however, the Shower Room view, which may be said to contain what is closest to *natural* feedback, is presented before the more Pipe System view with more *artificial* feedback. The photos of the characters with changing facial expressions in the Shower Room view could represent natural feedback, as these are the facial expressions of a real person and because it is likely that also the participants would have similar facial expressions if they were exposed to hot or cold water. Most of the feedback provided through the visualizations in the Two-Shower prototype is, however, artificial. Hot and cold water is visualized, as discussed previously, through red and blue colors while in real life it is colorless. In addition, the yellow warning pop-up that tells the participant that the water is too cold or too hot evidently does not exist in a real shower. The intention of the Two-Shower prototype is not to provide as much natural feedback as possible, but to be used as support for understanding the underlying model. It could, however be interesting to let the participants go back to the Shower Room view after having used the Pipe System view because it provides less artificial feedback. It may be that the amount of feedback in the Pipe System view is superfluous once they have developed an understanding of the underlying simulation model. Time restricted, however, such testing.

### 5.5.4 Interactivity and Game Structures

I also found game structures useful in designing the prototype. I discussed games and narrativity in Chapter 3, but return to it here when addressing design decisions related to learning theory. By including characteristics from computer games in an interactive learning environment, the intention is partly to provide means for engagement. At the same time the intention is to support understanding. Größler (1997; 1998), for example, presents a study of participants in a business simulator where they are given two goals: to obtain as good results as possible and to learn as much as possible. There is, however, as Alessi (2002) note important, partly incompatible goals between a computer game and an interactive learning environment which must be considered in the development process. He argues that in order to learn as much as possible the user would have to make both decisions that lead to good performance and *incorrect* decisions that lead to *bad* performance. The problem is that in order to *win the game* you only have to make good decisions. Alessi argues that if learning is the objective it is better to call it a *simulation* rather than a *simulation game* and to encourage exploration of both good and bad decisions. Similarly, if the participants are operating under the pressure of time, there may not be time to reason about and discuss the underlying model before decisions are being made. To jump ahead a bit, these competing goals became clear when Test 1 was carried out and some of the participants overlooked important information in the user interface as the desire to win took over (Frotjold, 2005). Test 1 will be discussed in Chapter 7.

Potential problems of competition between the participants also apply to the two-user version of the Two-Shower prototype where the objective is not to obtain the highest score or to stabilize the system in the shortest amount of time. The participants are not supposed to compete against time or each other, but should aim at controlling and obtaining an understanding of the system. When running the Two-Shower prototype, I found it important, therefore, to inform the participants explicitly of the objective of the session. They could be told, for example, that the objective is for them to find out how the underlying simulation model functions and not to compete to be the first one to obtain a comfortable temperature or manage to maintain a comfortable temperature for the longest period. Competition in a two-user version of the Two-Shower prototype could lead the system to oscillate and none of the participants would obtain a stable, comfortable temperature. In addition to instructions provided by facilitators, the problem of potential competition is partly overcome by making visualizations of the conditions of the other participant available on the screen of each participant. This is further discussed in the following section.

## 5.6 The Two-Shower Prototype – a Prototype of a Collaborative Interactive Learning Environment

As discussed in Section 5.3, according to constructivist theory, learning is seen as a social process and constructivist learning environments should therefore support collaboration (Jonassen et al., 2000). In this section, I consider collaborative aspects of learning and how the Two-Shower prototype was designed with the purpose of supporting collaboration between the participants in understanding the underlying model. I will first discuss some issues that led to the decision of developing the Two-Shower prototype as a two-user rather than single-user system. There is then a brief introduction to the field of computer support for collaboration and learning (CSCL). Theories of grounding are central in CSCL and, after an introduction to these theories, the chapter ends with a discussion of how the Two-Shower prototype is developed with the purpose of supporting grounding and the development of a mutual understanding between the participants.

### 5.6.1 Two-User Rather than Single-User

As mentioned in Chapter 1, an earlier version of the Two-Shower prototype was a single-user version (see for example Frotjold, 2005). In the single-user version, the tap setting of one shower is controlled by a participant while the other is controlled by the simulation model. There are, however, some limitations we felt with the implementation of the model as a single-user system. In order to obtain a stable, comfortable temperature, the actions in both showers should be coordinated. The simulated *participant* in the other shower, however, does not make optimal decisions in coordination with the real participant. A suggestion could be to construct a simulated *participant* that *learns* from its mistakes and gradually becomes *aware* of the other shower, *taking* the decisions and conditions of the real user into account when generating a change in tap setting. There is, of course, no way the *real* participant and the *simulated* participant can communicate to make sure they both develop an *understanding* of the system and to coordinate their decisions. The potential advantages of the communication process between two real participants will not be utilized.

Some form of communication is, of course, a prerequisite for collaboration. Consequently, a collaborative prototype, with the intention of supporting two participants in working together to try to understand the model and obtain a comfortable temperature was developed. An important issue in the design of the Two-Shower prototype was thus for the user interface to support communication between the participants. The development was partly based, as mentioned above, on theories of Computer Support for collaboration and Learning (CSCL). A brief outline of CSCL principles will be given in the next section.

### 5.6.2 Computer Support for Collaboration and Learning (CSCL)

CSCL is a research field that emerged in the end of the 1980s. It is a branch of instructional technology which *"...focuses on the use of technology as a mediational tool within collaborative methods of instruction"* (Koschmann, 1996b, p. 2). Lipponen (2002) states that the focus of CSCL is on *"...how collaborative learning supported by technology can enhance peer interaction and work in groups, and how collaboration and technology facilitate sharing and distributing of knowledge and expertise between community members* (p. 72). CSCL has its theoretical foundation in social constructivism, socio-cultural theories, and situated theories (Koschmann, 1996b).

In CSCL, learning is believed to be *“reflected in the language of learners”* (Koschmann, 1996b, p. 14). Empirical studies of learning therefore, to a large degree, focus on observation of participants in a learning environment rather than, for example, on cognitive representations of knowledge. Stahl (2002) argues that observations highlight the collaborative process:

> *“Because shared meaning exists in the observable world and collaborative meaning-making necessarily unfolds there, CSCL researchers can make learning visible by interpreting these meanings and practices. Collaborators must make their understandings of what they say, hear and see public in order for their partners to work together with them. Of course this does not mean that everything is made explicit* (p. 531).

He also points to the need to consider the artifacts created: *”the clues for making visible the learning that took place during the collaboration can generally be found in the externalizations and artifacts created then”* (ibid.).

Hutchins (1995), who may be placed within the situated perspective on which CSCL is also based, suggests that groups may have cognitive properties that are different from that of the individuals of which a group is made. The differences of the cognitive abilities may then depend entirely on the social organization of distributed cognition, not on differences in cognitive abilities of the group members. Collaboration may thus bring about other characteristics than that of the individual working alone.

When learning about complex ideas and concepts, collaboration has certain advantages over individual learning. Complex concepts may be simplified when explored and explained by an individual, and there are often misperceptions when multiple processes occur concurrently (Feltovich et al., 1996; Koschmann et al., 1996).

Based on this, I sought to implement support for collaboration in the Two-Shower prototype. The theories of grounding and shared understanding within the CSCL literature were of particular interest.

### 5.6.3 Grounding

An important aspect of collaboration is that actors develop a shared understanding of a system or topic. This is often referred to as *grounding*. Grounding is the process by which participants in collaborative learning construct and maintain some degree of mutual understanding or common ground (Baker et al., 1999). Common ground does not refer to the internal knowledge that the participants have in common, but to something that is actively negotiated and re-negotiated during the communication process (Arnseth & Solheim, 2002). In a more recent paper Koschmann and LeBaron (2003) critique the constructs of common ground and grounding. They argue that in the Computer Supported Cooperative Work (CSCW)[22] literature *common ground* has tended to be treated as an abstract but measurable entity. Analyzing a fragment of interaction from the operating room at a teaching hospital, Koschmann and LeBaron argue that Clark's contribution theory of discourse, which grounding theory arises from, fails to explain important aspects of the collaboration process. When interaction is analyzed as a goal-oriented process of establishing common ground, researchers may ignore the ambiguities and dynamics of everyday conversation and collaboration (ibid.).

In this thesis, however, I continue to use the terms grounding and common ground. By applying these terms, or the term *shared understanding*, I do not imply that the participants are expected to acquire some mutual *mental model* of the system that they will be able to express in words. Rather, the prototype is intended to support them together in reaching a sufficient understanding to be able to control the system

[22] CSCW is a research field which studies computer supported collaborative work that is supported by groupware. CSCL can be seen as the *younger sibling* of CSCW (Lipponen, 2002).

and communicate about its structural and behavioral characteristics. The medium that is used in collaborative learning must provide tools that support this communication process, such as chat-tools embedded in the software, or allow for spoken utterances among participants located in the same room.

Dyck, Pinelle, Brown, and Gutwin (2003) studied the interfaces of multi-player games in order to investigate whether innovations that were made in gaming interfaces may be used in the interfaces for computer supported collaborative work. They found that one of the key issues were that the interfaces were constructed in such a way that people can learn by observing others. Most other collaborative computer software contains these features, but lack a straightforward manner in which these observations can take place. In computer games, the users are provided real-time awareness and embodiment of the others' actions, which supports the users in comprehending the decisions and actions that are made by the others. Many games personify the users' actions by use of personal avatars that act within the game environment. According to Dyck et al. (2003), it is not the existence of the avatar itself that makes learning possible, but the mediation of embodiment, awareness, and task-based information. The avatar is then used as a means communicate important information to other users. Dyck et al. suggest that means for observational learning can be transferred to other types of applications.

All media used in collaborative learning impose constraints on the grounding process. Grounding through oral and written linguistic utterances may be limited by the participants' abilities to communicate their ideas and conceptions of the system and their ability to interpret the utterances of other participants. The grounding process may be improved by a tool that enriches the linguistic utterances. The Two-Shower prototype is intended to support grounding by providing visualizations of the conditions that each participant *experiences* during their use of the collaborative prototype. Visualization of complex, dynamic systems is thus intended to be used as a tool to support the grounding process and aid the participants' development of an understanding of the underlying simulation model (although to what extent this understanding actually takes place, is not yet tested).

### 5.6.4 Visualization to Support Grounding in the Two-Shower Prototype

To overcome conflicts of interest when it comes to shared resources, communication and cooperation is required. It is therefore necessary, to the extent that it is possible, to make decision-makers aware of each other, of the effects that their decisions have on others, and of the effect that the decisions of others have on themselves.

The interface of the Two-Shower prototype is constructed with the purpose of providing means for grounding. As discussed previously, grounding is the process of developing some kind of joint understanding of the system, which helps them to perform the actions required to control the system effectively.

The Two-Shower prototype is intended to support grounding through visualization of the participants' actions and the conditions of the other participant. The participant can then, at any time, monitor the conditions and decisions of the other participant and base her decisions on the other participants' actions.

The 2003 version of the Shower Room view had a Keyhole button which provided access to a view of the conditions of the other participant (Viste & Skartveit, 2003, 2004). In this view, only the behavioral characteristics of the system were displayed and the participant had only access to the slideshow of the other character in the shower (see figure 5-3). The photos and the background color of the window changed based on the temperature of the other shower. The intention was that this should support grounding between the participants. This function was removed from the current version of the prototype due to programming difficulties, but may be reintroduced in a future version.

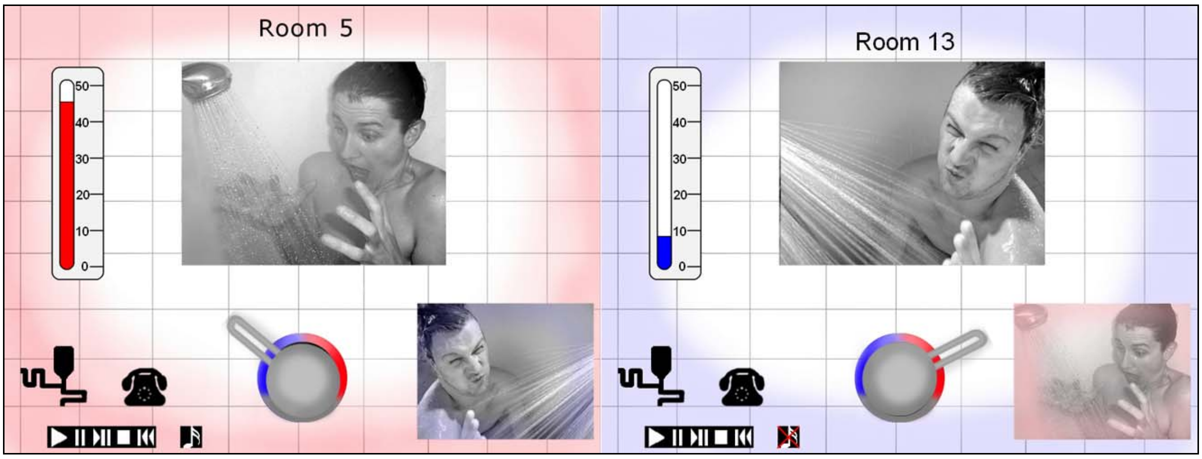


*Figure 5-3: An earlier version of the Shower Room view where it was possible to press a Keyhole button and observe the behavior of the other shower in a small window.*

As described previously, the structure of the underlying simulation model is revealed in the Pipe System view. Here, in addition to the conditions of the other participant, the first participant can see the decisions of the other participant as well, through a miniature visualization of the other participant's tap setting (see figure 5-4). As mentioned before, the effects of the other participant's decisions are also shown in this view. The intention is that the participants should use these visualizations to monitor the conditions of the other participant and to communicate verbally about the model and thereby be supported in the process of grounding.

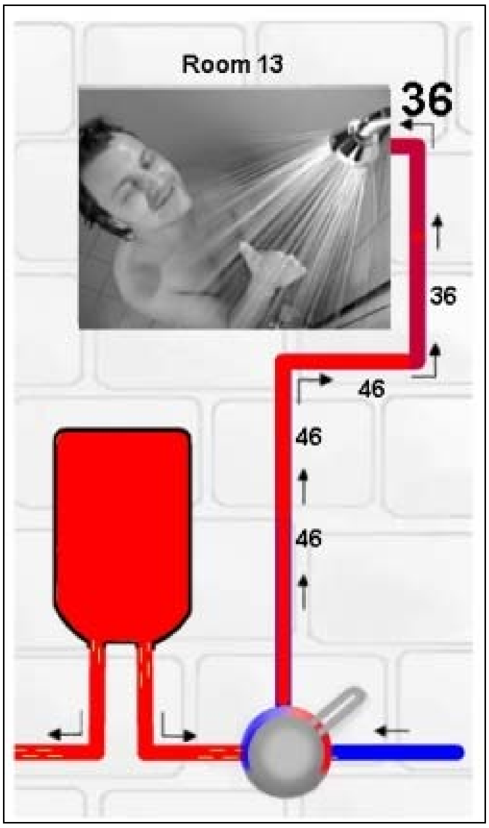


*Figure 5-4: Visualization of the situation and actions of the other participant in the Pipe System view.*

In this chapter, I have discussed how the Two-Shower prototype was developed with the purpose of supporting understanding. This was the last of three chapters on design. In the following three chapters, I turn to evaluation and begin with the first formative evaluation that was performed in the prototype by Frotjold (2005).

# 6. Test 1 of the Two-Shower Prototype

## 6.1 Introduction

The focus of Chapters 3-5 was the design of the prototypes. In Chapters 6-8 I discuss the issue of evaluation and, more specifically, usability testing. This chapter is mainly based on Frotjold (2005) and is a presentation of *Test 1* of the Two-Shower prototype. The design, delivery, and analysis of results of the evaluation test described in this chapter were conducted by Laila Frotjold under Associate Professor Weiqin Chen's supervision. The design and results of the test are summarized here. Although I was not involved in the development and implementation of this test, I include a description of this test because it was the first test of the prototype developed by the project group and because the results were important for the further development of the prototype and Test 2. While I was not involved in the design of Test 1, I do not include in this chapter a lengthy discussion of the theories behind her design. I do, however, include, in Chapter 7, an overview of theories on evaluation techniques that constituted the basis the test carried out by myself (Test 2).

The focus of Test 1 was to assess the functionality of the system and to identify design problems. There was a particular focus on whether the text generator would add meaning to the user's experience and whether the users were better able to understand the dynamics of the system when using the text generator (ibid.). In addition, there was a focus on whether the users were capable of using the text generator in the intended way. The test was performed as two subtests[23] where the participants used different versions of the prototype and the results were analyzed and compared based on these different versions of the prototype. Details of the test are

[23] Frotjold here uses the term *experiment*, not *subtest*. I have chosen to refer to the two parts of Test 1 as *subtests* as *experiments* would refer to tests performed for example to establish causal relationships through use of control groups, larger samples, and randomized treatments.

provided in Section 6.2. In the final section of this chapter, Section 6.3, I briefly discuss lessons learnt from Test 1 and the need for changes to the methods of testing.

## 6.2 Test 1: Frotjold's Evaluation of the Two-Shower Prototype

Frotjold's (2005) evaluation, Test 1, is the first formative evaluation test of the Two-Shower prototype and was designed as two subtests. This evaluation was used to point out some problems of the user interface and form the basis for further development.

Frotjold had two main goals for the test:

- *Assessing the extent of the functionality: I wanted to find out whether the addition of generated explanatory text to a graphical user interface would add meaning to the user's experience. Was the user able to understand the dynamics of the system more easily provided the explanation?*
- *Identifying problems with the system: If the users seemed to be using the explanation device in a proper way, yet seemed to be disturbed or confused by its presence, a redesign of the visual aspects of the tool would have to be conducted in order to perform the functionality test under adequate conditions*

(ibid. p. 51).

The aim of Test 1 was to evaluate the functionality of an earlier version of the Two-Shower prototype (a 2005 version) and to identify problems with the associated user interface. The main focus was on the evaluation of the textual explanations that are generated by the story generator and identify problems with the presentation of these.

Test 1 was conducted on a single-user version where one participant is replaced by a simulated participant. The intention was to give the participants an equal starting point (ibid.). The second reason to use a single-user system was that the results from a study with two participants would have been more difficult to interpret. Frotjold argues that *"different subjects would face different levels of difficulty, depending on the other user's comprehension"* (ibid. p. 56).

I describe the version of the Two-Shower prototype that was used for Test 1 in greater detail below so as to provide a more detailed description of the test and its results.

### 6.2.1 The Single-User Version used in Test 1

The version of the prototype that was utilized in Test 1 was deployed in 2005 (as mentioned in Chapter 1 the 2006 version, which is the latest version, is mainly described throughout this thesis). The starting point for the 2005 version of the prototype was a prototype that had been developed in 2002-2003 (see for example Viste & Skartveit, 2003). The main additions of the user interface are listed in figure 6-1.

The differences between the prototypes related primarily to design issues such as the inclusion of the text generator and single or dual users. Frotjold modified the prototype as described in Chapter 1, from a dual-user to a single user system. Moreover, the slider was further developed and implemented and the text generator was developed. At that time, the graphical animations of the story generator had not yet been implemented (the graphics representing the structure were there, but they did not change according to the values of the underlying simulation model when the participants went back in time to study previous behavior). The story generator, therefore, only consisted of the text generator and the time slider. The text generator was used both to create text commenting what had happened in the prototype over time, as it does in the current version, and to generate the *basic explanation*, that was displayed before prototype runs, serving as an introduction to the underling simulation model (see figure 6-2). The markers on the timeline with indications about which points in time important events had occurred had also not yet been implemented.

| |
|---|
| Basic textual explanation of the characteristics of the system was generated and displayed in a message box before the start of the simulation and could be shown again during the simulation run if the users pressed a button to repeat the basic explanation. |
| Time line |
| Text generator, which displayed textual explanations when the user held the mouse over the time line in places where important events had occurred during the simulation run. |
| Text box in lower left hand corner with:<br>- A message to the user that one could pause the system and start the text generator<br>- The explanations of the text generator<br>- A basic explanation of the system when the user pressed a button to repeat the basic explanation. |
| Important points in time on timeline (the important points were not indicated graphically on the timeline, only as changes in text in the text box. |

*Figure 6-1: Changes made to the user interface of the Two-Shower prototype before Test 1.*

In the 2005 version of the Two-Shower prototype, the basic explanation was generated and displayed in the Introduction text box rather than the introductory story (as described in Section 4.3.1). In the 2006 version the basic explanation is displayed when the participant presses the *Hints* button (see Section 4.3.2).[24] The 2005 version used for Test 1 also did not include the behavior graphs that were implemented in the

[24] As noted in Chapter 1 the version that is primarily described in this thesis is the latest version with additions in 2006 by myself with support from external programmer Olve Sæther Hansen.

2006 version of the prototype (see Section 4.3.3 or figure 6-3 for a representation of the Pipe System view in the 2005 version).

*You are involved in a system in which* your actions lead to consequences for you, in several ways, some time after you do it. You should take into account that the same action may lead to different consequences, and to different extents, at different times. *This can make your understanding of the system erroneous. In order for you to understand the system, we will take a closer look at parts of it.*

*When* your hot water setting *changes it takes time before* the temperature at your showerhead *changes accordingly. This delay arises because of* the time it takes the water to flow from the mixing battery to the showerhead. *You can therefore not expect that a change in* your hot water setting *will give immediate results* for you.

***In addition to this,*** the temperature in your mixing battery *will change. This depends on both* your hot water setting *and* the fraction of hot water available to you. The fraction of hot water available to you *is in turn dependent on both* your hot water setting *and* the other user's hot water setting. *It is therefore not certain that the same setting for* your hot water setting *will give the same effect* for you *every time.*

***But this is not all.*** *What you do will also bring consequences to the other user. When* your hot water setting *changes, it takes time before* the temperature in the other user's shower *changes accordingly. This delay arises because of* the time it takes the water to flow from the mixing battery to the showerhead. *You can therefore not expect that a change in* your hot water setting *will give immediate results* for the other user.

***Further,*** the temperature in the other user's mixing battery changes. *This depends on both* the other user's hot water setting *and* the fraction of hot water available to the other user. The fraction of hot water available to the other user *is in turn dependent on both* the other user's hot water setting *and* your hot water setting. *You can therefore not expect that a change in* your hot water setting *will give the same effect* for you *every time.*

*Figure 6-2: The basic explanation (the text used in the 2005 version of the Two-Shower prototype was in Norwegian).*

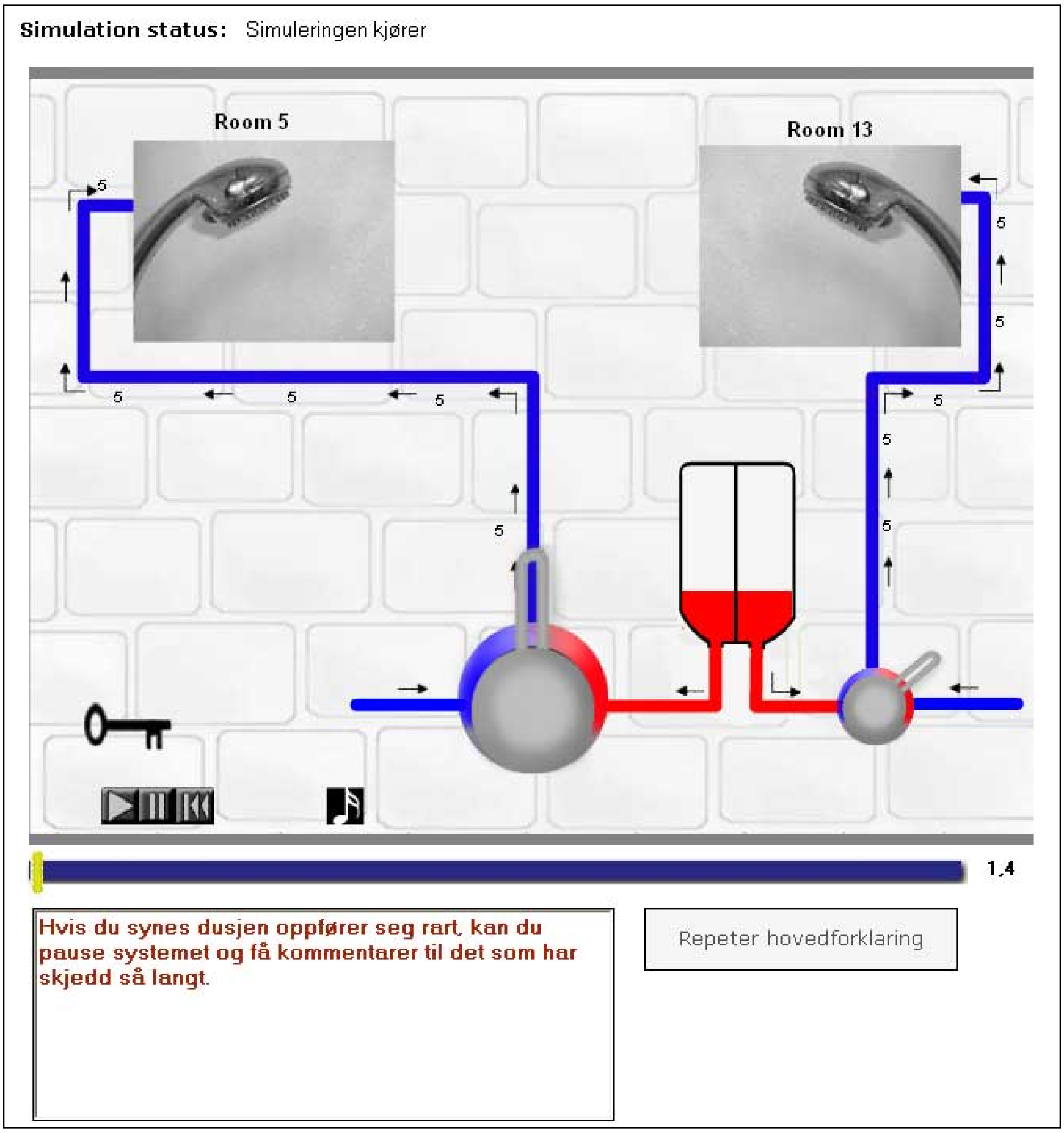


*Figure 6-3: The version of the Pipe System view of the Two-Shower prototype that was used in Test 1 (this was a Norwegian version of the prototype).*

### 6.2.2 Data Collection

In Test 1, Frotjold gave a PowerPoint presentation of the goals of the test, explanation of some terminology, and a relatively detailed explanation of the user interface and its features before running the prototype. The participants were informed that their main objective was to understand the system and that they were to explain how it worked after the simulation run (Frotjold, 2005). They were also told that one of the objectives was to obtain a comfortable shower temperature. The presentation revealed that the slide shows, the thermometer, and the numbers next to the pipes indicated the water temperature. The participants were also told that the Pipe System view was a representation of both showers and the hot water tank that they shared. The various buttons and the tap setting were also pointed out.

A combination of methods was used to gather data during the sessions: observation, interviews, system logs of stabilization times for the simulation model, and user input (through changes in tap setting). After the presentation of the system the participants were given a pre-test. They were asked the following questions:

> *What do you think will happen when you increase the tap setting?*
>
> *What do you think will happen when you decrease the tap setting?*
>
> *(ibid., p. 57)*

These pre-interview questions were given with the intention to test the participants' previous knowledge of the problem. When the prototype was running, computer logs of participant actions were recorded. Data regarding stabilization times were gathered in order to investigate how much time the participants spent before stabilizing the system. Each time the participants paused the system or changed their tap settings, the information was recorded in computer logs. Audio recordings were made from the sessions for analysis purposes. Notes were also taken to capture information that was not recorded otherwise, such as statements made when the audio recording device was switched off.

Post-test interviews were also conducted with the purpose of finding out what the participants thought about the user interface. The purpose of the interviews was also to study the participants' comprehension of the system and to clarify their observed behavior during the simulation runs. The questions asked were: *"How will you explain the temperature oscillations that you experienced?"* (Frotjold, 2005, p. 59) and *"can you explain how the system works as a whole*?" (ibid). These questions were supplemented by follow-up questions.

### 6.2.3 Two Between-Groups Tests

According to Frotjold (2005) Test 1 was based on the *between groups experiment* method as described by Dix et al. (1998). When using this method, the participants are divided into groups who are exposed to different versions of the system. The participants experience equal conditions with an exception of the value of an independent variable that is different for each group. In Test 1 the participants were divided into four test groups. One test group had neither access to the Pipe System view where the structure of the system is described graphically nor to the textual explanation of the system structure that was presented before the simulation was started. Test group number two only had access to the basic explanation, not the Pipe System view (the participants could also choose to read the basic explanation again while the simulation was running), while test group three had access to the Pipe System view and not the basic explanation. Test group number four had access to both, the basic explanation and the Pipe System view. All groups had access to the text that was generated by the text generator.

Frotjold conducted two subtests (here referred to as Test 1a and Test 1b) with two different user-groups. The participants of Test 1a consisted of Master students at the University of Bergen. The first subtest, Test 1a, had four test persons and there was , consequently, only one person in each test group, that is, one person who had access to each version of the user interface. Test 1b consisted of thirteen junior high school students at a school near Bergen. In this subtest there were three or four participants in each test group.

### 6.2.4 Test 1a

Frotjold (2005) analyzed the simulation results and comments made during the interviews and came to three main conclusions regarding the Two-Shower prototype based on the results from Test 1a:

1. The participants had problems understanding some of the graphics of the user interface.

2. The textual basic explanation, provided as an introduction, aided the participants in understanding the underlying model.
3. The participants did not use the continuous textual explanations (the text part of the story generator).

Frotjold concluded that only the participants who had access to the textual explanation managed to consider both the delay and the nonlinearity during the simulation run. She argued that the participants needed the text in order to understand the system properly. She noted that all participants were capable of reasoning about the system as the simulation was running, while only the participants who had the textual explanation managed to take the delay and nonlinearity into account during the simulation run (see Frotjold, 2005, p. 61).

The participants who had access to the basic explanation, -a description of the model structure, did not use the text that was generated by the story generator based on changes in the model during the simulation run. Some of these participants pointed out that this feature of the user interface should have been be emphasized more strongly. Based on this feedback, Frotjold improved the prototype before her second subtest (Test 1b) by including the yellow warning pop-up, which reminded the participants that they could press the pause button to receive an explanation about what had happened, as described in Section 3.4. In addition, she implemented a text that reminds the participants that they may pause the simulation to receive an explanation of what is happening. This text turns red and bold when the temperature is too cold or too hot. The pause button was also changed so that it turns red when there is an uncomfortable shower temperature. The introductory text (the basic explanation) was divided up to make it more readable.

### 6.2.5 Test 1b

As mentioned above, the participants in Test 1b consisted of thirteen junior high school students, divided into four test groups. Frotjold (2005) further categorized the participants as *strong*, *average*, and *weak* students.[25]

The participants' answers to the questions asked in the pretest were equal for both Test 1a and 1b. All participants were aware of the fact that the temperature of the water would rise or fall, not referring to the delay or the nonlinearity in the system.

In the post interviews of Test 1b there was a problem in making the participants express themselves about the system. Frotjold (2005) states: *"During the interview it became apparent that the students were not used to expressing thoughts and reflections,…. Several additional questions and re-phrasings were needed in order to elicit their thoughts"* (p. 63).

From this subtest (subtest 1b) Frotjold concluded that participants who had access to the basic explanation seemed to have a better understanding of the system than those who did not. All participants who had access to this explanation also managed to stabilize the system. Frotjold also found that there was a large difference in performance between the strong and weak students.

None of the participants in the subtest used the explanations that were generated by the story generator. Some of the participants, who managed to stabilize the system, claimed that they received sufficient information from the basic explanation.

The results from the subtest showed that the participants had problems understanding the graphics in the Pipe System view. Frotjold states: *"…they appeared to be*

[25] Frotjold describes the categorization process as follows: *"The teachers had picked students to be tested in advance, and the result was a very heterogeneous group, with some strong students, other weak, and some average students. The term 'weak students' refers to students whose reading and expressing capabilities were lower than the average presented to me. The term 'strong students' refers to students whose reading and expressing capabilities were higher than the average presented to me"* (Frotjold, 2005, p. 63). A classification of the participants in this manner could be questionable, but I remind the readers that the test was only part of Frotjold's Master thesis and the main goal was to pick up major difficulties with the text generator.

*overwhelmed by it, and only one student seemed to find it useful in order to understand part of the system"* (p. 69).

### 6.2.6 Conclusions from Test 1

One of the main conclusions from Frotjold's study was that the participants generally had a problem understanding the Pipe System view and she concludes that there is a problem with the design of the graphics in general (Frotjold, 2005). The respondents also pointed out that there was an incoherence between the Shower Room view and the Pipe System view.

One of the intentions of using a model of two showers that share a hot water tank was that the participants may have experienced a similar system in real life. Their previous experience was believed to enable them to better deal with the problems of such a system. In Frotjold's test, it seemed to work in an opposite manner. *"The familiarity element did however result in different behaviour than expected….some subjects would insist on leaving the tap setting on a setting similar to the one they would use in their own shower, even when the feedback from the system indicated that this was too hot"* (ibid. p. 70).

Frotjold also found that none of the participants used the explanations continuously generated by the text generator. Some of the participants stated that this was because they had already received the necessary explanation through the basic explanation (see figure 6-2, p. 179) while others reported that this was because they did not have time during the two minute simulation run.

## 6.3 Lessons Learnt from Test 1

I will now briefly point to some lessons learnt from Frotjold's evaluation. First, some important problems with the prototype became apparent through her subtests. Several of the participants had a problem in understanding the nonlinearity in the access to the hot water tank. This would have to be further developed before a second test.

Second, some of the participants commented that the prototype was too text-based. This was also something that could be improved before another test.

Regarding research design there were also some lessons that could be learnt. As mentioned in the previous section the participants in Frotjold's study were divided into four groups who received access to different parts of the user interface. Test 1a only had one test person for each of the four test groups. This is, as Frotjold (ibid.) also states, not a large enough group to support generalizations. The second group consisted of 13 people and this is also a rather small group for the type of research design that is used. With the division into test groups each version of the prototype was only tested on 3 or 4 participants.

The results between the different groups were compared and conclusions regarding the relevance of each part of the user interface were made based on the participants' performance and comments in the interviews. The separate value of each part of the user interface may not be as relevant. It is difficult to test each part of the user interface because when it is incorporated into the user interface, it may make a different contribution to the user interface as a whole. The question is not *which part is best?* It is rather how the different parts function together as a whole, and how they can be improved or better integrated in order to create a better user interface. Therefore, in Test 2, which will be discussed in Chapters 7 and 8, all participants were given access to the same version of the user interface. The problem of small test groups was thereby alleviated.

A further lesson learnt from this test originates from the explanation to the participants of what they were expected to achieve. As discussed in Chapter 5, according to Alessi (2002) one should be careful to avoid giving the participants of interactive learning environments conflicting goals, that is, both to learn and to win. In Test 1 the participants were asked to try to understand the system and to obtain a comfortable temperature in two minutes. Such a mixed focus could cause the participants to be more involved in obtaining a comfortable temperature than trying to understand the underlying model. Some of the participants also reported that they found the time limit stressing.

There was also a focus on stabilization time in Test 1 through computer logs of the temperature stabilization time. This which may indicate the assumption that there is a link between stabilization times and understanding of the system. The Two-Shower model is a relatively small model and it is possible to obtain a stable temperature through trial and error. Stabilization time, therefore, does not necessarily indicate that the participant has understood why she was able to stabilize the system. In a addition, a simulation run of two minutes is very short, and it would be difficult to distinguish between different stabilization times in such a short time period (for example, the difference between 1 minute and 1 minute and 20 seconds may be very small).

A final lesson learnt was a need to look carefully at the relationship between activity and understanding. In Test 1 it is assumed that there is a relationship between the number of times the participant adjusts the tap setting and his/her understanding of the delay between the tap setting and the showerhead. Simulation logs are used to draw conclusions about whether the participants were considering the delay when making decisions to change the tap setting. In hindsight, there is not necessarily a direct relationship between the number of changes made in tap setting and the understanding of the delay.

Test 1 pointed to some important problems of the user interface. Some of the graphics in particular were further developed, such as the representation of the hot water tank. The testing method was also reviewed as will be discussed in the following chapter.

# 7. Test 2 of the Two-Shower Prototype

## 7.1 Introduction

In Chapter 6 I revisited Test 1 of the Two-Shower prototype. In this chapter I present Test 2 that was designed and conducted by myself after making some changes to the prototype. The design of this test and a discussion of evaluation methods are included in this chapter. In Chapter 8 I present the results of Test 2.

The focus of Test 2 was to gain feedback from the participants about the user interface of the prototype and its various components. The aim of this formative evaluation was to identify usability problems in the user interface with particular emphasis on the participants' attitude toward the prototype and its various user interface components. The participants were also asked to describe some characteristics of the underlying model that they had identified during the session. The purpose of these questions was not to test whether learning had actually taken place, but to identify characteristics of the underlying simulation model that needed to be communicated better in the user interface.

I open this chapter with a discussion of evaluation techniques and a brief introduction to usability testing and two overlapping evaluation methods on which Test 2 is based, namely *query techniques* (Dix et al., 1998) and *beta analysis* (Alessi & Trollip, 2001) (Section 7.2). In Section 7.3 I discuss some changes that were made to the prototype between Test 1 and Test 2. In Section 7.4 I provide a description of the test design for Test 2. There were additional improvements that could have been made in the design of Test 2, which I will discuss in Chapter 9, - after my presentation of the results.

## 7.2 Evaluation Methods

The background for Test 2 of the Two-Shower prototype was that I wanted to test the functionality of the user interface and collect feedback from potential users about how they felt that the user interface of the Two-Shower prototype and its various components aided them in understanding the underlying model. A further goal was that this feedback could provide useful indications of which parts of the prototype that should be altered in order to develop a better interface.

The field of human computer interaction (HCI) embraces a number of methods for tests and evaluations of user interfaces. This section is a short introduction to some of the traditional evaluation methods within this field intended to highlight the range of characteristics that might be tested before narrowing the field to the methods chosen for Test 2.

### 7.2.1 Usability Testing

One of the aims within the field of human computer interaction is to develop computer systems, software, or web pages that are user-friendly and that support their users in accomplishing tasks effectively. A focus on *usability* should promote the user-friendliness of the system. Mack and Nielsen (1994) provide the following definition of usability:

> *"Usability is a fairly broad concept that basically refers to how easy it is for users to learn a system, how efficiently they can use it once they have learned it, and how pleasant it is to use. Also, the frequency and seriousness of user errors are normally considered to be constituent parts of usability" (p. 3).*

The term *usability inspection* embraces a number of methods for testing various usability aspects of a user interface (Mack & Nielsen, 1994). The goal of a usability inspection is to evaluate the design of the user interface based on the judgment of the inspectors. The various methods vary in terms of how the judgment is derived and in terms of the criteria on which the inspectors are making their judgments (ibid.).

Nielsen (1993) lists five components of a usability inspection: *learnability*, *efficiency*, *memorability*, *errors*, and *satisfaction* (p. 26). *Learnability* refers to ease of learning to use the system while *efficiency* should ensure a high level of productivity once the user has learned how to use the system. *Memorability* is the degree to which a system is easy to remember and whether a user is able to return to the system after a while and still remember how to use it. The system must also enable the user to operate it without committing *errors* and without experiencing errors in the system. If the user makes an error, the system should also quickly be able to recover. *Satisfaction* refers to the extent that the users *like* the system and are subjectively satisfied by its appearance and ease of use.

Learnability may be tested by letting the users perform a task and study whether they are able to perform the task successfully (ibid.). The performance may also be timed in order to study whether the user can perform the task in a specific amount of time. Learning in this sense mainly refers to be able to use the program.

Efficiency generally refers to how efficient the system is in use by experienced users (ibid.). Experienced users, however, may be difficult to find. Efficiency is therefore often measured by letting some users use the system for a specific amount of time and measure their efficiency after that time. Performance may also be monitored continuously to see if it improves after some time. Efficiency is thus typically evaluated based on the time it takes the users to perform the test tasks.

Memorability refers to how easy it is to return to the system after some time and is generally directed towards casual users (ibid.). Memorability may be tested by users who have not used the system for a while and measure the time it takes for them to perform specific tasks. A memory test may also be given after a test session where the users are asked to explain various features of the system.

The system should also be tested for errors. *"Typically, an error is defined as any action that does not accomplish the desired goal, and the system's error rate is measured by counting the number of such actions made by users while performing* *some specified task"* (ibid., p. 32). The system can, therefore, be tested for errors at the same time the other usability tests are performed.

Subjective satisfaction is most often evaluated by simply asking the users about their satisfaction, usually in a short questionnaire after the test session (Nielsen, 1993). Licensed instruments for satisfaction testing also exists such as QUIS™ (Questionnaire for User Interaction Satisfaction) 7.0, initially described in Chin, Diehl, and Norman (1988), or IBM's PSSUQ (Lewis, 2002). The users are typically asked to rate different features of the system on a rating scale from 1-5 or 1-7. Another option is to rate the system based on how much the users like or dislike the system on a *likert* scale, responding from "I strongly agree" to "I strongly disagree" to statements about the system (Nielsen, 1993). Nielsen (ibid.) emphasizes that the questionnaire should be tested before it is used to ensure that the questions are interpreted properly.

The ratings for satisfaction are usually calculated as a mean of the scores, however, more complex methods such as psychometrics may also be used (ibid.). (See Lewis (2002) for an example of psychometric evaluation.)

A setting or test environment must be chosen for the usability test. Laboratory studies have been most commonly used for usability studies. In later years, however, there has been a focus on testing the system in the settings in which the systems would ordinarily be used (Wichansky, 2000).

According to Mack and Nielsen (1994) a scenario for the test must be chosen, and the decision of whether to give the participants specific tasks that they should fulfill or give the participants open-ended instructions where the participants try out the system themselves must be made. They argue that giving specific tasks ensures testing of the features that one wants to test, but that open-ended tasks, on the other hand, may ensure that the program is tested more in the way it would be used in a real setting. Specific tasks may not be provided in real life. Mack and Nielsen (ibid.) argues:

> *"The use of scenarios involves the obvious trade-off between ensuring that inspectors inspect an interface in terms of a specific focus and task flow that a usability specialist cares about, and allowing a potentially open-ended and inspector-driven exploration and exposure to an interface" (p. 9).*

Nielsen (1994) suggests letting the user go through the program at least twice in order to get a feel for the flow of the interaction with the interface the first time and to evaluate the specific elements the second time.

In choosing the participants for a usability test Nielsen (1993) suggests choosing participants who are as close to the intended users as possible. He further argues that most programs need to be tested by novice users, but that in some cases it could be useful to obtain feedback from groups of both expert and novice users.

In later years there has been a growing critique against the traditional usability studies pointing to, for example, the lack of focus on the relationship between computer systems and the social practices in which they are used (see for example Chaiklin, 2007; Nocera & Sharp, 2007; Wilson, Galliers, & Fone, 2007). For a formative test of the Two-Shower prototype, however, elements from a traditional usability test may be useful and some ideas from usability testing in general were adopted for Test 2.

In Test 2 of the Two-Shower prototype I decided to have participants use the prototype before filling out a questionnaire designed to elicit their opinions of the prototype. In this way, I wanted to gather comments about the design ideas of the prototype for the purpose of illuminating potential problems in the design of the user interface. Through comments from the participants, it was also possible to obtain feedback about the participants' own opinion of the learnability and efficiency of the prototype. Some errors were also recorded informally through notes during the sessions and some errors were reported in the questionnaires. My evaluation design for Test 2 was also informed by query techniques and beta testing which will be discussed in the next subsections.

### 7.2.2 Query Techniques

Query techniques constitute a form of usability testing. Query techniques are based on collecting feedback about the system from users who participate in a test of the system. *"Query techniques are less formal than controlled experimentation, but can be useful in eliciting detail of the user's view of a system"* (Dix et al., 1998, p. 431). Query techniques simply consist of asking the users how they feel about the system. The results are necessarily highly subjective but valuable feedback and suggestions for improvements may be provided. *"The advantage of such methods is that they get the user's viewpoint directly and may reveal issues which have not been considered by the designer"* (ibid.). The system can subsequently be modified based on feedback from participants in the test.

According to Dix et al. (ibid.) there should be at least ten participants in order to ensure a group that is large enough to be considered representative. Information about the participants' opinions and experiences with the system may be gathered either through interviews or through questionnaires. An advantage of interviews is that the interviewer may ask follow-up questions and ask the participant to be more specific (ibid.). Questionnaires, however, require less administration and it is possible to conduct a more rigorous analysis because all participants answer the same questions. In developing the questionnaires, it is important to consider which information should be gathered and to develop questions that may provide that information. The questionnaires may include general questions about the background of the participants such as gender, education level, and age. There may also be open-ended questions where the users are asked what they think about the user interface or problems they experienced. Through these types of questions it may be possible to obtain an accurate description of problems that the participants encountered and suggestions for improvements. A third type of question is scalar questions. The participants may, for example, be asked to answer how well they agree or disagree with a statement about the program, or like or dislike some feature of what is being tested. The answers given to these types of questions are more quantitative and little detail is provided. They may provide indications, however, about parts of the system that should be improved. The results from using query techniques may provide useful indications as to which parts of the system that should be considered for change as well as useful suggestions for what those improvements may be.

The second test of the Two-Shower prototype, as noted above, was partly based on query techniques. The participants were asked about their opinion of the prototype through questionnaires that included both, open-ended questions and scalar questions. Background information was also gathered about the participants. A description of the questionnaires will be provided later in this chapter.

### 7.2.3 Beta Testing

A final testing method that informed my test design was beta-testing. Alessi and Trollip (2001) discuss the assessment method, *beta testing* to test a variety of interactive learning environments. The beta testing method has several characteristics in common with the query techniques described in the previous subsection. It contains seven steps:

1. *Select the learners*
2. *Explain the procedure to them*
3. *Find out how much of the subject matter they know already*
4. *Observe them going through the program*
5. *Interview them afterward*
6. *Assess their learning*
7. *Revise the program*

*(ibid. p. 550).*

Alessi and Trollip (ibid.), Dix et al. (1998), and Nielsen (1993) all suggest that the participants of an evaluation test should be as similar to the intended users of the program as possible. If it is not possible to test the system on participants who are part of the actual user group, it is important to test the system on users who are similar to the intended user group with regard to age, education, and experience with computers. Alessi and Trollip (2001) further suggest choosing weak, average, and strong learners. They do not, however, go into the details of how one is going to categorize the learners in advance. I will return to this issue of choosing test-participants for Test 2 in Section 7.4.3.

One of the steps of beta testing is to explain the purpose of the test and the prototype to the participants and to explain how the test will be performed and how data will be gathered. The participants should also be encouraged to criticize the system as this feedback will be used as a basis for improvements.

According to Alessi and Trollip (2001) the participants should be observed as they use the system. They suggest that also the body language of the participants should be observed. This, however, requires an extensive analysis that may be out of scope for many development projects. They further state that the participants should use the program alone because it is easier to observe one person at a time.

Subsequently, the participants are interviewed after having used the program and comments and questions from both the researcher and the participants should be addressed. In this process, the researchers should focus on the functionality of the system as well as its content (ibid.).

The participants should be asked how they feel about the system. Caution, however, should be shown when interpreting data regarding attitudes, in spite of the fact that they may provide some valuable information about potential needs for improvements. If most of the users dislike the system, for example, this may be a clear indication that a revision should be considered.

Alessi and Trollip (2001) reference Kirkpatrick (1996) who presents four levels of training evaluation and argue that it is also necessary to assess what the participants have learned. Test 2 of the Two-Shower prototype does not include testing of what the participants learned using the prototype, as that was beyond the scope of this thesis (although it would be a natural extension of the usability tests we conducted). Further issues addressed by Alessi and Trollip (2001) concern whether the participants actually use what they have learned in real life situations and whether the development of the system can be defended based on its costs. These issues are not yet of great relevance to the Two-Shower prototype, because the prototype has not yet been used in an actual educational context. Moreover, the final analysis of costs and effectiveness is also not included.

Test 2 of the Two-Shower prototype is partly based on steps 1 through 5. The details of the test will be explained later in this chapter. Step 7, revision of the program was not performed, but changes to the prototype based on the results from the test will be suggested and discussed in Chapter 9.

The following section is a description of the main changes that were made to the Two-Shower prototype between Test 1 and Test 2.

## 7.3 Changes to the Prototype and its Interface

Several changes were made to the Two-Shower prototype before Test 2. The version of the Two-Shower prototype that was used in Test 1 was discussed in Chapter 6. The version of the Two-Shower prototype that was used in Test 2 is the one that is described throughout this thesis. Some of the changes that were made to the prototype between Test 1 and Test 2 were based on the findings from Test 1, - others were parts of the initial plans for the user interface (see Table 7-1 for an overview). The changes were made by Olve Sæther Hansen and myself.

The results from Test 1 indicated that some of the participants had problems in understanding the nonlinear affect of the hot water pressure on the temperature at the showerhead represented by an animation of the hot water tank in the graphical user interface. A new version of the changing hot water pressure was therefore implemented in the Pipe System view (the old and new versions are both described in Section 4.3.3). The variations in the hot water pressure are now illustrated by yellow lines flowing at varying speeds through the pipes over time.

In Test 1, the basic explanation provided a relatively detailed description of the system, available to those who had access to it, before running the prototype (see figure 6-2, p. 179). The basic explanation describing the system in advance was removed from the new version of the Two-Shower prototype that was used for Test 2. Although Test 1 showed that the participants who had access to the textual explanations performed better, there was a concern that it would remove part of the *excitement* of using the prototype and thereby maybe some of the participants' eagerness to explore the prototype during the sessions. In addition, some of the participants in Test 1 found the prototype to be too text-based and found the text to be a somewhat redundant. The basic explanation was replaced by the introductory story of the person who arrives at a hotel after an exhausting journey (see Section 3.3.1).

Another difference between the version of the prototype that was used in Test 1 and the one used in Test 2 was that behavior graphs of important variables associated with the delay and nonlinearity that were included in the user interface. These were part of the initial plans for the user interface and were included because it could be useful to have access to an overview of the historic behavior of the underlying model.

The story generator was different in the two versions used for Test 1 and Test 2 in that the version used for Test 2 included changes in graphics and animation, not only text. The important points in time were also marked by yellow markers on the timeline. The reason the changes in graphics and animations had not been implemented earlier was due to programming capacity. The text box was moved towards the right under the graphs (see Section 3.4).

The Two-Shower prototype was also turned back into a two-user system because it was my initial intention that the Two-Shower prototype should provide support for collaboration between users.

In figure 7-1, I offer a summary of the main changes that were made to the user interface between Test 1 and Test 2 of the prototype.

| | |
|---|---|
| Graphical representation of the hot water tank | Changed based on results from Test 1 |
| Introductory story rather than basic explanation | Part of initial plans of including narrative but also to make the system less text-based as was commented in Test 1 |
| Behavior graphs included | Part of initial plans |
| Inclusion of graphics and animations of the story generator | Part of initial plans |
| Important points in time marked on timeline | Part of initial plans |
| Turned into a two-user system | Part of initial plans |

*Figure 7-1: Changes made to the user interface of the Two-Shower prototype between Test 1 and Test 2*

Even though several changes were made to the prototype between Test 1 and Test 2 the version that was used in Test 2 was still a prototype and therefore in some sense incomplete. Essential features that were captured were the ideas and concepts of the visualizations, but further refinement of these would be needed. Some of the problems that were not yet resolved related to technical problems regarding the server and the control of the Flash animations. Some of these problems will be discussed in Chapter 8.

## 7.4 Test Design for Test 2 of the Two-Shower Prototype

In this section I describe changes in the research design for Test 2 of the Two-Shower prototype, the questions that were asked, and how I intended to provide answers to these questions through the evaluation. The goal of the test is discussed before I address my choice of participants. The setting in which the test took place is then discussed, followed by a discussion of the explanation of the procedure to the participants. The last four subsections consider how questionnaires, notes, and video recordings from the sessions were used to gather data about the participants' opinions of the system.

### 7.4.1 The Goal of Test 2

As mentioned earlier in this chapter, Test 2 of the Two-Shower prototype is based on usability testing (Nielsen, 1993), the beta test described by Alessi and Trollip (2001), and the query techniques described by Dix et al. (1998). As the Two-Shower prototype is only a prototype, the focus of the test was mainly on characteristics of the user interface that should be improved in order to provide better support for users. My main concern was whether the design ideas of the Two-Shower prototype worked as intended and to identify features that needed improvement. The user interface is evaluated based on feedback from participants in the test and their comments about the prototype.

When administrating Test 2 I wanted to the following questions answered:

- Is the user interface designed in such a way that the participants find that it aids them in obtaining an understanding of the underlying model?

- What are problems associated with the user interface as seen from a usability perspective?

- Are there particular elements or characteristics of the representation of the underlying model in the user interface with which the participants report they have problems?

- What is the participants' attitude towards the prototype and which parts of the prototype do they have problems with or would they like to have changed?

The intent was, therefore, that the evaluation should result in feedback about the parts of the user interface that should be considered candidates for improvement.

### 7.4.2 The Number of Participants - A Collaborative System

As discussed previously, only one person at a time participated in the sessions in Test 1. In Test 2 a collaborative two-user version of the system was used and the participants worked together in pairs at all times throughout the sessions.

Alessi and Trollip (2001) argue that it may be better to observe one learner at a time to avoid missing important information. For this evaluation, I do not agree. Important information may be missed, rather, when the system is used by one participant at a time because the participant does not have anyone with whom he/she may communicate (except for the researcher). Observing facial expressions as encouraged by Alessi and Trollip (ibid.) may provide some additional information about how the participant feels about the system. It may, however, be time-consuming and difficult to interpret facial expressions may be. Observing the conversation between two people as they use the system may provide more information about their experience with the system and how they use it to try to obtain an understanding of the underlying model.

The data gathered through video recordings were not extensively utilized in Test 2, but further suggestions regarding how these kinds of data may be utilized will be made when I address the proposed future study of the prototype in Chapter 9.

### 7.4.3 Pre-Questionnaire and the Participants

As discussed previously, the intended users of the Two-Shower prototype were first year students in system dynamics. The participants of Test 2, however, were not part of an introductory course in system dynamics. This is because the test was carried out towards the end of the semester and no novice students were available. The downside of choosing participants that are not part of a course about complex systems or system dynamics is that they may lack motivation for using it. They may not understand the goal of the prototype when it is not experienced as part of a larger educational context. Their participation in the test was strictly on a voluntary basis.

Alessi and Trollip (2001) suggest to divide the learners into three categories according to their capabilities as learners and their previous knowledge about the problem. *"One should be representative of the best of the potential learners, one an average learner, and one similar to the slowest of the learners that will use the program"* (ibid. p. 550). I found this categorization rather difficult because there are many factors involved in positioning the participants into these groups, and I did not find this strict categorization useful for this evaluation. However, the participants were asked some introductory questions regarding their prior experience in mathematics and system dynamics, and experience with computers in order to obtain some information about their background (see Appendix A). Moreover, the participants were given a pre-questionnaire in which questions about their:

- Gender
- Age
- Educational background
- Previous experience with computers
- Previous experience with mathematics
- Previous experience with system dynamics

In particular I found it important to record whether any of the participants had extensive experience with system dynamics. Previous experience in mathematics, system dynamics or systems thinking may influence how participants approach the problem. The background of the participants could also influence how they interact with each other and with the prototype. Their experience with computers, for example, could influence their ability to navigate in the system and would probably also have impacts on how comfortable they were with using the system.

There were twelve participants in Test 2. Two of the participants had taken a course in system dynamics or a course where system dynamics was one of the subjects. Five of the participants said they knew what system dynamics was, while five of them had only heard of it before. None of the participants had extensive experience with system dynamics, but they had all heard about it.

All participants had completed or were part of a 5-6 year masters program. Ten of them were Masters or Ph D students at the Department of Information Science and Media. These ten, therefore, all had a technology-orientation in their education. The other two participants had higher education (equivalent to Master's degrees) in the fields of natural sciences. There were nine women and three men in the study. All participants relatively frequently used computers for other than text editing and browsing. Three of the participants had taken mathematics courses at university level, while the others had taken mathematics at a high school level.

As discussed in Section 6.3 Test 1 was composed of two subtests consisting of two different user groups. One consisted of junior high school students and the other one of master students in information science. Frotjold (2005) found that it seemed that junior high school students were too young and found that they were unable to express themselves about the model. Some of the participants of Test 1, however, were younger than the intended user group.

The average age of the participants in Test 2, on the other hand, turned out to be higher than the intended user group. The intended group ranges from undergraduate students to graduate students, which would probably indicate an age group between about 18 and 30. The average age of the participants in Test 2 was 33.25.

### 7.4.4 The Context

As discussed previously, the test sessions in Test 2 were not part of an actual instructional context and thus not part of the context for which the prototype is intended. At this stage in the development process, it was considered better to use a laboratory test. I expected to find some problems with the prototype that should be solved and it would have been too early in the development process to use it in a real educational context.

The results of laboratory tests of the kind that I have undertaken do not lend themselves to wide generalizations. In an educational context, there are many additional factors that influence the outcome, - factors that are missing in the laboratory context. A lab evaluation is, however, less time consuming, easier to set up, does not require the time of lecturers or the availability of large computer labs. I therefore concluded that a lab was the appropriate context in which to conduct my tests at this stage of the development, but I underscore that the results must be interpreted as such, and not as if they were part of a larger, more user-realistic context.

The lab was set up with two laptops located next to each other, with the intention of letting the two participants discuss the prototype and be able to look at the screen of the other participant as the prototype was running. One video camera was used to capture the screen and face of one participant but only parts of the screen and face of the other participant. Even though one of the participants was only partially present in the video recordings this was seen as sufficient as the video material was mostly used as a supplement to the notes that were taken during the sessions. Speech from both participants was captured on video. In a more extensive test it would have been necessary to have at least one camera that captured the facial expressions of the two participants in addition to recordings of the screens of both participants.

### 7.4.5 Explanation of the Procedure to the Participants

According to Alessi and Trollip (2001) the purpose of the test should be explained to the participants prior to their use of the software. In Test 2 of the Two-Shower prototype, each group was given a short, oral introduction to what the test was about (see Appendix A for details). They were told that they were going to take part in an evaluation of a prototype of an interactive learning environment and that the purpose of the study was to test the design ideas behind the system. They were also told that their goal was to control the system through collaboration and to try to understand the underlying model. They were asked to discuss what they were doing in order to make it easier for the researcher to understand the problems they would encounter.

I debated whether the participants should be able to ask for assistance during the sessions. If they were allowed to ask for assistance, it would more effectively expose the problems they had running the prototype in real time rather than having them remember the issues until after the simulation run. In a classroom context, for which the learning environment is intended to be used, it would also be common for participants to be able to ask the teacher for assistance. The problem of allowing them to receive assistance would be that they would not use the user interface in the same way, and it could be difficult to find problems of the user interface if they asked questions even before trying to figure out the solution by studying the user interface themselves. It was decided to ask the participants, to the extent it was possible, not to ask for assistance. But I did choose to interfere in cases where there were some obvious usability issues that could not be resolved by the participants. In those cases I made a note of the problem and helped them to move on. The participants were informed that the session would be recorded on video and that notes would be taken during the simulation runs. They were also told that they were going to fill out a pre-questionnaire with some background information and a post-questionnaire with feedback about the system and problems that they encountered and that they were expected to respond to some questions regarding the underlying model.

### 7.4.6 Observation during the Simulation Runs

As discussed previously, the sessions in our test were videotaped, but the material was not extensively used due to time limitations and the scope of the test. Rather than transcribing and using all the video material in the study, the video tapes were used to return to certain moments in the test to recall what had happened. In this manner the video material was used as a supplement to notes that were taken during the sessions. A more thorough study of the video material could be conducted at a later stage.

Each group participated in two simulation runs. Not all participants noticed the pipe system button and in the cases where it was not noticed during the first simulation run, this option was pointed out to the participants before the prototype was restarted.

In cases where the participants did not pause the system during the first simulation run to employ the story generator, this feature was also pointed out before the second simulation run.

### 7.4.7 Post-Questionnaires: the Participants' Attitude towards the Prototype and How They Reported that the Prototype Aided their Understanding

After running the prototype twice, the participants were asked to fill out a post-questionnaire about their opinions of the prototype and some questions related to the structure and behavior of the underlying model (see Appendix A).

The first question concerned what they thought about the prototype and gave the participants an opportunity to offer some general comments. The remainder of the first part of the questionnaire consisted of a number of questions regarding how they felt that specific parts or characteristics of the user interface aided their understanding of the underlying model. The alternatives ranged from whether they felt that these parts made the underlying model easier or more difficult to understand. After each question with fixed alternative answers they were asked to comment on particular parts of the prototype in their own words.

The parts of the prototype that the participants were asked to evaluate and comment on were:

1. The Hotel view
2. The Shower Room view
3. The Pipe System view
4. The graphics
5. The slide shows
6. The animations
7. The behavior graphs
8. The textual explanation which becomes visible when the hints button is pressed
9. The narratives or explanations that appear when the pause button is pressed

10. Organization of the use interface
11. Navigation
12. The collaborative aspect

I wanted the participants to provide feedback on these particular characteristics as they had been the focus of the development process.

### 7.4.8 Post-Questionnaires: Questions about Model Characteristics

For the Two-Shower prototype it was decided to ask the participants some questions regarding the underlying simulation model after it was run. The purpose of these questions was to supplement the participants' reported views on the components of the model and to use their descriptions of the system structure and behavior to find potential problems with the user interface and to find parts of the user interface that needed to be more explicitly exposed. The intention was to reveal whether most of the participants clearly provided wrong answers to these questions, in which case that could be considered a clear indication that the prototype did not represent the characteristics of the underlying simulation model as well as intended.

Two questions regarding the underlying model were asked:

1. *Can you explain why the temperature oscillated?*
2. *If you did not change your tap setting for a while you could still experience changes in shower temperature. Can you explain why?*

There is, however, a problem with these questions, in that there is an implicit assumption that the answers can be directly linked to the quality of the user interface. This problem will be further discussed in Chapter 8.

The participants in Frotjold's (2005) evaluation test, Test 1 had difficulties in describing the model after the simulation run, and she had to ask leading questions about their understanding of the system. I therefore found that it could be helpful for the participants to review the user interface of the prototype when they answered these questions after the simulation run. The participants in Test 2 therefore had access to the user interface as they were answering the questions in the post-questionnaire.

With this description of the research design and testing process (for Test 2), I turn in Chapter 8 to a review of the results of Test 2 and indications of changes that could be considered based on the results.

# 8. The Results from Test 2 of the Two-Shower Prototype

## 8.1 Introduction

In Chapter 7, I presented the research design of Test 2 of the Two-Shower prototype. In this chapter, I discuss the results from the test and design changes they indicate.

Section 8.2 is a presentation and discussion of the participants' opinion of the Two-Shower prototype resulting from Test 2. Each subsection deals with a question in the first part of the post-questionnaire (see Appendix A). The answers are to some extent discussed in light of the background information provided by the participants in the pre-questionnaires and observations and notes that were made during the sessions. In Section 8.3 I review the second part of the post-questionnaires, where the participants were asked questions regarding the underlying simulation model of the prototype. The intention behind these questions was, as discussed in Chapter 7, to obtain indications as to what parts of the prototype should be improved or emphasized differently. As also mentioned this method is questionable because there may not necessarily be a direct relationship between their answers and the quality of the user interface of the prototype. The answers from this section of the test are therefore not to a great extent used as part of indications for improvements on the prototype.

## 8.2 Results from Part One of the Post-Questionnaire

This section considers the answers regarding the participants' attitude towards the prototype. As described in the previous chapter, I expected, from the results of Test 2, to receive feedback about the prototype that could be used as a basis for future design changes. Some usability problems of the prototype were already apparent, such as problems turning the mixing battery, but they had not been improved due to limited resources and time constraints. In this section I present the results and consider briefly what changes could be made to the design to increase usability.

The version of the Two-Shower prototype that the participants used was in English. In the questionnaire all questions were asked and answered in Norwegian and have subsequently been translated into English. Translating the answers was rather difficult because some comments were not written as complete sentences. It was also difficult to make the translation as close to the original as possible because sometimes the meaning would be lost if the sentences were not translated into complete sentences. Some of the comments are, therefore, transformed into complete sentences in the English translation. The words that were included in the translation are denoted by brackets. The problems associated with the interpretation of answers written as incomplete sentences is one of the disadvantages resulting from the use of questionnaires with open questions and will be discussed further in Chapter 9.

### 8.2.1 General Comments about the Two-Shower Prototype

I open with some general comments and then move on to specific interface views and components. The first question of the post-questionnaire was:

> *What is your opinion of the Two-Shower prototype?*[26]

This question was intended to give the participants the opportunity to make some general comments about the prototype. There seemed to be a difference in the perception of the prototype between the various participants. Some seemed to like the prototype while others did not. Some positive comments were:

> *"The idea is exciting – nice illustration of the relationship between two partly overlapping systems."*
>
> *"[The prototype] showed the principles well."*

---

[26] As described previously, the Two-Shower prototype was originally called the Two-Shower Model, but is referred to as a prototype in this thesis to emphasize that it is not a traditional system dynamic model, but that it has a different interface.

> *"It was a well-arranged model regarding the representation of what was happening."*

Three of the participants were more critical in their comments:

> *"Interesting ☺ It was maybe a little difficult to understand why things were happening."*
>
> *"It was difficult to relate to so much information that is provided in different places on the screen."*
>
> *"Seems like a fun idea. But the application is characterized by hastiness, with a little/medium user-friendly interface. An unnatural location of several of the elements hinders natural linking of related elements."*

The third comment is likely to refer to some of the usability issues of the prototype (this interpretation is based on later comments by the same participant). Some of these issues, such as the problems of the mixing battery, we were aware of in advance, but, as indicated above, we did not have the resources to adjust the prototype accordingly in time for the test. Others had not been considered, such as difficulties of finding the Pipe System view. Various usability issues will be discussed more in detail throughout this chapter.

One of the intentions of choosing the problem of two showers that shared a hot water tank to illustrate complex systems and resource sharing, was to allow the participants to utilize their previous experience with tap settings and shared hot water supplies to develop an understanding of the underlying model. In this respect, one of the participants noted the following:

> *"The problem is highly current in all households. May seem like a good example of simple systems."*

Another participant was concerned that her prior knowledge and experience would affect her understanding and dealings with the prototype:

> *"Ok model, but is the user affected by previous visits to old houses etc. and am I therefore prejudiced?"*

Using a system that would be familiar to the users was as mentioned, part of the initial intentions of the Two-Shower prototype so if this comment refers to a concern that the participants may be able to interpret the structure and behavior better based on prior knowledge, then this is not a concern. In Test 1 of the prototype, however, the familiarity aspect constituted a problem (Frotjold, 2005). Some of the students would insist on choosing tap settings that were similar to that in their showers at home and continued to turn the tap setting higher even though the feedback from the prototype signaled that the temperature was too hot. This problem was not reported or observed in Test 2.

A further intention with the prototype was to make the participants identify with the characters in the shower and the *feeling* of hot and cold water. One of the participants noted the following:

> *"Fun to try to act based on knowledge about known stimuli, without these being present → the feeling of warmth."*

Another purpose of the prototype was to support collaboration. Although Test 2 is not designed to test whether the prototype actually supports collaboration, one of the participants pointed out the following advantage of being able to collaborate:

> *"It was ok. I quickly understood the principles since I had the opportunity to talk to the person next to me and to look at her screen".*

One of the problems of the test may have been that because the Two-Shower prototype was not placed in a real educational setting and the participants were not system dynamics students, it may have been difficult for the participants to understand the objective of developing or using such a prototype. This is reflected in the following comment from one of the participants:

> *"Nice graphics, funny pictures. It was a little difficult to understand the point of such a model – what use you may have of it."*

This comment may also indicate lack of information about the intention behind the prototype. In an actual educational setting or with participants form the target user group it may have been easier for them to recognize the purpose of such a prototype.

In the following sub-sections, I will address particular interface views and some of their components individually.

### 8.2.2 The Hotel View

The participants were asked two questions regarding the Hotel view and the introductory story (repeated in figures 8-1 and 8-2). As discussed in earlier chapters the Hotel view is the introduction to the prototype and contains little important information about the characteristics of the underlying simulation model. The intention behind the introductory story about the tired guest who arrives at a hotel is to contribute to the narrative aspects of the prototype and to introduce some structural elements and some potential problems of controlling the shower behavior. The first question regarding the Hotel view was:

*3 a. The Introduction view [the Hotel view] was:*

*Difficult - A little difficult - Neither difficult nor easy - A little easy – Easy*

*to understand.*

*3 b. Comments about the Introduction view:*

Figure 8-3 is a graph that shows the distribution of the responses from the participants. I emphasize that the data represented in figure 8.3 and the subsequent graphs in this chapter do not carry statistical significance because the sample is too small. The graphs are used only as visualizations of the general opinions that the twelve participants reported about various parts of the interface. Seven participants answered that the Hotel view was *easy* to understand, three that it was *a little easy* to understand, and two that is was neither *difficult nor easy* to understand.


*Figure 8-1: The Hotel view of the Two-Shower prototype, repeated.*

*Imagine the following scenario: You arrive at a hotel for the night, and tired from an exhausting journey you long to take a refreshing shower. The bathroom looks OK and the shower as well - until you turn on the water... You struggle to find the right temperature, turning the tap more and more desperately towards hot or cold with not immediate results. Finally, you realize that a little patience waiting for the temperature to change before you turn on even more hot or cold water helps more than cursing. Then you hear the shower in the next room being turned on, and it is the same story all over again. What happened? Remember that you can press the pause-button at any time during the simulation and receive comments about what has happened so far!*

*Figure 8-2: The introductory story, repeated.*

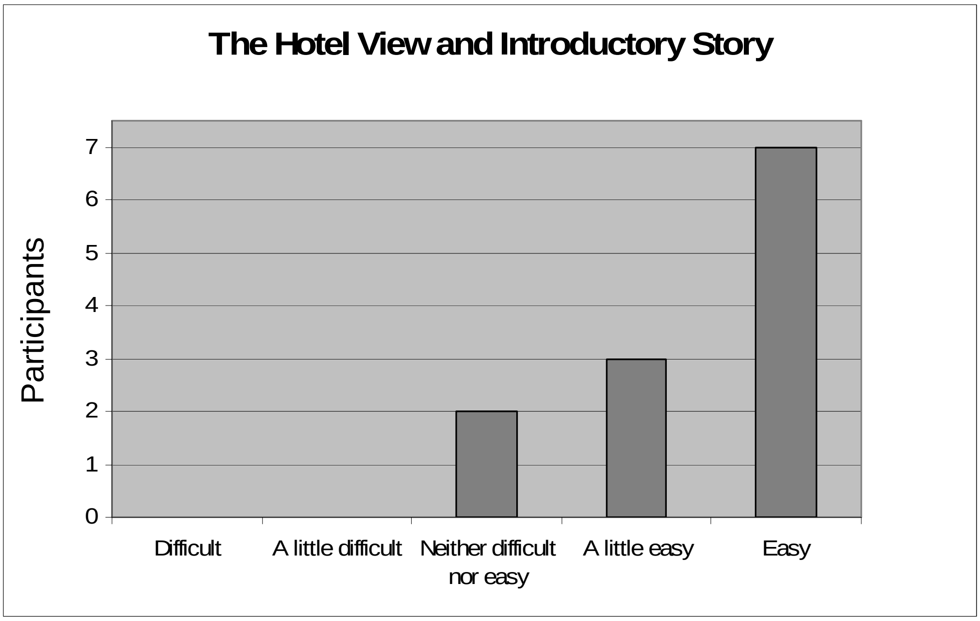


*Figure 8-3: The distribution of answers from the participants regarding how they felt about the Hotel view.*

Most of the participants did not report that they experienced problems with this view and some of their general comments were:

> *"Clear and precise introduction, I quickly understood what it was about."*
>
> *"Ok explanation of the scenario."*
>
> *"Simple and ok – logical story to attach it [the problem] to."*
>
> *"Good explanation."*
>
> *"Relatively well explained."*
>
> *"Quite ok. Reading could be done at your own speed, not controlled by the system."*

The last extract probably refers to the Introduction text box containing the introductory story allowing the participants to proceed at their own speed (see figure 4-3, p. 120).

Some comments were more precise about improvements that could be made. During some of the sessions there was a problem with the server that ran the underlying simulation model and the prototype had to be restarted several times. Some of the participants, therefore, had to go through the Hotel view and the introductory story several times. One of the participants suggested that it could be left out if the prototype were to be run more than once.

> *"It should be possible to skip it, very good the first time."*

As described earlier, the introductory story is presented in message boxes where the participants can click *next* and *previous* at their own pace and, thereby, navigate their way through the story. One of the participants suggested that, instead of using message boxes, the introductory story should rather be integrated as a part of the user interface using Flash:

> *"[You] should have used Flash for the introduction view – not message boxes. They are not perceived as a part of the system, but as a message about something."*

Some of the participants found that there was too much text in the introductory story, and that it could have been easier if it was extended with illustrations:

> *"A lot of text – or not a lot of text but very text-based. Ok, but could be sparked up a little visually. I didn't see the picture of the hotel… The intro-story could have been followed by pictures (cartoon, comic strip)."*

Another participant had the following comment regarding the introductory story:

> *"[It] could have been extended with pictures. Some lack in relationship/connection between the different pages."*

If these suggestions had been followed up in the design of the prototype, they would have added to its narrative aspects. The problem with too much text in the prototype, which was noted by several of the participants, was also commented on in Test 1 of the Two-Shower prototype (Section 6.3 and Frotjold, 2005). In the development of a future version, a reduction of the amount of text and an improvement of the visual support, therefore, should be considered.

### 8.2.3 The Shower Room View

Questions about the Shower Room view were asked in a similar manner as the questions about the Hotel view. The questions were:

> *3 c. The Shower Room view was:*
>
> *Difficult - A little difficult - Neither difficult nor easy - A little easy – Easy*
>
> *to understand.*
>
> *3 d. Comments about the Shower Room view:*

As discussed in Chapters 3 - 5 one of the intentions using the Shower Room view was to represent the behavior of the underlying simulation model, illustrating only certain characteristics of the model at a time (see figure 8-4). In part this was to trigger the attention and curiosity of the participants. The participants' feedback about the Shower Room view was more varied than in the case of the Hotel view. One of the participants found it to be easy to understand, four found it a little easy, four found it neither difficult nor easy, while three found it to be a little difficult to understand (see figure 8-5). A problem regarding this question is that it is not clear whether the answers apply to difficulties of the view or difficulties in understanding the underlying. As mentioned before, the goal of this view was not that the participants were supposed to find the underlying model easy to understand, but to trigger their curiosity. The comments that were given by the participants are therefore more valuable than their ranking of the interface across the various levels of difficulty illustrated in figure 8-5.


*Figure 8-4: The Shower Room view, repeated.*

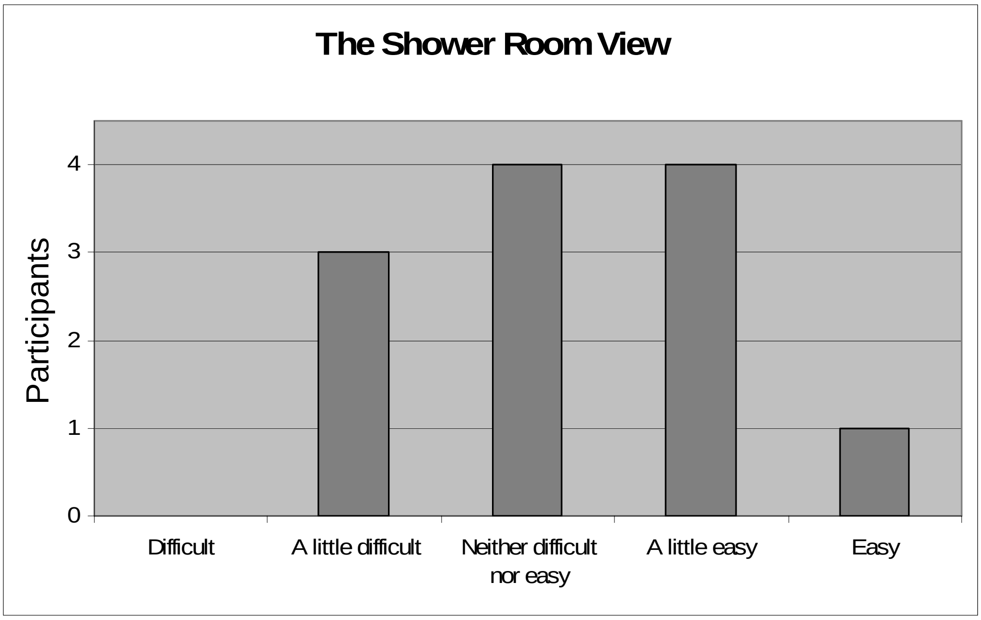


*Figure 8-5: The distribution of answers from the participants regarding how they felt about the Shower Room view.*

Some general comments on the Shower Room view were:

> *"The part of the model I liked the best."*
>
> *"Fun graphics, fun that it was actually two different people".*

The second comment may indicate that this participant may have liked the characterization in the prototype, which was implemented with the intent to incorporate narrative aspects of the prototype.

Some other comments were:

> *"After some trial and error, and communication with the neighbor, it was easy to understand how the Shower Room view functioned".*
>
> *"It took a couple of seconds before I found out that it was smart to coordinate the turning [of the tap] with the other that was showering. Easy to understand, but maybe easy to turn the 'knob' a little 'agitated' before the temperature stabilized."*

These comments may be linked to the intention of adding support for collaboration in the prototype. Although the participants did not provide indications as to whether the prototype actually supports collaboration.

One of the participants was, however, not particularly satisfied with the Shower Room view:

> *"The location of several of the elements was illogical. Little standardization. The elements do not have the same style, for example borders around some things etc."*

This comment points to how the organization of each view in the Two-Shower prototype should have been better. Part of the problem was that because of the nature of a prototype, not all standard buttons and functions were incorporated into the user interface. Moreover, because the project group was so small (there were at any time no more than two people working on the prototype) more attention, at times, have been given to solving technical difficulties and programming aspects of the prototype rather than attending to usability aspects. The organization of each view (such as for example the location and appearance of the simulation control buttons and navigation between the different views), should probably have been prioritized.

In the Shower Room view, two of the main usability problems of the prototype became apparent. One was that in order to leave the introductory story (close the Introduction text box, see figure 4-3, p. 120) the participants had to press a button named *Enter shower*. During the sessions it was observed that it seemed like several of the participants then expected the simulation model to start because they would immediately try to change the tap setting. This may in part be due to the fact that they not only had to press a *Enter shower* button in order to enter the Shower Room view; - they subsequently had to press the *play* button in the Shower Room view itself. The location of the play button was also not intuitive because the user interface had been extended in a way that called for a relocation of the play button (see figure 8-4). The play control buttons had originally been placed in the left hand corner but were in the end of the development process, after additions to the user interface, left further up on the screen.

Another usability problem concerned the control of the tap setting. The hit point was on the lower part of the mixing battery and not the handle (see the mixing battery in figure 8-4). Comments on the problems of the mixing battery were expected and, in hindsight, it should have been changed before the evaluation as nearly all the participants had problems turning the tap. The problem of the mixing battery was observed during the sessions and reported in the post-questionnaire. The participants had the following comments:

> *"I didn't immediately understand that I couldn't control the temperature by turning the part [the handle] that was sticking out of the mixing battery."*
>
> *"It was not intuitive to press play. OK with temperature regulator, but also not easy to know how it was regulated."*
>
> *"Would have been natural to pull the 'handle', not turn the wheel [the lower part of the mixing battery]."*
>
> *"Control of the battery was difficult. The extra 'arm' [the handle] on the actual switch [the lower part of the mixing battery] was just for decoration and can be dropped."*
>
> *"What is the point of making the tap so difficult? In the beginning it doesn't seem to have anything to do with the functionality."*
>
> *"Turning the tap could be a little more intuitive."*

There were also some comments about the control buttons. In the system there are only buttons for *play* and *pause*. Running the underlying simulation model generated a large amount of data and if the model was to be restarted the server would also have to be restarted in order for the capacity of the computer processor to become available for the new session. The *stop* and *reset* buttons were therefore removed before the experiment with the intention of avoiding confusion due to control buttons that were not functioning. The lack of standard control buttons caused some confusion for some of the participants and that may have been, in addition to their location, part of the reason the participants had difficulties finding the *play* and *pause* buttons. One of the participants made the following comment:

> *"When there is a play button you expect that there is a stop button too (pause button + play could have been one button). And an exit button (leave the hotel) is always good"*

The incorporation of the play and pause buttons into one button had not been considered. Some of the participants had problems with the play and pause buttons during the simulation runs and it seemed that they did not find it intuitive to press the play button again after having pressed the pause button. This problem may also be due to a problem with the story generator, which will be discussed later in Section 8.2.10. The prototype also lacked an exit button that must be added in future versions.

The next subsection continues with feedback that was given on the Pipe System view.

### 8.2.4 The Pipe System View

The next questions asked the participants about the Pipe System view. The questions were:

> *3 e. The Pipe System view was:*
>
> *Difficult - A little difficult - Neither difficult nor easy - A little easy – Easy*
>
> *to understand.*
>
> *3 f. Comments about the Pipe System view:*

As discussed in Chapters 3-5 the Pipe System view was intended to provide further transparency of the model structure through multimedia representations of the pipes, the hot water tank, and the two showers (see figure 8-6). Two of the participants found the Pipe System view easy to understand, four found it a little easy, three neither difficult nor easy, and three a little difficult (see figure 8-7). The answers here are even more widely distributed than the question regarding the Shower Room view. Again, most important are the comments that were provided by the participants. Several important issues were pointed out regarding the interface view.

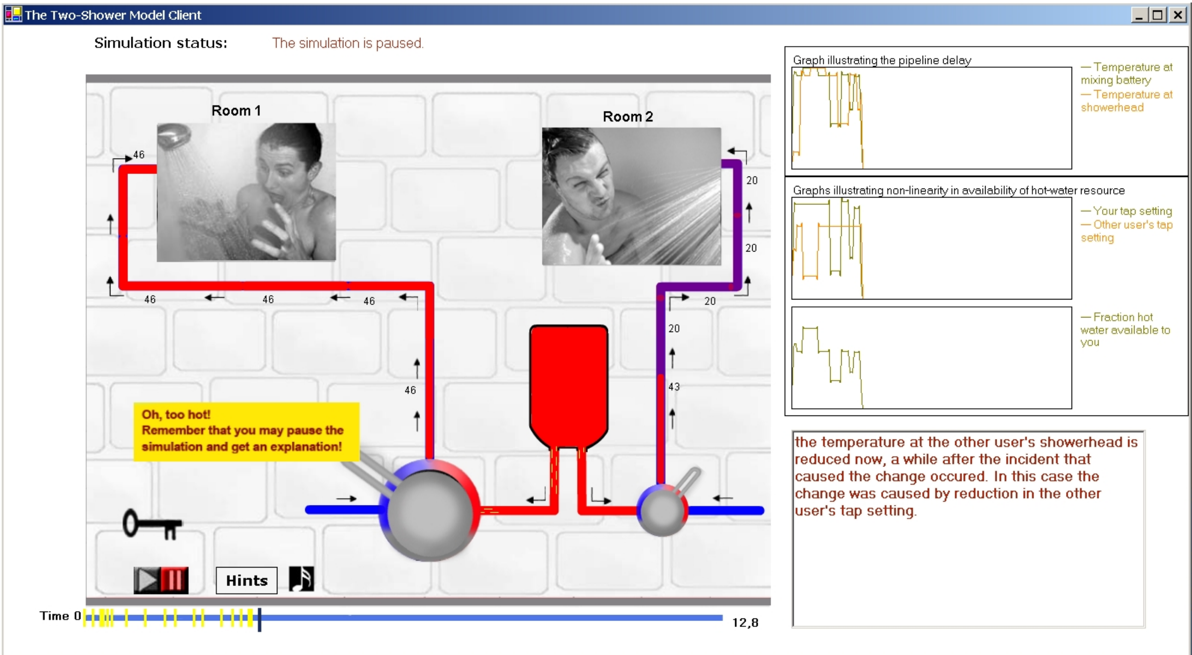


*Figure 8-6: The Pipe System view, repeated.*

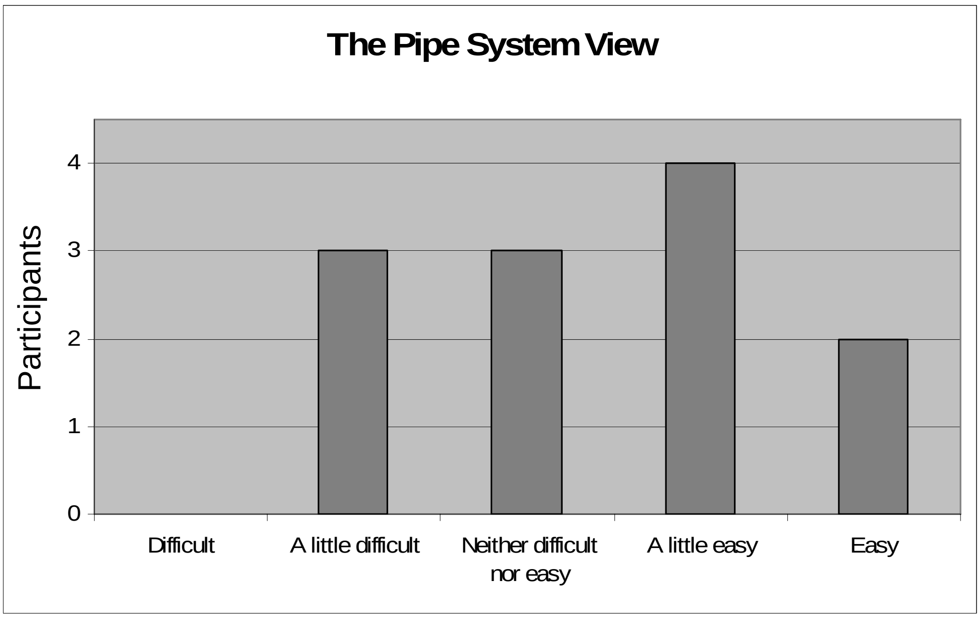


*Figure 8-7: The distribution of answers from the participants regarding how they felt about the Pipe System view.*

One of the main problems of the Pipe System view actually turned out to be the trouble of finding it. Almost none of the groups managed to find it by themselves. When they did not manage to find it in the course of the first simulation run it was pointed out to them between the first and second simulation run. One of the comments was:

> *"It took time before I discovered this possibility because the Pipe System button was experienced as partly hidden by the yellow text field [the yellow warning pop-up]."*

The yellow warning pop-up was not present at all times, but almost all participants still had difficulties in finding the Pipe System button. The problem of finding it may indicate that the button is not well integrated into the user interface and that the opportunity to navigate between the Shower Room view and the Pipe System view should have been made clearer.

Some more general comments about the Pipe System view were:

> *"Visualization of the pipe system was the best part of the model. A good explanatory model of the shower room."*
>
> *"On the second try I saw more use of using this picture [view]. By use of this picture [view] it was actually easier to understand why it turned cold and hot. A useful tool.*
>
> *"The best and most intuitive part. Shows nicely how the water moves and the temperature changes."*

These responses may indicate that some of the participants found that the Pipe System view made the characteristics of the underlying model clearer.

As discussed in Section 5.6, one of the goals of the Pipe System view was to support collaboration between the participants and that the visualizations could be used as a means for grounding. One of the participants made a comment regarding how it was helpful to see the conditions of the other participants, although this does not give any indications about whether the visualizations actually supports grounding:

> *"It was easier to see the relationship between changes in the tap setting and both showers. It also took some time here to turn the 'knob' back and forth in the beginning, but [it was] easier to see that it [the temperature] was stable for both."*

As discussed in Chapters 3 and 4 the pipeline delays were represented by graphics and animations of water with changing temperatures that was flowing through the pipes. Two of the participants made comments regarding the pipeline:

> *"It [the model] was simplified considerably by the indication of the temperature in the pipe section. [I] could see the difference in the length of the pipe etc."*

> *"[It was] easier to see how the temperature slowly changes and how it took time before hot/cold water reached the showerhead."*

In the Pipe System view the nonlinearity of the access to the hot water was represented by yellow lines flowing out from the hot water tank with changing speeds according to the share of hot water available to each participant (see figure 8-6). There was, however, a comment that indicated that at least one of the participants may have had problems with the graphics explaining the nonlinearity:

> *"All over ok, but didn't quite catch the point of the yellow dots in the tank [the pipes from the tank]. Otherwise ok."*

This was the only comment by any of the participants about the yellow lines from the hot water tank and, although sharing of the hot water was discussed by all participants during the sessions, they did not discuss the actual representation of the nonlinearity in the user interface. With the test design that was utilized it is impossible to infer whether they did not discuss the yellow lines because they did not notice them, they did not understand them, they immediately understood what they represented, or some other reason.

The two client applications (Room 1 and Room 2) had the same structure and appearance except for the room title and character in the shower. The data about the *current participant* was on the left hand side of the view while data about the *other participant* was on the right hand side in the Pipe System view. The location of Room 1 and Room 2 (controlled by participant 1 and participant 2) caused some problems for a few of the participants during the sessions and one participant had the following comment:

> *"I think it is more logical that Room 2 comes after Room 1 even if one plays Room 2."*

By this comment she probably means that the user interface could have been horizontally reversed for Room 2 so that the mixing battery of room 2 was on the right hand side of the screen for both Room 1 and Room 2. Another participant was also confused by the location of mixing batteries for the two rooms:

> *"A little confusing by the choice of regulator [the mixing battery] compared to shower view, where the location was opposite."*

The location was not opposite in the two views but she may have been confused by the two mixing batteries for Rooms 1 and 2.

### 8.2.5 The Graphics

After having answered questions about the various views of the Two-Shower prototype, the participants were asked questions regarding various elements and characteristics of the prototype. The first question dealt with the graphic elements in the prototype. The questions were:

> *4 a. I think the graphics made the model*
>
> *Difficult - A little difficult - Neither difficult nor easy - A little easy – Easy*
>
> *to understand.*
>
> *4 b. Comments about the graphics:*

Part of the intention of the graphic representations was to visualize the causal structure of the underlying simulation model and its influence on the behavior.

The answers to question 4a is represented in figure 8-8. Here it must be noted that this question could be interpreted in several ways. It could refer to whether they found that it was easier to understand the underlying simulation model with the graphic user interface compared to for example equations (which the participants in this case did not have access to and would therefore be impossible to answer). Another alternative could be whether they found that the graphics aided them in understanding the prototype and its underlying simulation model. It could also be how they found the graphics useful, for example, compared to the textual elements of the prototype.

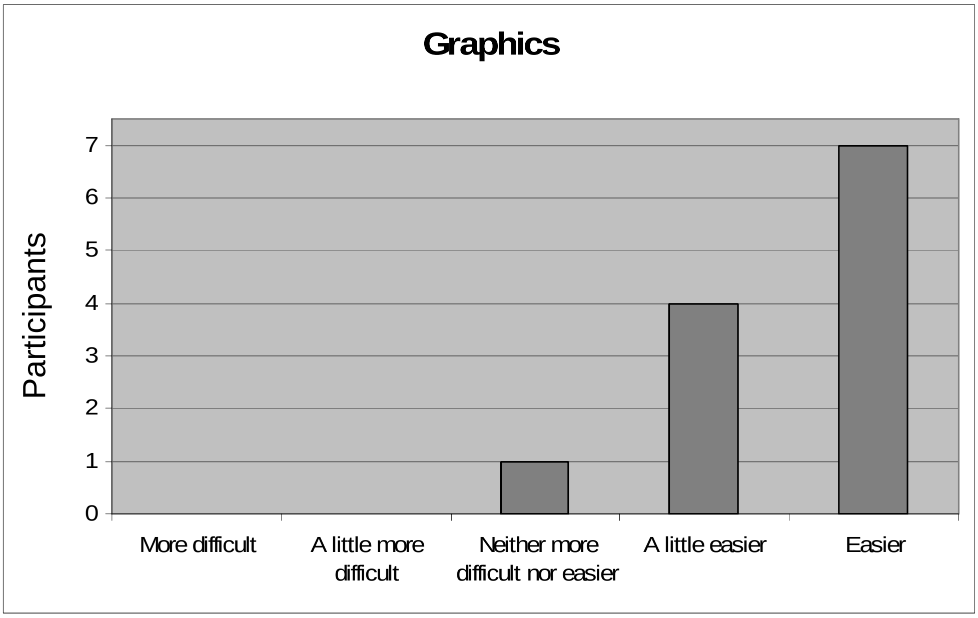


*Figure 8-8: The distribution of answers from the participants regarding how they felt the graphics aided them in understanding the prototype.*

Some general comments about the graphics were:

> *"Quite ok. Well arranged and understandable."*

> *"It was absolutely easier to understand the model with graphics."*

> *"The graphics were simple, but very relevant and explanatory."*

> *"Nice indication of temperature."*

One of the participants was not satisfied with the graphic representations:

> *"Little standardization, unclear. Some places there are lines, others not. Illogical locations. The tap should draw more attention, it almost looks disabled."*

One of the participants had the following comment on the graphics that represented the delay:

> *"That the color changed according to how hot/cold it was and that one could see how far the water had come in the pipes made it easier to choose which way to turn the knob [the mixing battery]"*

This comment indicates that one of the participants may have found that the graphics aided her in linking cause and effect and to see how input affected the behavior of the temperature in the model.

The graphics of the background in the Shower Room view represented tiles (see figure 8-4, p. 218). In the Pipe System view the background represented bricks, as if the inside of a wall (see figure 8-6, p. 222). The colors were similar in that the background of the Shower Room view was white and the background of the Pipe System view was light grey. One of the participants had the following comment regarding the graphics of the background:

> *"Fine, but further work from a 'graphical design' point of view may be done. But it works as a simple simulation. The distinction between the different views can be made clearer by light/darkness of the background 'wall'".*

Another of the participants had the following comments about the graphics:

> *"The graphics were quite ok. [I] did not understand the symbols automatically – the key for example. [I] had to hold the mouse cursor on top in order to understand what it was.[It was] nice with an explanatory text!"*

The key probably refers to the Key button that could be pressed in order to return to the Shower Room view when the Pipe System view was present (see Section 4.3.3). The button had a mouse-over text, which indicated that the participants could press it and return to the Shower Room view. The key button is again an example where the participants had problems with the buttons and navigation in the prototype and is a further indication that improvements should be made in a future version.

One of the participants commented that she wanted to know some more about the underlying model:

*“Clearly nice to have graphics to illustrate but with later simulation runs it could be useful to get more information about what lies beneath in order to understand more.”*

One solution to this would be the inclusion of a stock and flow diagram or a causal loop diagram. The inclusion of a stock and flow diagram in an interactive learning environment that is intended for learning about system dynamics was discussed in Section 5.4.

### 8.2.6 The Photo Slideshows

The participants were also asked about the influence of the photo slideshows and whether they felt that the photos supported their understanding of the system. The following questions were asked:

*5 a. I think the photo slideshows made the model*

*Difficult - A little difficult - Neither difficult nor easy - A little easy – Easy*

*to understand.*

*5 b. Comments about the photo slideshows:*

As discussed in Section 3.3 the intention of the photo slideshows was to make the participants empathize with the characters in the system and to visualize the behavior of the model in a more intuitive way (see figure 8-9). The response to the question about the photo slideshows are summarized in figure 8-10.

*Figure 8-9: Example of a photo slideshow, repeated.*

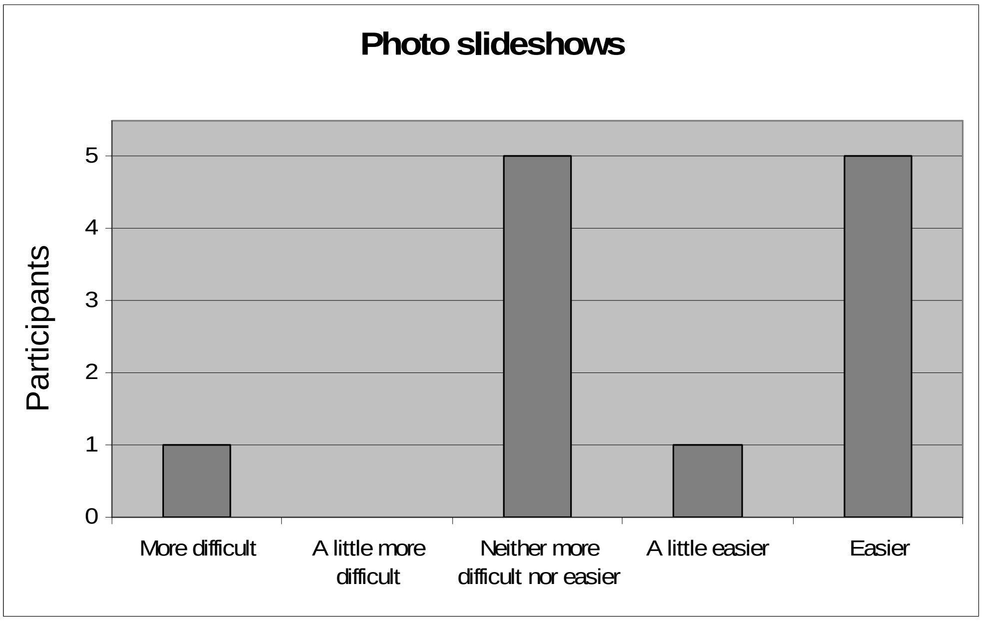


*Figure 8-10: The distribution of answers from the participants regarding how they felt the photo slideshows aided them in understanding the prototype.*

Some general comments were:

> *"Super photos."*
>
> *"Funny and good pictures. Easy to understand if it turned too hot or too cold based on the photos."*
>
> *"The photo slideshow responded quickly to my temperature changes and was therefore crucial for my temperature choices. [It was] easier than if it had been text-based. Photos are easier to perceive quickly."*
>
> *"It was good – it was well-functioning! It immediately became a real problem situation one immediately could act in accordance with and 'figure out' (everyone who has taken a shower has experienced this). So recognition of experience is also central for the intuitive access one may have as a user to a simulation."*

These comments are in line with our intentions of the photo slideshows (to elicit empathy), however, not all participants had the same view of the usefulness of the photos. In some of the sessions there was a problem with the server that went down while the client applications were still running. The wrong photos were then shown for the wrong temperatures. In this case the photos only made what was happening in the prototype confusing. One of the participants in one of the sessions in which the server went down made the following comment:

> *"Most of the time the photos made it easier to predict what temperature that was comfortable. However, when the photos did not correspond on the two computers it became confusing. When we both had a relatively stable water temperature between 36 and 38 degrees (something I think should be good) one photo showed displeasure and the temperature was unpleasant for the same room (and the opposite for the other room). That became a little confusing."*

Even though this comment was made based on an error in the program it shows that this participant used the photo series to get information about the temperature of the shower.

During the sessions, some of the participants seemed to discuss to a large degree the photo slideshows, while others seemed to spend more time discussing the graphical animations. One of the participants also commented that he did not focus on the slideshows:

> *"Focus was not put on the photos but on other feedback in the system."*

Some of the participants also had some critical comments regarding the photo slideshows:

> *"Maybe a little difficult to understand the difference between hot/cold!"*

> *"[They] changed maybe a little too often. [They are] at times difficult to interpret"*

The comment from the participant that found that the photo slideshow made the system more difficult to understand was:

> *"Has no function as long as <u>one single photo</u> is not attached to a particular temperature. Are only perceived as pictures that go on and on, and the effects are directly disturbing."*

During the discussion with this group it was mentioned that they found the intermediary or neutral photos (such as for example the showerhead with running water, see figure 8-9) which were included as transitions between the photo series confusing. This group found the neutral photos to be worthless because they did not add any value. These neutral photos were included in the beginning of each slideshow sequence so that there would be no direct jump from photos that represented hot and cold water. These were included with principles of film editing in mind. Some participants may at times have experienced frequent exposure to these neutral photos because of frequent changes in the tap setting and thereby frequent changes between cold, comfortable, and hot water temperatures. In the case of an interactive learning environment, if there is a contradiction, the focus should probably be on the representation of the underlying model rather than principles of film editing. Removing these narrative elements, of which the editing of the photo-slideshows is part, however, may also remove part of what makes the prototype interesting. Our assumption is that maintaining the attention of the participants is important for their engagement with the prototype. But this test did not give conclusive support for this assumption. A revision of the neutral photos, however, should be considered for a future version of the Two-Shower prototype.

### 8.2.7 The Animations

The participants were asked how they felt the animations, such as changes in the color of the background and changes in the color of the pipes) aided them in developing an understanding of the system (see for example figure 8-11). The following questions were asked regarding the animations:

> *6 a. I think the animations (such as for example change in color of the pipes, change in color of the background, flow of water from the hot water tank) made the model:*
>
> *Difficult - A little difficult - Neither difficult nor easy - A little easy – Easy*
>
> *to understand.*
>
> *6 b. Comments about the animations:*

Part of the intention of the animations in the prototype was to provide a closer visual link between cause and effect and for the participants to be able to observe how their changes in tap setting influenced the system. The animations were quite well received by the participants (see figure 8-12).

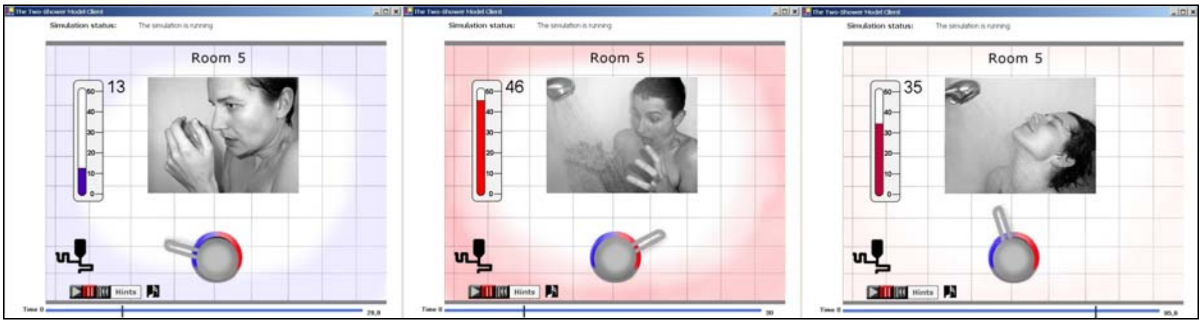


*Figure 8-11: Animation of the background in the Shower Room view, repeated.*

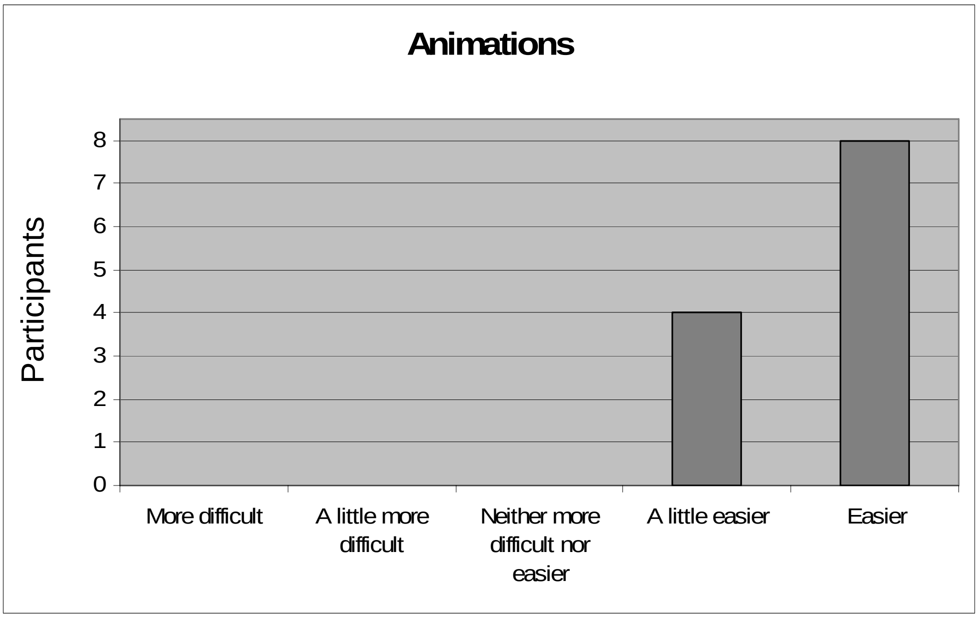


*Figure 8-12: The distribution of answers from the participants regarding how they felt the animations aided them in understanding the prototype.*

Some general comments regarding the animations were:

> *"[They] made it easier to follow the development [the behavior of the model]."*

> *"The animations are well-functioning and seem to contribute with something extra to the user interface. They actually increase the information value, and are so discrete that they do not steal the attention."*

> *"A good tool and a supplement to the Shower Room view."*[27]

> *"Function well! The right choice of colors ☺"*

[27] This participant may be referring to the Pipe System view as a supplement to the Shower Room view.

*"Very nice illustrations!"*

*"Easy to see what is happening ☺ At least ok in the pipe system. Especially in relation to making preventive adjustments of the temperature."*

*"[I] did not notice the change in the background, but the colors of the pipes were good."*

*"Especially the flow of water and the color of the pipes were good"*

One of the participants commented that there was a problem regarding finding a comfortable temperature based on the color of the water in the pipes:

*"A little unclear what 'color' the shower person preferred"*

In a future version it may be more clear which color the person in the shower prefers. A potential problem, however, could be that the participants would be more involved in obtaining the *correct color* through trial and occasional failure rather than discussing how to obtain a comfortable temperature. These are, however, only speculations.

### 8.2.8 The Behavior Graphs

The participants were also asked how they felt about the behavior graphs which were located in the upper right hand corner of the user interface (see figure 8-6, p. 222). The questions were:

*7 a. In the upper right hand corner there were some graphs indicating the behavior of the model over time. I think the graphs made the model:*

*More difficult - A little more difficult – Neither more difficult nor easier - A little easier – Easier*

*to understand.*

*7 b. Comments about the graphs:*

The intention of the behavior graphs was to provide a means for the participants to study how the behavior of certain variables had changed in relation to each other over time. The graphs were implemented only hours before the testing started, so there was no time to include units in the graphs and they were a little *ruddy* and *sketch-like* looking (see figure 8-13). The participants were also not as satisfied with the graphs as they were with the animations (see figure 8-14).

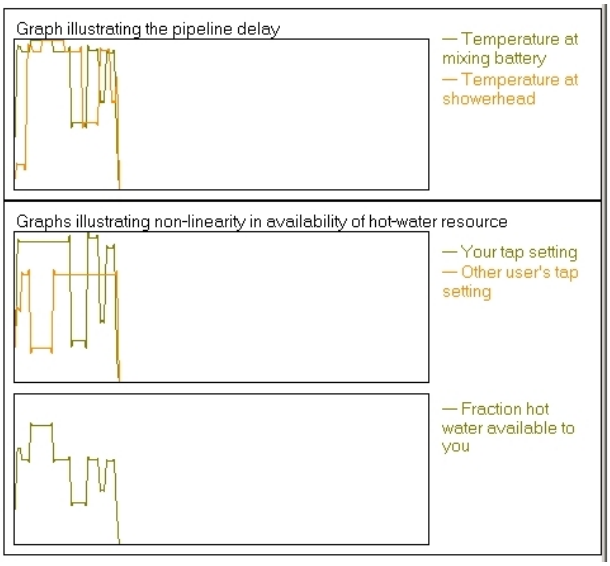


*Figure 8-13: The behavior graphs of the Two-Shower prototype.*

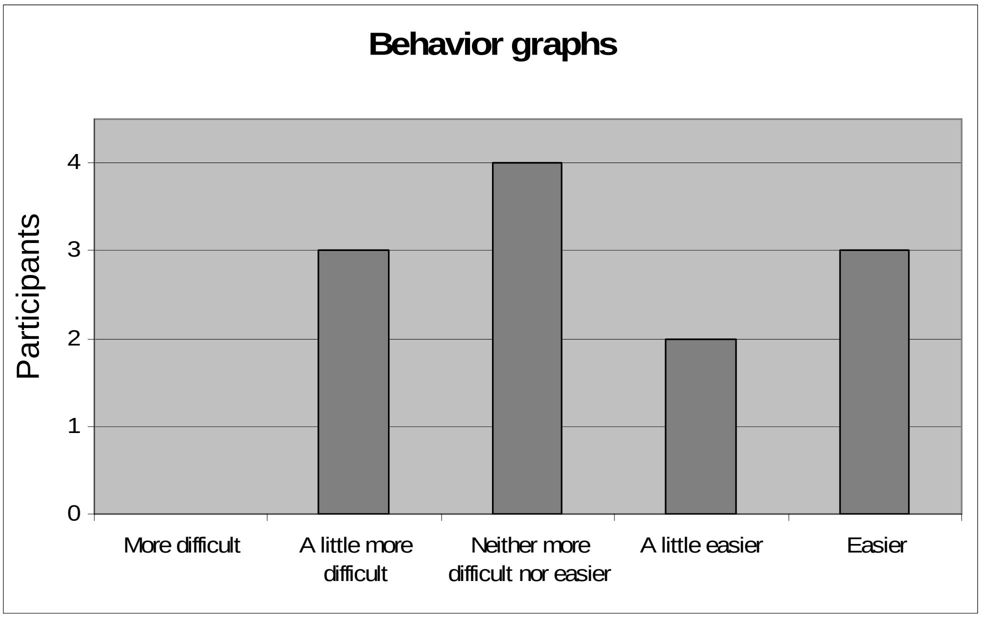


*Figure 8-14: The distribution of answers from the participants regarding how they felt the behavior graphs aided them in understanding the model.*

During the sessions, there was quite a large difference between the participants in the references that were made to the behavior graphs in their discussions. Most of the participants did not reference them directly. A few groups discussed them for a little while, but quickly moved on to discuss other parts of the interface. Two of the groups, however, used them to a large extent and discussed both the delay and the nonlinearity in the system with reference to the graphs. One of the participants from these groups had the following comments regarding the behavior graphs:

> *"Here I first answered 'a little more difficult' because it was complicated to relate to this information, but then I decided to answer 'a little easier' because the graphs, as I started to understand [the model], actually contributed to a better understanding of how the model worked."*

> *"They showed the uneven turning of each shower [tap setting] in the beginning well and it was nice to watch the graphs to control that I managed to hold a stable temperature. [I] don't think I saw the graphs until the end after I had obtained 'control' of the temperature…"*

Some comments about the graphs indicated that they were difficult to use when the prototype was running:

> *"[They were] difficult to use during the simulation."*

> *"There was not time to use them, the first impression is that I had no need for them, they look too complicated."*

> *"I chose to overlook them in order not to be disturbed in my 'decoding' of the model. The feeling was 'I don't have time for that'".*

For a novice audience it could be better to draw attention to the behavior graphs when the model is paused. On the other hand it may be important to monitor the behavior over time as the simulation is running so as to respond to undesirable changes and to monitor how the variables act in relationship to each other and how the model responds to input.

There were also some comments about how the behavior graphs failed to present the information properly:

> *"It wasn't intuitive what the graphs showed. 'A lot' of text, difficult to follow them at the same time. It would maybe have been ok afterwards in order to take a closer look at what happened, something I did not really do."*

> *"What the graphs showed was maybe a little insignificant compared to the simulation model and was to a larger degree illustrated in the Pipe System view."*

> *"I did not notice them in the beginning. [I] looked at them after the program was over/locked – they were of no help on the way. When I see them afterwards I don't quite understand them. Maybe [I am] not used to reading graphs, but they don't tell me that much. [I] cannot interpret anything from them."*

These last comments may also be influenced by the fact that the participants were not part of the intended target audience. They were not novice system dynamicists. If they had been, they may have felt that the behavior graphs were an important part of their system dynamics education.

As expected, some of the participants commented that the behavior graphs lacked units:

> *"[The behavior graphs] lack values. A little frail and boring, but they function. [I] miss linkage to the time axis."*

> *"I would have preferred time units marked on the time axis. Maybe also units on the temperature axis. The last graph I don't have a clear picture of what is showing – my portion, the remaining portion (or content of tank)."*

As mentions above, the behavior graphs were implemented in the last minute before Test 2 so there had been little time to work on the appearance of the graphs.

One of the participants had a suggestion for an improvement:

> *"Maybe a bar graph with my portion (graph 3) had been easier to understand, or that my share in the existing diagram was somehow emphasized."*

The suggestion of a bar graph is interesting because the representation of the hot water tank in the version of the Two-Shower prototype that was utilized in Test 1 was similar to a bar graph within the outline of a hot water tank (see figure 4-12, p. 130). This representation was changed, as discussed previously, after the results from Test 1.

### 8.2.9 The Textual Explanation / The Hints Button

The participants were also asked whether the textual explanation that was given when the Hints button was pressed had contributed to their understanding of the prototype. The following questions were asked:

> *8 a. You could press a button in the model to receive a textual description of the system. Did you use this button?*
>
> *Yes - No*
>
> *8 b. If yes:I think the text made the model:*
>
> *More difficult - A little more difficult - Neither more difficult nor easier - A little easier – Easier*
>
> *to understand.*
>
> *8 c. Comments about the textual description:*

As described in Section 4.3.2 the Hints button triggers a relatively detailed textual description of the structure of the system. Only seven of the participants answered that they had used the Hints button. The response on the textual explanation is represented in figure 8-15. There were only three of the seven participants who had used the textual explanation who reported to find it useful.

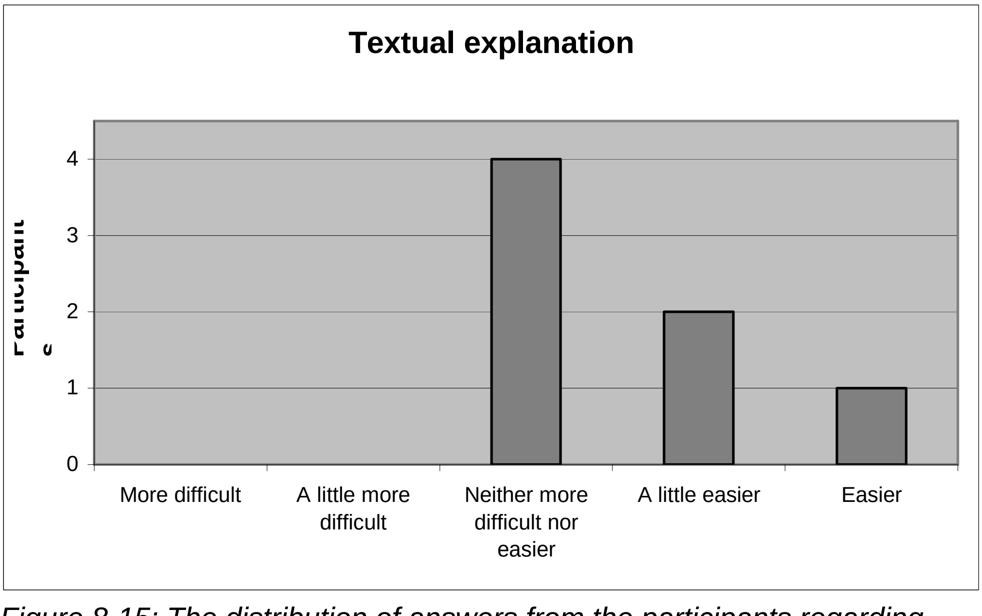


*Figure 8-15: The distribution of answers from the participants regarding how they felt the textual explanation aided them in understanding the prototype.*

A general comment about the textual explanations was (this could refer to the story generator):

> *"Short and ok explanations."*

Two of the participants commented that the text was superfluous in the context of the other information that was provided in the user interface:

> *"[I] mostly received messages I had already understood, but [it was] nice to have them confirmed."*
>
> *"The text actually describes what one expects during such a course."*

Some of the participants commented that there was too much text and that the language was too complicated:

> *"A lot of text – the language was a little difficult. [I] think the point was that it took time and that it depended on the neighbors."*
>
> *"[It was] a little interruptive to change to 'reading mode' (a relatively large amount of text) when you were otherwise performing actions. Less text with a larger font that is more included in the graphical room view [Shower Room view] and Pipe [System] view would have been good."*
>
> *"[There was] too much text. Illogical placement, deviates from the standard for where information is placed. I miss linkage to the time scale."*

In addition to the amount of text, which was also discussed in Section 8.2.2, these excerpts demonstrate that some of the participants felt that the text to a larger degree should have been integrated into the user interface, not be displayed in a separate text box. Linking the text to the time line is not impossible as the text in the Hints button only explains the structural characteristics of the system. This textual explanation could, however, be enhanced through graphics in a future version of the Two-Shower prototype. This will be discussed further in Chapter 9.

There was also a comment regarding the font size of the text in the text box, which was quite small, leaving some empty space when there were short texts to be displayed:

*"The text in the Hint-menu could be larger (fill out the field)"*

These comments show that the Hints button may be reconsidered and that the text may be better incorporated into the user interface. There should also be a total revision and reconsideration of the use of text in the user interface.

### 8.2.10 The Story Generator

The details of the story generator were described in Section 3.4. When the participants pressed the pause button important events or points in time were indicated on the timeline by yellow markers (see figure 8-6, p. 241). When the participants moved the mouse cursor over these points, textual explanations of what had happened would be displayed in the text box and the animations and graphics would change according to the state of the system at that particular point in time.

Questions were asked regarding how helpful the participants found the story generator and the text that appeared in the text box (when the simulation was paused) to be. The questions were:

> *9 a. You could pause the model and, by holding the mouse pointer over the time line, you could get an explanation of what had happened earlier in the simulation run. Did you use the pause button?*
>
> *Yes - No*
>
> *9 b. If yes: I think this opportunity made the model:*
>
> *More difficult - A little more difficult - Neither more difficult nor easier - A little easier – Easier*
>
> *to understand.*
>
> *9 c. Comments about the explanations one received by pressing the pause button:*

Nine of the twelve participants used the pause button that triggered the story generator. The answers to question 9b are summarized in figure 8-16).

In Test 1 of the Two-Shower prototype (Frotjold, 2005) there was a problem in that the text generator (the earlier version of the story generator) was not being used. The participants of Test 2 reported to find it more useful. Changes had been made to the story generator between Test 1 and Test 2, but because the sample is so small, it is difficult to state whether the increased use and satisfaction is due to these changes or because the participants in Test 2 were older than those of Test 1.

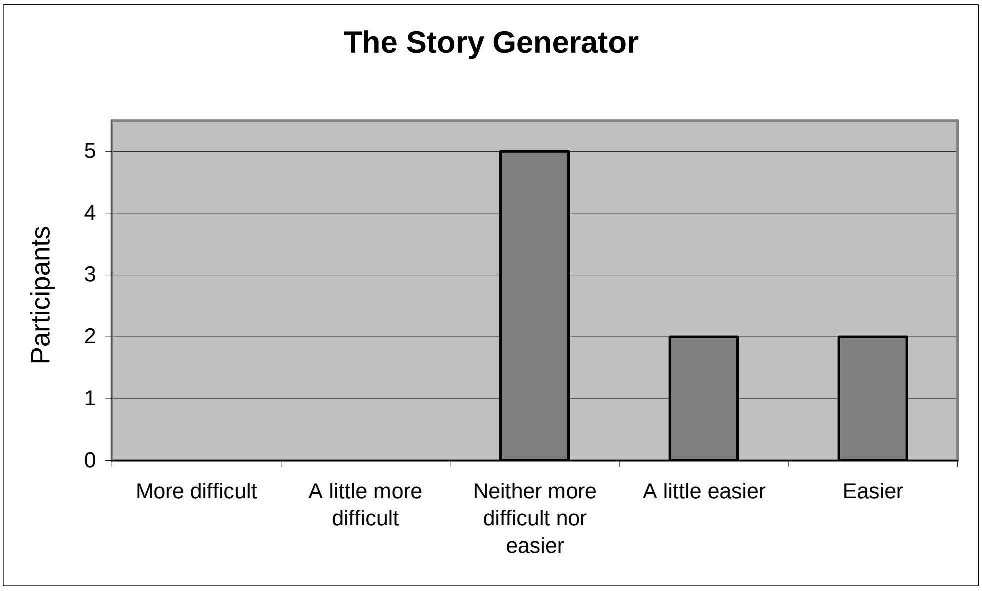


*Figure 8-16: The distribution of answers from the participants regarding how they felt the story generator aided them in understanding the prototype.*

Some positive comments regarding the story generator were:

> *"Ok explanation. It made us understand more quickly how we were going to solve the task."*

> *"[The story generator] helped me understand the model better."*

Not all participants agreed on this. One participant was confused because the text disappeared or changed when the mouse cursor was moved to another yellow marker on the time line:

> *"A little complex, because they [the text fragments] disappeared as new text was displayed they were less informative."*

Another participant reported that it was difficult to understand why the particular points in time were chosen to moment on important events:

> *"Quite ok, but it should have been better explained why the yellow lines came where they came."*

The message that was given to the participants previous to and during the simulation run was: *Remember that you can press the pause button at any time during the simulation and receive comments about what has happened so far*. Explaining that the participant's actions were recorded and analyzed with reference to the model structure and behavior could have helped the participants better understand why the text occurred at different places on the timeline. Such an explanation, however, would have added to the amount of text in the system.

The problem with the large amount of text, which was also discussed in Sections 8.2.2 and 8.2.10, was also commented on here:

> *"The text didn't say that much. It is possible that it had been better if I had used more time, but now there was too much text compared to the rest. It is possible that the text required a little more knowledge about the model/what happens in order to obtain complete benefit."*

The last part of this comment is important because it points to how the text could be made more useful with a better overview of the structure of the model. Information about the causal links and the direction of an influence on this relationship should be more useful with a more detailed overview of the model as a whole. This is, however, only a speculation.

During the sessions, there was a difference in the groups that spent time trying to figure out the story generator and those that did not. Some groups did not try it at all, while others seemed to study it relatively thoroughly. Some of the participants paused the system, had a quick look, and resumed the simulation.

As discussed in Section 8.2.3 some of the participants seemed to have problems in determining whether the simulation was running or not and suggested that the pause and play buttons could be incorporated as one button. However, observations during the sessions indicated that the problem of knowing whether the simulation was paused or not may have been because the animations with the photo-slideshows continued also in the story generator view which was present when the simulation was paused. The animations may have created the impression that the model was still running. This problem was anticipated in advance but it was not solved due to time constraints.

### 8.2.11 Organization of the User Interface

The participants were asked questions regarding the organization of the user interface with particular emphasis on the gradual introduction of complexity. The gradual introduction of complexity was, as discussed in Section 5.4 part of an intention of making the underlying simulation model gradually more accessible. The following questions were asked:

> *10 a. I think the organization of the user interface (with a gradual introduction of complexity) made the underlying model:*
>
> *More difficult - A little more difficult – Neither more difficult nor easier - A little easier – Easier*
>
> *to understand.*
>
> *10 b. Comments about the organization of the user interface:*

The response to question 10 a is represented in figure 8-17.

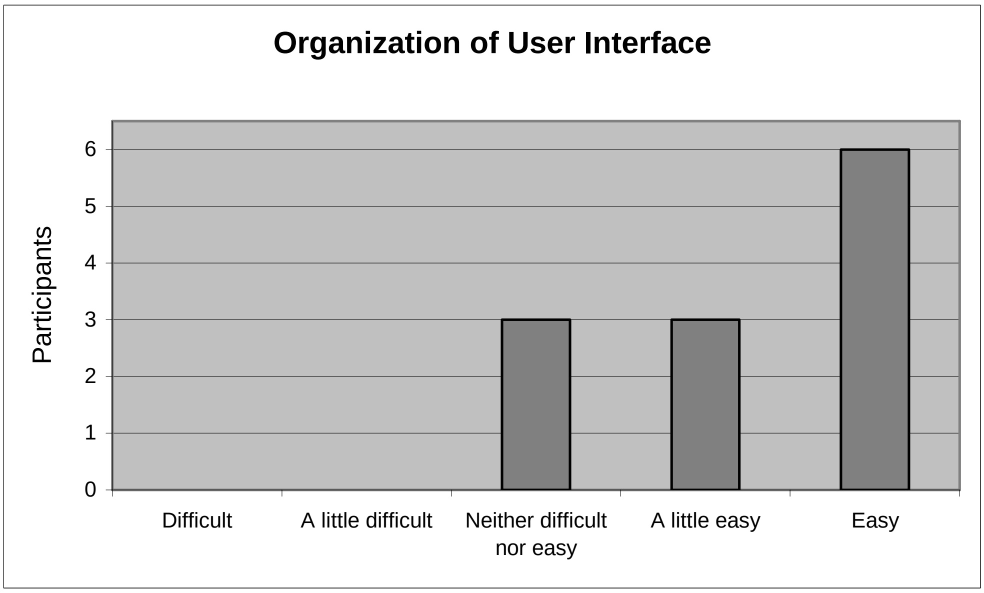


*Figure 8-17: The distribution of answers from the participants regarding how they felt the organization aided them in understanding the prototype.*

Some general comments were:

> *"Easily understood."*

> *"Nothing particular to criticize about it."*

Some of the participants commented that the Pipe System view was the most informative view. Part of the intention was to present information about the underlying simulation model gradually before providing more details. The following comment was made by one of the participants:

> *"It may have been easier to understand if the first view after the introduction had been the Pipe System view, but the way it is now there is more room for independent thinking."*

> *"It was maybe easier in the last view [the Pipe System view] – easier to get an overview when seeing both showers on ones own screen."*

Some of the participants did not have the feeling that the different views introduced more complexity:

> *"[I] didn't really feel that there was a change in complexity – maybe that means that I haven't understood the underlying model?"*

> *"I had no immediate feeling of great complexity so that means that it made it easier."*

One of the participants who found that gradual introduction of complexity made the system easier to understand had the following comment:

> *"But I still think that too much complexity was introduced at a too early point in time. This particularly applies to the graphs."*

These comments indicate that the perception of complexity of the various views and the prototype as a whole varied greatly among the participants. As designers, it is impossible to meet all the needs, requirements, and desires of all users. In a future version of the prototype the organization of the views could be changed by, for example, introducing the graphs at a later stage.

One of the participants, who found the organization of the user interface to make the prototype a little easier to understand, had the following suggestion for improvement:

> *"Ok. There could maybe have been a little more information/explanations of the graphical elements/'rollover' text with definitions / explanations."*

The lack of roll-overs and text designed to explain the various user interface elements was also commented on by another participant when addressing the graphics of the prototype. More use of rollover texts, however, could contribute to the problem of too much text in the interface.

### 8.2.12 Navigation

The participants were asked what they thought about navigation in the user interface. The questions asked were:

> *11 a. I think navigation in the system was:*
>
> *Difficult - A little difficult – Neither difficult nor easy - A little easy – Easy*
>
> *to understand.*
>
> *11 b. Comments about navigation:*

In the design of the user interface effective navigation had not been prioritized to any significant degree, The focus had rather been on transparency and the information that should be provided in the various views, not that much on how the participants could navigate between these views. It was therefore expected that the participants in the test would experience some navigational problems. There was a rather large variety in the responses (see figure 8-18).

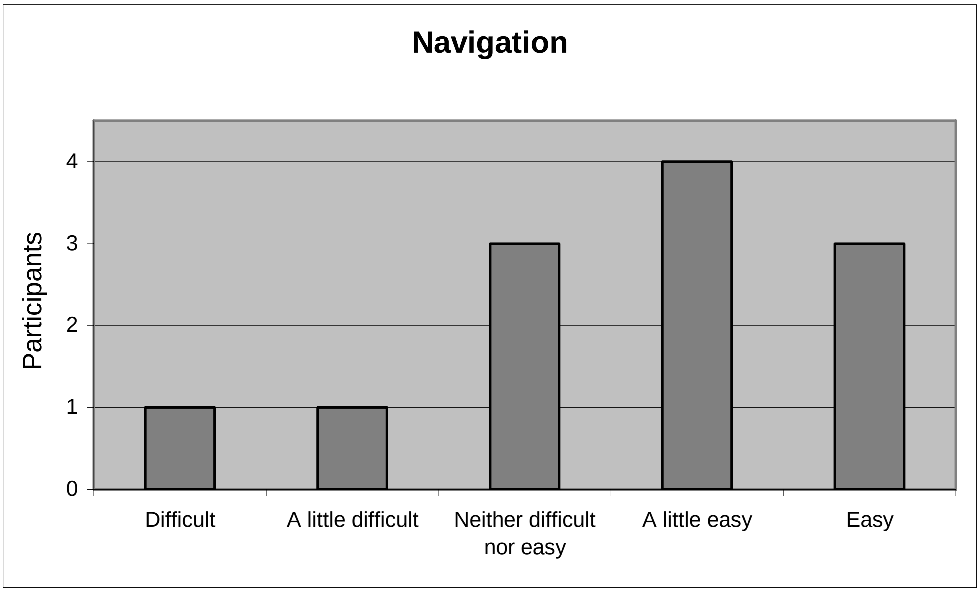


*Figure 8-18: The distribution of answers from the participants regarding how they experienced navigation in the system.*

Some general comments about navigation were:

> *"All over relatively easy, but everything is not as intuitive."*

> *"It was ok to navigate in the system. I may not have used all the functions (like the explanations [the story generator] ...). Intuitive."*

These comments and the responses that are shown in figure 8-18 are not surprising. Most of the participants experienced some navigational problems during their session . As mentioned earlier in this chapter, most of the participants had problems finding the Pipe System button and did not understand that they had to press the play button when they entered the Shower Room view. The problem of the pipe system button was also commented on by some of the participants:

> *"The buttons had nice symbols and were easy to understand. I did not, on the other hand, get that the 'Pipe System' was a button until it was pointed out to me."*

> *"Not intuitive what [which buttons] should be pressed."*

There was also a problem in understanding the Key button that brought the participants back to the Shower Room view:

> *"The key in the Pipe System view should be something that points more to a shower or hotel room. A key is often a help function in a user interface."*

One of the problems of navigation in the system was that the buttons were not located in a coherent manner (this is mostly due to the evolutionary development process). The location of the buttons was criticized by one of the participants:

> *"Buttons etc. are not standardized and placed wherever. Some buttons are 3D, others not. Some have mouse over text, others not."*

In an improved version of the Two-Shower prototype the navigational aspect of the prototype must be reconsidered and developed more thoroughly with the purpose of supporting the participants in navigating through the system.

### 8.2.13 Collaboration

As discussed in Chapter 5, one of the intentions of the Two-Shower prototype is to support collaboration. The intention was that the visualizations of the conditions and actions by the other participant could be used as a basis for collaboration and grounding. Test 2 of the Two-Shower prototype was not designed to study how the prototype supports collaboration, but the participants were asked some questions regarding what they thought about collaborating in their attempt at controlling and understanding the model. The following questions were asked:

> *12 a. I think collaboration with another person made the underlying simulation model:*

> *More difficult - A little more difficult – Neither more difficult nor easier - A little easier – Easier*

*to understand.*

*12 b. Comments about collaboration:*

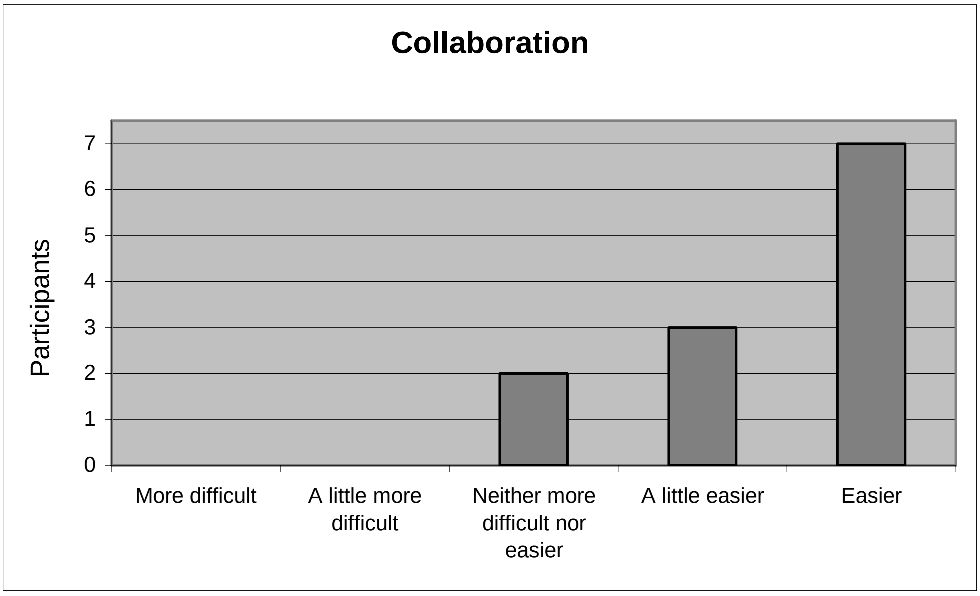


*Figure 8-19: The distribution of answers from the participants regarding how they experienced collaboration.*

The response to the first question regarding collaboration is represented in figure 8-19. A problem with the phrasing of the question ("I think collaboration with another person made the underlying simulation model more difficult/a little more difficult/neither more difficult nor easier/a little easier/easier to understand") is that there is no reference to the collaborative aspects of the user interface, only the fact that they could collaborate with another person.

Most of the comments about the collaboration were positive:

> *"Good collaboration – good discussion."*

> *"It went well and easily. The second time we ran the model we had understood the model and were more like playing with it."*

Another comment was:

> *"Nice to be able to communicate with the other [participant] – get explanations and to try to express my understanding of the model to the other [participant]."*

This comment corresponds to the intention that the participants should use the prototype as a base for discussion.

Some further comments were:

> *"Without collaboration it would have taken a long time to solve the task. Very nice to be able to collaborate when we are able to see each other's screen."*
>
> *"Nice to hear how the conditions were in the other shower while I was focusing on obtaining control of the temperature in my own shower."*
>
> *"We could each manipulate our knobs [mixing batteries] and discuss why we observed what we did."*
>
> *"The collaboration made it possible to see that the actions of the other had an effect, but it was less clear exactly how the collaboration had an effect (direct effect)."*

These comments show that the possibility to collaborate with another person was generally well-received by the participants. Two of the participants, however, were not as positive about the collaborative aspect of the sessions:

> *"It neither made it easier nor more difficult. I think it would have been just as easy to understand alone."*
>
> *"We didn't really have a task, so it was just small talk, not really collaboration."*

One of the participants pointed out during the sessions that the collaboration would have been different if they knew each other before using the model (although some of the participants did). This is an important point, and will of course influence the success of collaborating. In this test for example, some groups consisted of one Master's student and one PhD student. The difference in academic level of the participants may also have affected the conditions for collaboration. In a real educational setting, the participants in such an interactive learning environment are likely to have at least met before and will be part of the same class or group.

## 8.3 Part Two of the Post-Questionnaire

As discussed previously, the participants were asked some questions about the underlying model after the sessions. The original purpose was to find indications of characteristics of the underlying model that needed to be represented differently or made more explicit in the user interface. The problem is that in using the answers to these questions as indicators of the quality of the user interface, there is an implicit assumption that the answers are representative of the understanding of the underlying model gained by the participants. There is also the assumption that the questions are phrased in such a way that the participants can provide *correct* representations of their understanding of the model. The participants may have understood the basics of the underlying model even though they were not able to describe it in words or did not mention particular elements in the system. In spite of these problems, I will discuss the answers to these questions and try to relate them to the first part of the post questionnaire and observations during the sessions. Although speculative, an indication of design functionality will also be discussed briefly.

In the first part of the questionnaire, the participants were asked whether they found the underlying principles of the simulation model difficult to understand. The question was phrased accordingly:

> *I think the underlying principles of the Two-Shower prototype were:*
>
> *Difficult - A little difficult - Neither difficult nor easy - A little easy – Easy*
>
> *to understand.*

The answers to the question are represented in figure 8-20. These answers, of course, do not indicate whether the participants actually developed an understanding of the underlying model but are only their own reports on the level of difficulty that they experienced.

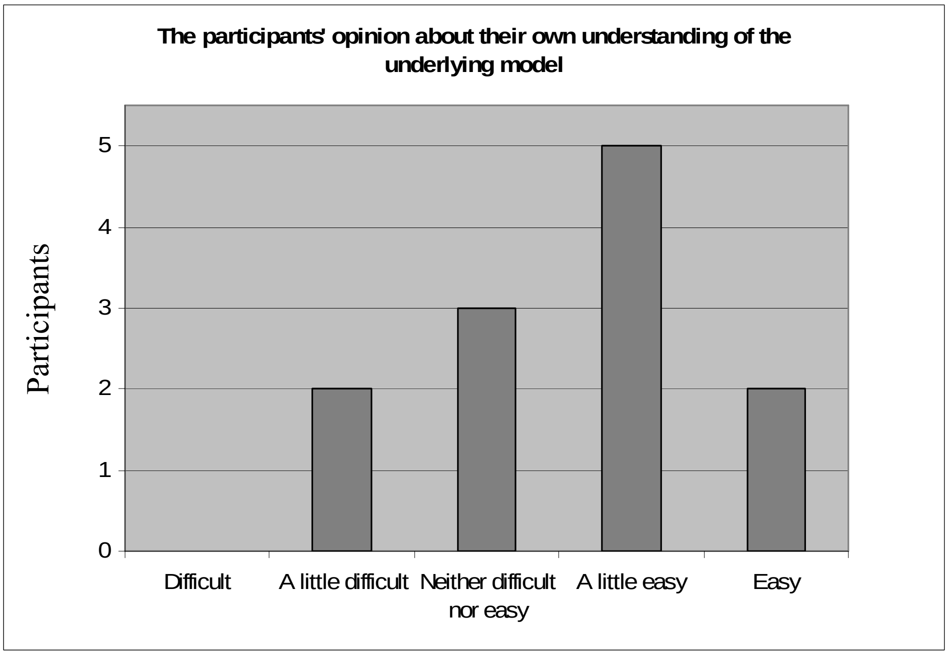


*Figure 8-20: The participants own view on the difficulty of understanding the underlying principles of the Two-Shower prototype.*

### 8.3.1 Oscillations and Changes in Shower Temperature

In the first question of the second part of the post-questionnaire the participants were asked to explain why the system oscillated:

*"Can you explain why the temperature oscillated?"*

Almost all of the participants asked what *oscillates* (*oscillerer* in Norwegian) means. This term is generally not used in everyday Norwegian and is therefore a typical example of how researchers should be careful to use jargon that the participants may not understand. If the participants had been part of the target group of fresh system dynamics students they would have been more familiar with the term.

An answer to why the system oscillates could be that the model contains two balancing loops (the adjustments made to each shower based on their own temperature compared to their desired temperature) and a reinforcing loop in which the two showers strive to obtain more or less of the hot water resource (see figure 2-11, p. 65 for the causal loop diagram of the original Two-Shower prototype). More specifically, there is a delay from a change is made in the tap setting until the water with the new temperature reaches the showerhead. If the delay is not taken into account and exaggerated changes in the tap setting are made it will be impossible to find a comfortable temperature. The hot water supply is also shared with another participant/shower, who experiences the delay. The access to the hot water supply depends on the tap settings of both showers. If one shower turns the tap up, there will be less hot water for the other shower, which, after some time, leads to a lowering of the temperature in the other shower. When the other person turns her tap setting up, the temperature of the first shower will consequently go down and the first person will try to adjust the temperature gap by turning the tap setting even higher (the reinforcing loop). The combination of the two balancing loops, the reinforcing loops, and the delays that are not taken into account will result in oscillatory behavior.

The second question concerned why they could experience a change in water temperature even if they had not changed the tap setting for a while and was as follows:

> *"If you did not change your tap setting for a while you could still experience a change in temperature. Can you explain why?"*

This question refers to the nonlinearity in the access to the hot water resource and to the delay from a change in tap setting is made until the water with the new temperature reaches the showerhead. The short answer to the question could be that the temperature changes because the other participant changes his tap setting or that it takes time from a change is made in the tap setting until the temperature at the showerhead changes.

The participants had also not been introduced to system dynamic terms (delays, nonlinearities etc.), which they probably would have if it was used as part of an introductory course in system dynamics, - depending on what stage there were at in the course. Although most people would be expected to have experience with sharing a hot water tank, the participants did not have experience with complex systems theory or system dynamics and they were therefore not expected to provide detailed answers regarding the relationship between the structure and behavior of the model using the terms of system dynamics.

The focus of the analysis of the text was to see whether they referred to the delay between the tap and showerhead, the delay between the other persons tap and showerhead, and the nonlinearity in the access to the hot water resource and how this was determined by the tap settings of the two showers.

None of the participants referred to only the delay between the tap and the showerhead (see figure 8-21). Four of the participants only referred to the nonlinearity in the access to the hot water resource and not the delay between the tap and the showerhead. Eight of the twelve participants referred to both the delay between their tap setting and the showerhead and the nonlinearity in access to the hot water resource.

| | |
|---|---|
| Number of participants that only referred to the delay | 0 |
| Number of participants that only referred to the nonlinearity | 4 |
| Number of participants that referred both to the delay and the nonlinearity | 8 |

*Figure 8-21: The number of participants referring to the delay and nonlinearity.*

The following is an example of a reference to both the delay and nonlinearity in the system:

> *"The temperature in each showerhead is dependent on both mixing batteries. All movements have a delay in the Pipe System."*

Another example is:

> *"The temperature was dependent on one's own and the neighbors changes in tap setting. Very hot at the neighbor's also lead to less hot water for me even if I do not change the tap setting. Much oscillation up and down is probably caused by impatience – expected that the temperature should change immediately and moves back and forth to provoke a change."*

Only two of the participants also referred to the delays that were experienced in the showers of the other participants:

> *"Time difference/delay from I made changes at the battery to the water reached the showerhead. Shared hot water tank, dependent on the other shower. This enforced the effect of the time delay -> there are two people that experience it."*

This participant, as well as all of the other participants, does not use the terms *delay* (*forsinkelse* in Norwegian) or *nonlinearity* (*ulinearitet* in Norwegian) in her description of the system. Instead, she uses descriptions like "*dependent on one's own and the neighbors changes in tap setting* " and "*caused by impatience – expected that the temperature should change immediately"*, here referring to the fact that that one should have patience because the temperature does not change immediately. As mentioned previously, however, the participants did not have experience with complex systems or system dynamics and could, therefore, not be expected to use complex systems or system dynamics terminology. They were as also mentioned previously expected to have experience with sharing a hot water resource.

The representation of the hot water tank was also hardly discussed or referred to during the sessions, although several groups discussed the relationship between the two showers. None of the participants, however, described only the delay and not the nonlinearity in the second part of the post questionnaire. It may also be that the pipeline delay was visualized relatively explicitly and that the participants therefore took it for granted.

An example of an answer that only refers to the nonlinearity is:

> *"[The system oscillated] because we used the same hot water resource. When I wanted a colder temperature it turned warmer in the other shower because it then got access to more hot water and the other way around. I think…"*

Even though there was much discussion about the sharing of the hot water, as mentioned before, almost none of the participants discussed the yellow lines that were flowing out of the hot water tank at varying speeds during the sessions. These were intended to portray how the flow of the hot water changes according to the tap settings of the two showers. The participants also did not emphasize this part of the user interface in the answers provided in the post-questionnaires. It seemed like some of the participants were still able to explain the nonlinearity and some even explained it in detail.

The observations that were made during the sessions and the focus of the answers in the first part of the post-questionnaire differed from the answers that were provided in the second part of the post-questionnaire. Although speculative, it could be that the visualization of the variations in flow of the hot water (the yellow lines that change speeds) is unnecessary, and that the visualizations of two tap settings is sufficient to visualize the changes in the access to the hot water resource. As noted before, part of the intention was to use a problem that would be familiar to many and that most people had experienced in real life. Detailed visualizations of the nonlinearity for other participants, less familiar complex systems, could be necessary. These are, as noted above, only speculations.

In this chapter I have discussed the results from Test 2 of the Two-Shower prototype. The intention was to test the design ideas and to receive feedback that can be used as a basis for improved design. Suggestions as to how the prototype may be improved have been mentioned briefly here and will be discussed more thoroughly in the next chapter. Chapter 9 also contains a discussion about the research design for Test 2 of the Two-Shower prototype and lessons learnt for further work on the evaluation of the prototype. Some of the feedback that was provided by the participants in the Test 2 will also be related to the Quito prototype discussed in Chapter 10.

# 9. Future Work on the Two-Shower Prototype

## 9.1 Introduction

In the last chapter, I discussed the results of Test 2 of the Two-Shower prototype. Several issues that needed further consideration were found. The focus of this chapter is on lessons learnt and future work on the Two-Shower prototype including further evaluation. Section 9.2 picks up the design issues that emerged from Test 1 and Test 2. I consider some of the general usability and navigational problems of the prototype and present suggestions for improvements. Section 9.3 is a discussion of lessons learnt regarding the research design for Test 2. Future testing of the Two-Shower prototype is suggested in Section 9.4.

## 9.2 Lessons Learnt and Future Work on the Design of the Two-Shower Prototype

Test 2 of the Two-Shower prototype showed that there was both, a large difference in which parts of the user interface the participants chose to focus on and which parts they reported to aid them in developing an understanding of the system. There was also a large difference in the feedback from the various participants about the different components of the prototype. For further development of the Two-Shower prototype, one would therefore have to choose to focus on only some of the feedback from the test and to suggest improvements based on these. Some of the participants, for example, found that there was too much text while others wished for more information and explanations (see Sections 8.2.9 and 8.2.11). In addition, some had difficulties interpreting the behavior graphs of the prototype while others reported that these aided them in developing an understanding of the underlying model (see Section 8.2.8). A problem with basing suggestions for future design on the results from Test 1 and Test 2 may be that the participants of the tests were not students who were attending a course in, for example, system dynamics. Their feedback was therefore characterized, not only by the lack of experience with complex systems (which novice system dynamics students would also lack), but by that the fact that the participants probably did not have a wish or intention to learn about complex systems.

The suggested changes must be based, of course, on the intended audience of people who are going to learn about system dynamics. The problem here is that also this audience differs in terms of experience with complex systems. There may be a difference between those who have not yet had their first lecture, those that have, and those that has studied system dynamics for a couple of months. The difference in experience could influence the design decisions of a future version of the Two-Shower prototype. How the prototype is used as part of an actual educational context will also have implications for the design of the prototype.

Some of the elements that received mainly positive comments were:

- Collaboration
- Visualization in the Pipe System view.
- The use of color, animations, and photos to communicate conditions of variables in the system.

As discussed in Chapter 9 most of the participants answered that they found the possibility to collaborate helpful. The Pipe System view, where the delay and the water with changing temperatures were represented, was also generally well received among the participants. Several of the participants also expressed that they were satisfied with the indication of temperature by color, photo slideshows, and animations.

Some lessons regarding general usability issues have been learnt, however, and should be dealt with in an improved version of the prototype. The tests showed that the following elements need to be revised:

- Control of the tap setting.

- Location and appearance of the buttons.
- Size of text font.
- Help or information about objects with mouse-over text.
- Improved design of behavior graphs.

These problems are not discussed in detail in this chapter. They are mainly programming issues and can be dealt with relatively easily without a complete reconsideration of the visual design of the interface. These problems must be taken into account if detailed plans for further development of the Two-Shower prototype are to be made.

Through Test 2 of the Two-Shower prototype some problems more closely related to the visualization of the underlying model also became apparent. These included:

- Too much text-based information introduced at an early stage.

- Too many neutral transition photos, with no information about the water temperature, are displayed.

- Few of the participants seemed to focus on the graphics which represented the nonlinearity in the sharing of the hot water resource. New design of graphics that represent this nonlinearity may be considered.

- Most of the participants reported that the story generator neither made the underlying model easier nor more difficult to understand. In addition, the story generator was found to contain too much text and that it was confusing that the text disappeared. In addition one participant commented that it was difficult to know why the yellow lines on the time slider appeared at certain points.

- Almost all of the participants had problems with navigation in the prototype.

- Some of the participants reported that it was difficult to monitor and make use of the behavior graphs as the model was running.

- One participant reported that she was confused because in the client applications Shower 1 was on the left of one and on the right of the other. The same was the case for Shower 2.

The following subsections consider how the issues summarized here may be improved in a future version of the Two-Shower prototype. Note that these are not yet implemented in the prototype and that the figures and screenshots merely represent suggestions regarding the design of the future prototype. A new organization and navigational system is suggested for the prototype, in addition to several new views that are intended to provide additional levels of graduated complexity. It is also important to note that not all design ideas for the improved version of the prototype are illustrated by screen shots and graphics even though they may be discussed in the text.

### 9.2.1 Suggestions for Future Navigation and Revised Organization of the User Interface

Test 2 of the Two-Shower prototype showed that one of the participants reported that too much information was introduced at an early stage (see Section 8.2.11). Although this was only reported by one participant, considering this problem in a revised version of the Two-Shower prototype, some additional views in which information about the prototype was better distributed could be developed. The participants may thus use a number of views that represent characteristics of the underlying model in different ways.

Another lesson learnt from Test 2 of the Two-Shower prototype was that there were some navigational problems, especially in switching from one view to another (see Section 8.2.12). The participants did not seem to notice the pipe button that could be pressed to enter the Pipe System view and the key button that could be pressed in order to return to the Shower Room view.

The participants in Test 2 also had some problems of finding the control buttons in the system and understanding that they had to use these to control the system. The control buttons would therefore be enlarged and moved to the lower left hand corner in a future version of the prototype.

As discussed in the previous chapter, almost all of the participants in Test 2 had problems regarding turning the tap setting. The tap setting should therefore be changed both in order to make it look more like a real tap setting and with the purpose of improving its functionality. The handle in the current version of the tap could be replaced by a yellow indicator that shows where on the scale between cold (blue) and hot (red) the tap setting is placed.

To ensure the implementation of graduated complexity not all the views should become active immediately. The views may be implemented so that they become active only gradually or that for the first run the participants are only able to activate the Pipe System view, in the second run they may activate both the Shower Room view and the Pipe System view and so on. The gradual activation of views, however, may be annoying for a participant who has run the model before and wishes to study the details during the second run. Alternatively, the participants could indicate whether they had run the simulation previously and thereby be provided access to views according to their previous experience. Another option is to give the participants assignments where they are told which view to focus on at certain times. The participants may be asked, for example, to run the model one time before entering the Pipe System view.

Some of the suggested views should only be active when the system is paused. These would therefore appear as dimmed when the simulation model is running but become active when the model is paused. If the mouse cursor is placed over these views on the menu when the model is running, there could be a mouse-over message indicating that the view may only be accessed when the model is paused.

Based on these problems I suggest a different organization of the views for a future version of the Two-Shower prototype (see figure 9-1, p. 263). An improved version I suggest would include the following views: the Introductory view, the Shower Room view, the Pipe System view, a new Hot Water Tank view, a revised Story Generator view, a separate Textual Explanation view, and a Stock and Flow or Causal Loop Diagram view. In addition, to aid navigation in the prototype, there could be a small text field in the lower right hand corner of each view with a short explanation of what the view is representing. For the remainder of this section I discuss the various views suggested for a future version of the Two-Shower prototype. A possible downside of dividing the prototype into a relatively large number of views may be that it could be difficult for the user to link the different parts of the underlying model. It would therefore be important to seek to integrate the different components with an appropriate navigation tool.

### 9.2.2 The Revised Introductory Story View

One of the lessons learnt regarding the Hotel view was that the introductory story did not seem to be incorporated well enough into the user interface (see Section 8.2.2). Suggestions that pictures could be included to break up the text were also presented.

In the revised version of the Two-Shower prototype the Hotel view would therefore essentially be the same as before. Rather than presenting the introductory story in a message box, however, it would be incorporated as a separate view in the prototype. Into the story of the person who arrives at a hotel there could also be photos or graphics of a person who arrives at a hotel and receives a key. The next picture could show a photo of the character who is undressing with the shower in the background. In this way the text would be illustrated by graphics, adding to the narrative characteristics of the interface, and hopefully present the introduction of the problem in a more integrated manner. Adding more visual illustrations to the text could also contribute to making the interface less text-based. As discussed previously, the large amount of text was commented on by participants of both, Test 1 and Test 2.

### 9.2.3 The Revised Shower Room View

A further way of addressing information distribution is through changes to the Shower Room view. This view could be simplified to show only the shower room (see figure 9-1) and the behavior graphs could be introduced at a later stage. Access to the story generator could also be denied at this stage. A staggering of information in this way would also be more in line with the initial intention of a gradual introduction to the complexity of the system.

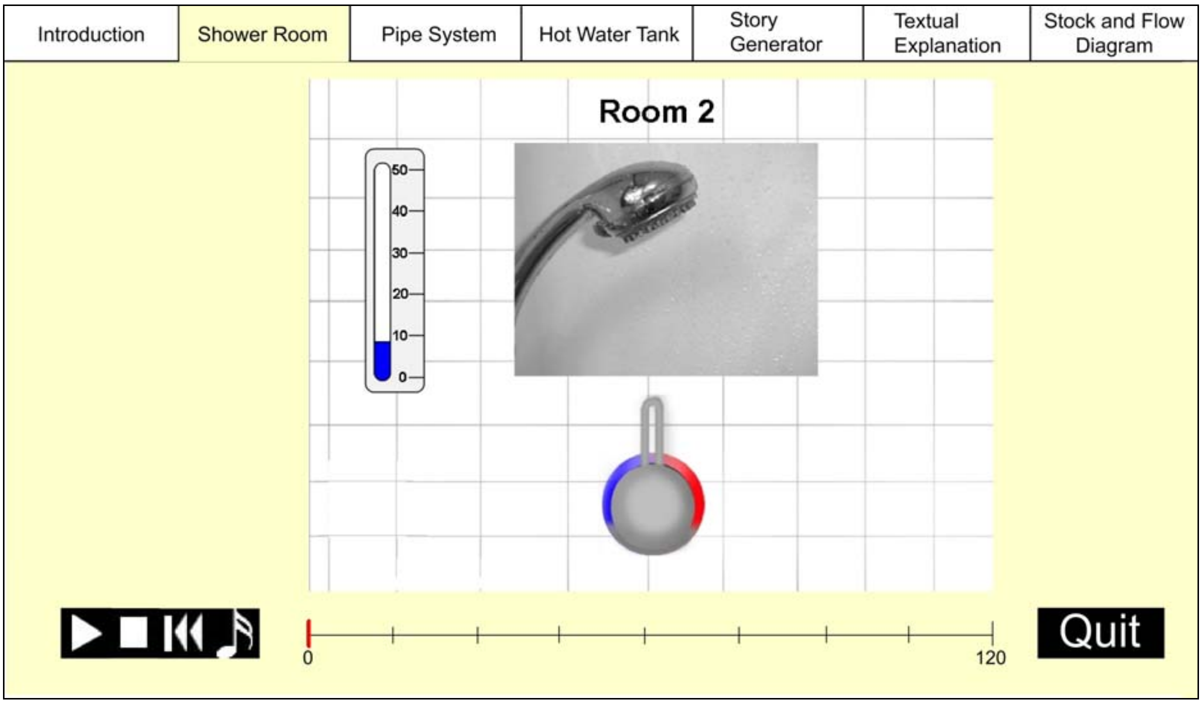


*Figure 9-1: The Shower Room view without the behavior graphs.*

### 9.2.4 The Revised Pipe System View

As discussed in Section 8.2.4 one of the participants in Test 2 had problems understanding which mixing battery in the Pipe System view belonged to her and which belonged to the other participant. I also noted how one of the comments by another participant may be interpreted as if she also was confused regarding which mixing battery belonged to whom. One of the participants suggested reversing the interface so that Room 1 was on the left and Room 2 was on the right hand side on both screens – a suggestion that should be considered in a revised version of the prototype.

One participant commented that there should be a larger distinction between the background in the Shower Room view and the Pipe System view. The background colors could also be reconsidered in a revised version of the prototype. Changing the background colors to emphasize the distinction between the two views, however, may not be necessary as also the proposed change in navigation may create a clearer distinction between the different views of the prototype.

### 9.2.5 A New View: The Hot Water Tank View

As discussed in Chapter 8 the participants of Test 2 of the Two-Shower prototype did not discuss or comment on the yellow lines that represented the changes in the access to the hot water resource. Due to the design of the test, it is impossible to know whether the participants paid attention to the yellow lines or not. Furthermore, even if they had noticed the yellow lines, the test did not focus on learning so it is impossible to evaluate whether the participants understood the nonlinearity of the system. In addition, the test design also did not allow me to detect whether they actually did pay attention to the yellow lines or not.

In a future version of the prototype, however, a separate view, a *Hot Water Tank* view, could be integrated into the user interface. The purpose of this view would be to emphasize the details related to the nonlinearity. These details, I believe, are important in developing an understanding of the underlying model.

### 9.2.6 The Revised Story Generator as a Separate View

The following are two suggested versions of the story generator that could be implemented in a future Two-Shower prototype. One suggestion would be to simplify the appearance of the story generator by developing a new view where the text-part of the story generator is associated with behavior graphs (see figure 9-2). A second suggested version relates to narrative aspects of both text and graphics.

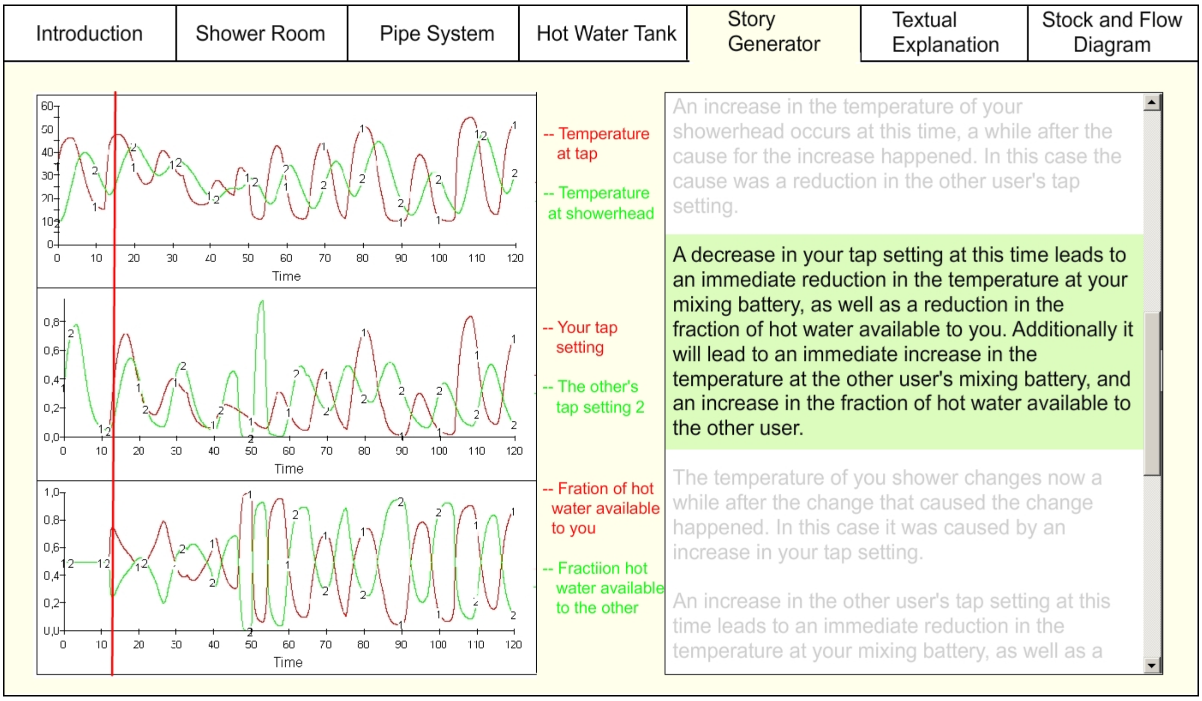


*Figure 9-2: The revised Story Generator view.*

Because both, the text part of the story generator and the behavior graphs, contain historical information about what has happened in the model, a suggestion could be to place them together, in a separate view. The intention of the view would be to support the participants in analyzing the behavior of the underlying model over time in relation to its structure.

Some of the comments from Test 2 indicated that the story generator contained too much text and that it was confusing that the text disappeared the moment the mouse cursor moved away from the marker on the time line (see Section 8.2.10). A suggested revision of the text of the story generator could be to generate one text, containing all the text fragments, but where only the current, relevant text is in focus (see figure 9-2). When the mouse cursor is moved to a different point on the timeline of the behavior graphs a different section of the text would come into focus. The text would roll up or down according to the movements of the mouse over the timeline. The proposed revised view would still contain a large amount of text. The simulation would be paused, however, whenever it was visible, and the participants could take their time to study the behavior and how it is related to the structure.

A suggestion could also be to include a line that traverses the behavior graphs (see the red line traversing the behavior graphs in figure 9-2). The participants could drag the red line along the time line and the text in the text box would change according to the time. The purpose would be to aid the participant in comparing the historical behavior of the system and in linking important events to the behavior over time.

One of the participants in Test 2 commented that it was difficult to know why the text of the story generator appeared where it did. In an improved version of the prototype there could be a text or a help button in connection to the Story Generator view explaining that the text is generated based on the underlying simulation model and the previous actions of both users. It should also be made clear to the participant that the intention of this view is to point out some important characteristics of the underlying model. The inclusion of such an explanation would add, however, to the amount of text in the user interface.

A second alternative for further development of the story generator would be to create more interesting and varied texts and graphics. The text generator could be extended by use of various text structures. The current version uses the same text structure and similar wording every time. More *exciting* text fragments, combined in a manner that makes the story more interesting, could be implemented. The logical construction of such a story generator would be highly complex, however, as would its implementation. This issue is therefore not elaborated further here. If, however, a more complex story generator were to be implemented, it would be necessary to consider and study how it supported the participants in understanding of the underlying model, and not just whether it made the prototype more *fun* or *exciting* to use.

### 9.2.7 The Textual Explanation as a Separate View

Only seven of the twelve participants in Test 2 of the Two-Shower prototype used the Hints button that gave a textual explanation of the user interface (see Section 8.2.9). Some of those who did use it remarked that it contained too much text and that they were not ready to turn to *reading mode* in the middle of a simulation run. Some of the participants also found it difficult to relate it to the graphics of the prototype because the text was implemented in a separate text box in the lower right hand corner of the user interface.

In a revised version of the Two-Shower prototype, the textual explanation could therefore be implemented as a separate view that could only be activated when the simulation model was paused. The text could be supported by graphical representations that were similar to the other graphic representations in the interface. The graphics would also contribute to breaking up the text segments. There would thus be a text and graphics-based explanation of the main structural characteristics of the underlying simulation model.

### 9.2.8 A New View: The Stock and Flow / Causal Loop View

As discussed in Section 5.4 when learning about system dynamics, learning the different representation techniques of the research field is an important part of the education. If the Two-Shower prototype were to be used in system dynamics education it should also include a view with either the causal loop diagram of the two-shower model or a stock and flow diagram. A separate view with one or both of these representations could therefore be included in the revised version of the prototype. There could also be an opportunity to study the model equations and it could be possible to click the elements in the stock and flow diagram and obtain information about equations, units, and definitions of variables.

This section has considered a number of suggested improvements for a future version of the Two-Shower prototype. In the next section, I discuss lessons learnt from the test design of Test 2 of the Two-Shower prototype.

## 9.3 Lessons Learnt from the Test Design of Test 2 and Ways Forward

This section includes a critical discussion of the test design of Test 2 of the Two-Shower Prototype. One of the main problems with the research design was that it was not pre-tested before the main test was performed. Ideally, a pre-run-through with one or two groups should have been made performed, and the test design should have been evaluated and changed based on the experiences from that pre-test. Some problems with the test design became apparent during the first couple of sessions and, rather than using the results from these sessions as part of Test 2, they could have been used to improve the design by, for example, rephrasing some of the questions in the questionnaires.

### 9.3.1 Lessons Learnt from Choice of Participants

Ideally, the group of participants should have been a little larger than twelve individuals, but it may not need to be much larger. The group of twelve was able to provide useful feedback that could be used for further development.

As discussed in Chapter 8 some of the participants reported that they failed to see the purpose of the prototype. This may have been caused by the choice of participants for the test. Ideally the participants should have been novice system dynamics students and part of an educational program or course in system dynamics. Because it was late in the semester and no such students were available, I chose to use graduate students among who most were students in information science. With a group that was closer to the intended users it could have been easier for them to recognize the purpose of the prototype and to relate it to the course they were taking. An advantage of using participants with a background in information science, however, may be that they were trained in user interface design and that they, therefore, could provide valuable feedback regarding the interface from such a perspective.

Another problem regarding the choice of participants may have been that I knew some of them beforehand. According to Nielsen (1993) users may also have a tendency to provide the answers they think are expected or wanted. This may especially be the case when the participants know the researcher. Again, the choice of participants was a matter of timing because many students were either busy with exams or had gone home for the summer holidays.

With the group that was used for Test 2 it was difficult to distinguish between the different participants as is recommended in usability testing because most of them had no experience with system dynamics and they were all graduate students, most of them in information science (see Nielsen, 1993). If I were to do the test again with actual system dynamics students, I could choose students who were about to start a course in system dynamics, students who were in the middle of a course in system dynamics, and students who had just finished or were just about to finish a course in system dynamics. With such a classification of participants, it could possibly be easier to analyze their responses according to their level of proficiency.

### 9.3.2 Lessons Learnt Regarding the Location

Test 2 of the Two-Shower prototype was conducted in my office at the Department of Information Science and Media Studies. These are only speculations, but this choice of location could have both positive and negative effects. Test 1 was performed in a computer lab. Frotjold found that the several of the participants seemed nervous as if they considered the it as some kind of exam (Frotjold, 2005). The location in an office would be more personal and they would not maybe be as nervous as in a lab setting. On the other hand, a more personal location could prevent the participants from being honest in their critique about the prototype, giving it more positive feedback than otherwise. These are, as mentioned, only speculations.

At a later stage in the development process it would probably have been an advantage if the prototype had been tested as part of a course in system dynamics.

### 9.3.3 Lessons Learnt from the Pre-Questionnaire

The main lesson learnt regarding the questions asked in the pre-questionnaire was that they were difficult to use and that they were especially difficult to use in the discussion about the answers from the post-questionnaire. In addition to being a small test group, the test group was relatively homogeneous since they were all graduate students and only two of the students had taken a course in system dynamics or a course that was in part about system dynamics. There was therefore little basis for interpreting the answers in the post-questionnaire based on, for example, the lack of or completion of a mathematics course or their familiarity with system dynamics. In a future test of the Two-Shower prototype I would probably also use a pre-questionnaire but with a different test group I could find more use of it. If the participants were taking a course in system dynamics I could, for example, analyze their answers in the post-questionnaire based on the stage they were at in the course, reported in the pre-questionnaire.

### 9.3.4 Lessons Learnt from Observation during the Sessions

As discussed previously, notes were taken during the sessions. One of the problems of not having tested the test design beforehand was that it became clear that I was unsure what I should look for or note in particular. If the test design had been properly tested, I could have identified some of the main problems that I should be prepared to look for or some main actions or discussion topics on the part of the participants that was worth noting. In this way I would have had more comparable notes from the various sessions and they would have been more useful for the analysis.

As discussed in Chapter 7 the participants were videotaped during the sessions. Because of time constraints, most of this video material remained unused. If I were to do the test once more, I would have used the video recordings to a larger extent to observe which parts of the user interface they chose to focus on and their communication during the sessions. I will discuss video recordings in more detail in the next section.

### 9.3.5 Lessons Learnt from the Post-Questionnaires

A problem with the closed questions in the post-questionnaire for Test 2 was that it seemed like the answers were often positive in spite of observations that would suggest that there had been problems with certain parts of the user interface. As discussed previously, participants may often provide the answers they think they should provide (Nielsen, 1993). Another problem with the closed questions was also the way they were phrased. Several of these problems were discussed in Chapter 8 and will not be repeated here. If I were to do the test again the questions would have to be reformulated more in relation to the answers that I wanted from each of the questions.

I found the answers to the open questions more valuable as feedback because they were more specific about problems with parts of the user interface. The answers to the open questions were, however, difficult to interpret because the participants did not answer all the questions and some answered in incomplete sentences. This problem is also noted by Nielsen (1993). In a future test the participants could be asked to explain the answers that were difficult to interpret or the session could rather include a post-interview where follow-up questions could be asked to provide further meaning to the response from the participants.

To avoid the problem of open questions Nielsen (1993) suggests that most questions should be closed and there should be a number of alternative check boxes (such as a number of opinions on a rating scale). Such questions were also included as part of the post-questionnaire for Test 2. The problem was that the test group was small so that it was not possible to generalize to a large extent based on answers. The open questions actually turned out to provide the most useful feedback regarding the design of the user interface even though the answers were not statistically credible and difficult to interpret.

There were also some more particular problems in the post-questionnaire. As discussed in Chapter 8 there was a problem with use of the word “oscillates” in one of the questions. If the questionnaires had been tested before running the actual test the question could have been rephrased and the problem would have been avoided.

As briefly discussed in Chapter 8, most of the participants reported to be satisfied with the opportunity to collaborate during the sessions. A problem with the questions asked in Test 2, however, was that it was not formulated in such a way that it was possible to evaluate what the participants thought about the support for collaboration in the user interface.

As discussed in more depth in Chapter 8 there was also a problem with the last part of the questionnaire which included questions about the structure of the underlying model. This has been discussed previously and will not be dealt with again here.

The weaknesses in the design of the test suggest changes to further evaluation. I would consider drawing on interaction analysis to test a future version of the prototype.

## 9.4 Ways Forward: Future Testing of the Two-Shower Prototype

Chapters 6-8 considered Test 1 and Test 2 of the Two-Shower prototype. The focus of both these tests was on the user interface design. The tests gave indications of parts of the user interface that the participants had problems with and the results were used as a basis for suggestions regarding future development of the prototype. This section briefly considers future testing of the Two-Shower prototype based on the experiences from Test 1 and Test 2 and the assumption that there is a need to study how participants use the Two-Shower prototype.

For a future usability test of the Two-Shower prototype I would use a larger group of participants. They should be system dynamics students as they would then be part of the actual intended user group. The test could be performed at three stages in a semester by some students at the beginning, some in the middle, and some at the end of the semester. The test would be performed in a computer lab as this would be closer to the location in which the students would normally work. To solve the problem of incomplete sentences in the questionnaires I would use open interviews or questionnaires with closed questions. Closed questions, however, would require a larger group for the answers to be representative. If questionnaires were used they would have to be tested in advance.

For a future test I would also to a larger extent utilize the video recordings, drawing on some of the ways in which video is utilized in, for example, *interaction analysis*. In interaction analysis video recordings are typically utilized to study interaction between people and how they use technology (Jordan & Henderson, 1995). Groups of researchers analyze the video recordings together sometimes including the participants in the study to clarify what has happened (ibid.). The video material is played over and over again and a finer and finer analysis is made.

In Test 1 of the Two-Shower prototype there was a problem with the students' inability to express themselves about the system after they had run the simulation (Frotjold, 2005). Frotjold had to ask leading questions in order to make them describe their understanding of the system. The students were not able to communicate in words the problems they had experienced in retrospect. A more thorough analysis of video recordings could have been used to study how they interacted with each other and the system. For an evaluation of the user interface, however, it would not be necessary or efficient to analyze the video recordings in as much depth as is the norm within interaction analysis.

Interaction analysis also suggests considering in greater detail the interactions between participants and between participants and artefacts: Interaction analysis is an ethnographic method where the interactions among people, artifacts, and their environment in a natural setting, are studied. Interaction analysis is *"…an interdisciplinary method for the empirical investigation of the interaction of human* *beings with each other and with objects in their environment"* (Jordan & Henderson, 1995, p. 39). Human activities are at the center of the research and participants are studied as they interact with each other and handle artifacts. Speech and nonverbal interaction are studied to identify practices and problems that they have in common and how they draw on available resources to solve their problems.

Particularly if the Two-Shower prototype, when fully developed, were to be used as part of an educational setting a more extensive study based on interaction analysis should be performed. The focus would then be on how participants interact with the Two-Shower prototype in order to see how the technology supports or hinder their participation and interaction. Interaction analysis in other words may be a suitable method for studying how a simulation-based interactive learning environment is actually used and how people deal with the underlying model through the user interface.

In sum, a future study of the Two-Shower prototype should study the following more fully:

- What types of interaction and activities does the prototype engender?
- How do the participants communicate based on the prototype?
- How is their understanding and learning reflected in their communication about the prototype?

In this chapter future work related to the Two-Shower prototype has been discussed. The next chapter is a continuation of how the experiences from development of the Two-Shower prototype, and the findings from Test 1 and Test 2, may be used in the development of a prototype of another simulation-based interactive learning environment. I discuss a prototype based on a larger system dynamics model: the Quito prototype. Examples of how the design ideas and findings may be transferred to the Quito prototype are discussed through examples and suggestions for interface design. Some additional considerations that must be made in relation to working with a larger and more complex prototype are also discussed, such as navigational problems and some additional issues relating to models that represent social systems.

# 10. Applying the Experiences from the Two-Shower Prototype to a Larger Model: The Quito Model

## 10.1 Introduction

The previous chapters of this thesis have been on the development, evaluation, and future work of the Two-Shower prototype. In this chapter I discuss how the experiences discussed in the previous chapters may be transferred to a second prototype - the Quito prototype. The Quito prototype portrays some of the problems of the development of a city. Part of the added complexity lies in the diverse characteristics of such a system.

In the first sections of this chapter the underlying simulation model of the Quito prototype, the Quito model, which was developed utilizing a system dynamics approach, is described. Section 10.2 is an introduction to the Quito model. In Section 10.3 I discuss how the characteristics of the Quito model are more diverse and thereby more complex than the Two-Shower model and point to considerations that must be made in the visualization of such a model. The current user interface of the Quito prototype, developed primarily by Hanne-Lovise Skartveit and Nils Magnus Djupvik, is presented in Section 10.4. Some additional concerns regarding the placement of system dynamics in relation to social theory are raised in Section 10.5. This discussion is important as the Quito model portrays actors in a social system. Implications for the development of an interactive learning environment are also considered. Section 10.6 is a discussion of how the Quito prototype can be made into an interactive learning environment for students in complex systems. This section is partly based on arguments presented in Sections 10.2-10.5 and on experiences from the development and testing of the Two-Shower prototype. The incorporation of narrativity, visualization of particular model elements, and how the model may be further visualized, is discussed.

In the following I will refer to two *versions* of the Quito prototype. The first is the *current Quito prototype*, which is the model with the existing user interface developed as part of the VOCS project predominantly by Skartveit and Djupvik. The second is the *future version of the Quito prototype,* which is the prototype of the interactive learning environment that is proposed as future work in this chapter.

## 10.2 The Characteristics of the Quito Model

This section is a description of the Quito model on which the Quito prototype is based. In the following subsections, the main components of the Quito model and its main structure are discussed. A presentation of some of the delays in the system is then provided and an example of nonlinearity in the system is presented in the end of the section. I do not go into the particular details of the model structure, the behavior, or details about all considerations that were made in choosing the variables and units of measure. This was more thoroughly explained in the project reports by Guevara-Chaves and Perez-Bennet (2002) and Djupvik and Frotjold (2003).

Unlike the underlying model of the Two-Shower prototype, which was based on a model presented in an article by Morecroft et al. (1995), the underlying model of the Quito prototype was developed within the VOCS project. The initial version of the Quito model was developed by Antonio Perez-Bennet and Porfirio Guevara-Chaves in 2002 in a term project as part of the Master's program in system dynamics at the University of Bergen. Nils Magnus Djupvik and Laila Frotjold made an improved version of the model as a term project for the same course in 2003. Both projects were partly supervised by Hanne-Lovise Skartveit and myself. Skartveit and I also developed the requirements that the model should fulfill. The description and analysis of these models in this chapter are mine and inform the discussion of design of a future Quito prototype in Section 10.6. The interface created by Skartveit and Djupvik, described in Section 10.4, is informed by the discussions and analysis of the underlying model. The choice of audience for this prototype, citizens and planners in Quito, was made by Skartveit (note, however, that this prototype was not completed and translated into Spanish). As this audience is different from the audience I have primarily been concerned with (students of system dynamics), the design of the current interface is different from the one I propose in Section 10.6. Due to these differences in audiences and intended user-groups the proposed prototype in Section 10.6 is determined to a greater extent by the need to understand the structure of the model.

The Quito model describes problems related to the urban development of Quito, the capital of Ecuador. Quito is inscribed on the UNESCO World Heritage Cultural Sites list because of its historic buildings in the city center. According to UNESCO, the city has the best preserved, least altered historic center in Latin America.[28] Being both a historic and a modern city, it faces the challenges of mixed use: a city for locals, tourists, different social groups, commerce, and housing.

The dynamics of cities is a prevailing issue within the system dynamic community, as well as for city planners, government officials, and citizens of any city. Urban dynamics was first introduced to the field of system dynamics by Jay Forrester (1969). He studied the patterns of rapid population growth and following economic decline, as has been observed in cities such as Manhattan, Detroit, St.Luis, Chicago, and Boston. Forrester constructed a system dynamic model that represented a city as a system of interacting industries, housing, and people. Urban development and related issues has since been a central problem analyzed by several practitioners in the field of system dynamics (see for example Backus, Schwein, Johnson, & Walker, 2001; Mayo, Callaghan, & Dalton, 2001; Piattelli, Cueno, Bianchi, & Soncin, 2002; Sudhir, Srinivasan, & Muraleedharan, 1997).

---

[28] From the UNESCO web-site: http://whc.unesco.org/sites/2.htm (accessed March 12, 2003).

The urban problems of Quito were chosen as a basis for our second prototype as the problem of the dynamics of a city is suitable for visualization of complex systems and a number of studies were already performed within system dynamics. The problem is also suitable for a narrative approach as the city contains actors and characteristics such as buildings and traffic that may be suitable for visualization.

Although the Quito model is relatively large in terms of number of variables, it is a relatively simple representation of the problems related to the development of a city and is not intended for, for example, city planning. Rather, we sought to portray some main characteristics of these urban problems for education about complex systems.

The Quito model describes a variety of socially related problems regarding the maintenance of the historic buildings and issues of concern for the citizens and tourists. The model includes pollution, traffic, unorganized street vendors, and *attractiveness* for tourists and citizens. The objective for the participant using the prototype of the learning environment is to manage the cultural heritage, that is, to preserve the buildings, but at the same time look after the interests of the citizens and the tourists. In order to gain control of the model, the participant would have to gain some understanding of how the underlying structure causes behavior.

Similar to the Two-Shower prototype, the future version of the Quito prototype is not intended as a standalone product, but should be used as part of an educational setting. The target audience, similar to the Two-Shower prototype, would be people who are going to learn about system dynamics and the age group would be undergraduate and graduate students. I discuss this target group and their needs further in Sections 10.5 and 10.6.

### 10.2.1 Some Main Differences Between the Two-Shower Model and the Quito Model

There are some differences between the Two-Shower model and the Quito model that bring forward some additional problems to be considered in the design of the user interface of the Quito prototype.

The Quito model is not only a larger model with respect to the number of variables that are incorporated into the model. It is also more complex and has more diverse characteristics than the Two-Shower model. The implication of the added diversity is that there are some additional requirements to the visualizations in portraying and communicating this diversity. The problem of model diversity is considered in Section 10.3.

The Quito model is also more complex with respect to the assumptions that are made in the model. A condition that may be seen as advantageous by some participants may not be seen as advantageous by others. In the Two-Shower prototype it is relatively evident that both participants try to obtain a comfortable temperature and that the comfortable temperature is somewhere between hot and cold. In the Quito model the effect that the number of street vendors have on tourists, for example, may not be as evident. There was a discussion about this in the VOCS project group where some felt that the street vendors would add an exotic atmosphere and character to the city, while others felt that the street vendors would be bothersome and hinder the tourists from seeing the historic buildings. This brings new challenges with respect to visualization of the assumptions that are made in the model. It is necessary to inform the participants of these assumptions through the visualizations. This will be discussed in Sections 10.5 and 10.6.

There is also a need to consider the effect of using such a model on people’s understanding of what is being portrayed. It may influence the participants to believe that changes made in the model will have an equal effect when made in the real system (video could for example enhance the sense of the model being a *copy* of the real system).

The Quito model is a model of a social system. The problems of a social system brings forward some additional concerns that must be taken into account in the design of the user interface. Social theory and implications for design will be discussed in Section 10.5.

### 10.2.2 The Main Components of the Quito Model

The components of the Quito model were chosen based on how they were assumed to affect the problem of maintaining historical buildings, living conditions and *satisfaction* for the citizens and the *attractiveness* of the city for the tourists.

The Quito model has ten major interconnected components:

1. Population
2. Formal employment
3. Informal employment (street vendors)
4. Tourism
5. Human pollution
6. Pollution from industry
7. Garbage
8. The historic center
9. Crime
10. Government

The *population* part of the model describes how the population of Quito is controlled by various factors such as births, deaths, immigration, and emigration. Immigration and emigration is again controlled by factors that make the city more or less attractive to live in, such as employment opportunities, the crime rate, and pollution. Data about this was gathered from the Latin American Center of Demography (CELADE)[29], which showed a problem that many Latin American cities face of continuous migration from rural to urban areas. From 1970 to 1995 the urban population of Ecuador grew from 40% to 60% of the total population.

The section of the model that describes the *historic center* deals with the quality of the buildings in the historic center and how that quality may be improved through investments in restoration. According to Jones and Bromley (1996) who refer to Illustre de Municipio de Quito (IMC, 1991) efforts to restore and maintain the

[29] CELADE: http://www.eclac.cl/celade/default.asp?idioma=IN (accessed August 8, 2006).

buildings began in the 1920s, but by the mid 1980's 14% of the buildings were still classified as *under threat* (Jones & Bromley, 1996, p. 376).

The *formal employment* sector of the model deals with how many jobs are available and how the industry is influenced by the skilled labor that is available and the capacity of the industry. Data about this was gathered from INEC (Instituto Nacional de Estadistica y Censos, the National Statistics Institute of Ecuador, 2001), which is an Ecuadorian governmental institution gathering population data..

*Informal employment* refers to unorganized street vendors who sell goods in the historic center of Quito. The level of street vendors is influenced by the size of the population and the number of tourists. Data for this part of the model was based on Mangurian (1999) and INEC (2001).

The *tourism* sector describes how the number of tourists who come to Quito varies according to the quality of the historic buildings, the crime rate, tax on tourism, number of travel agencies, street vendor density, tourism capacity, and garbage in the streets.

The *pollution* sector of the model describes how the pollution level of the city changes based on investment in public transport, pollution initiatives, the size of the population, and industry. Only air pollution is taken into account in the model, not for example, ground water pollution. This sector was partly based on data from Southgate and Lach (1995).

*Garbage* in the streets is influenced by the number of street vendors and the number of tourists who visit the historic city center.

*Crime* is described in the model as being influenced by the percentage of employment and investments in law enforcement. The data is based in part on Glaeser and Sacerdote (1999).

The model also has a section that describes the city *government* of Quito. This section has a government income which is dependent on industry, population, and number of tourists. The income is used as a basis for development of a budget, which influences the city investments in law enforcement, restoration of the historic buildings, and public transportation.

As discussed previously, the Quito model represents a simplified version of the problems related to the development of the city and is not intended to be used for creating plans about the development of the city. The income and expenses of a city is, for example, much more complex than what is described here. A rise in population will, for example, not only represent more income for a city, but also represent an increase in expenses, as more people require medical care and education.

### 10.2.3 The Reinforcing Loops of the Quito Model

In this section, I describe the major *reinforcing* loops of the Quito model. The loops may be seen in the causal loop diagram in figure 10-1. The Quito model has nine main reinforcing loops.

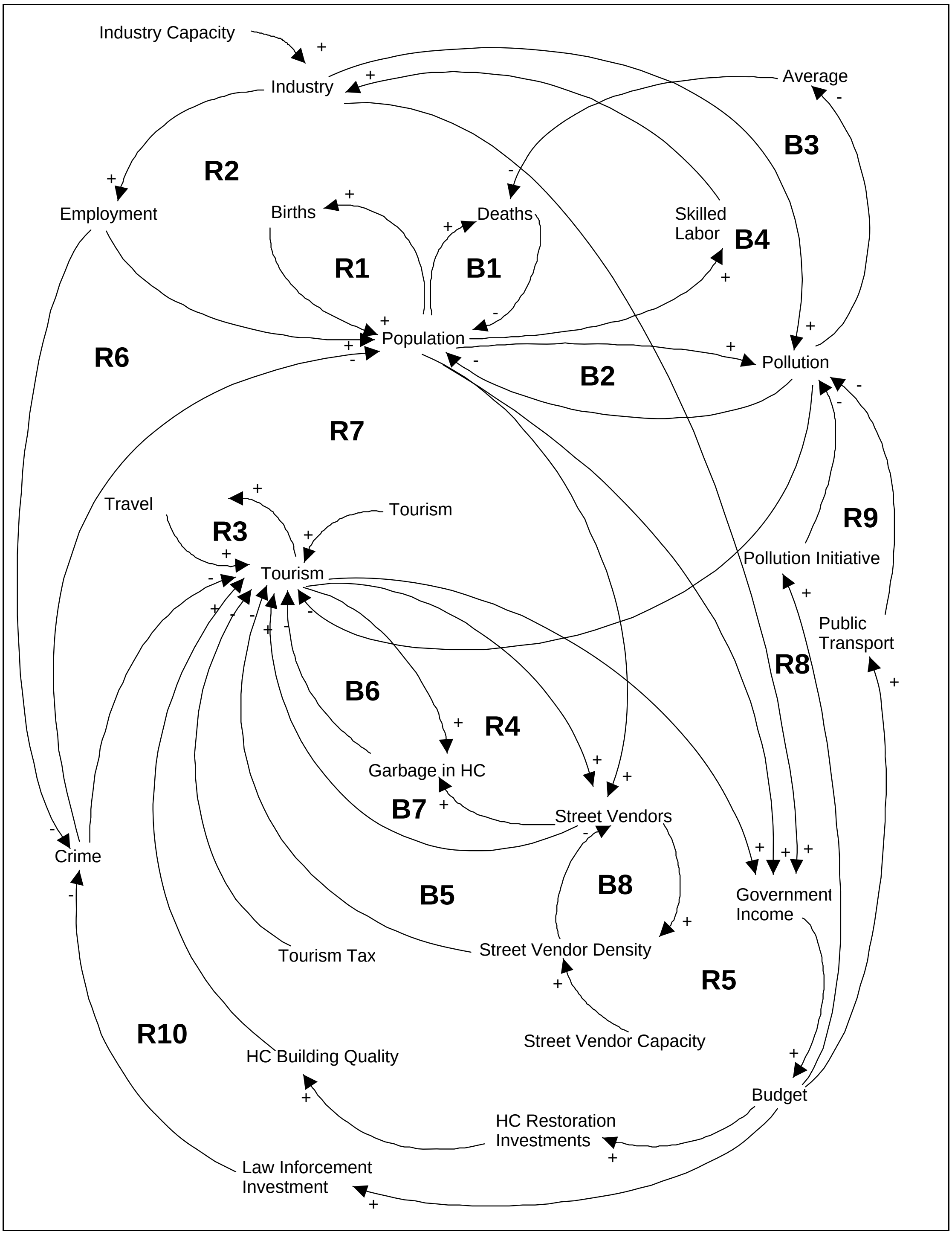


Figure 10-1: Causal loop diagram of the Quito model.

*R1: The effect of births on the population*. The number of births is dependent on the size of the population. The larger the population, the higher the number of births will be. The higher the number of births, the higher the population will be. The effect of this loop will have a delay of 20-30 years. Here it is important to note that the birth rate will always have a positive effect on the population, that is, it will always represent an increase in the population. If the birthrate decreases this does not cause the population to decrease. This is one of the problems of causal loop diagrams as discussed in Section 2.6.

*R2: The effect of available jobs on population and industry*. The population is also controlled through emigration and immigration based on the number of available jobs. If the employment rate is high, Quito may seem more attractive and more people may move to the city. An increase in the population may lead to an increase in skilled workers that are available to the industry. This may make investments in industry more attractive and thereby lead to a higher employment rate.

*R3: The effect of travel agencies on tourists*. The more travel agencies that promote Quito as one of their destinations, the more tourists will visit the city. The more tourists who travel to Quito, the more attractive it is for travel agencies to promote trips to the city.

*R4: The relationship between street vendors and tourists*. An increase in tourists may lead to an increase in the number of street sellers, as there will be more people to buy their goods. In the model we have made the assumption that the street vendors also have a positive effect on the number of tourists. This may be debated as some tourists may prefer to walk around in clean, calm, museum-like surroundings, while others may enjoy the authenticity and excitement of the crowded streets.

*R5: The effect of tourists on the government income and building restoration*. If the number of tourists increases, the city government income will increase. This means more resources may be spent on restoration of the historic buildings, which affects the quality of the historic buildings and again the attractiveness of the city for tourists.

*R6: The effect of employment on crime*. If the industry of the city increases, the employment rate will also increase. In the model, we have assumed that this leads to less crime and that a lowering in crime again makes the city more attractive for immigrants. A higher population will again mean more skilled labor which, as discussed earlier, leads to more industry.

*R7: The effect of the crime rate on the population*. If the number of crimes decreases the attractiveness of living there also increases and more people may want to move there. This increases the income of the government through taxes. The increased income may be used to improve law enforcement which again may lead to less crime in the city.

*R8: The effect of pollution on the population*. If the pollution in the city decreases, it may be more attractive for people to live there and more people may move to Quito. A higher population will lead to a higher income for the government, which again can be spent on pollution initiatives to decrease the pollution even further.

*R9: The effect of the population on public transport*. If the population increases, the government income will increase, and more money can be used for public transportation. This may lead to less pollution in the city and may make the city more attractive for people to move there.

*R10: The effect of tourism on government income and crime.* An increase in tourism will lead to an increase in government income. This further leads to an increase in the budget, which may be used for investments in law enforcement. In the model it is assumed that an increase in law enforcement will lead to a decrease in crime (although this may not always be the case in reality). A decrease in crime leads to an increase in tourism.

### 10.2.4 The Balancing Loops of the Quito Model

The last subsection was a description of the *reinforcing* loops of the Quito model. This section is a description of the eight *balancing* loops in the model.

*B1: Deaths as a regulation of the population.* The size of the population is regulated by the death rate of the population.

*B2: The effect of the population on the pollution level*. If the population increases, there will be more pollution in the city such as, for example, from heating or cars. As the pollution increases living in the city will become less attractive and people may migrate from the city or the growth of the population may stagnate. This will again lead to a lower pollution level or a stagnation of the growth of pollution.

*B3: The effect of pollution on average lifetime*. If the pollution increases, then the average lifetime of the population will decrease and the death rate will increase. The death rate drains the population and a decrease in the population again leads to a decrease in the pollution level.

*B4: Pollution from industry*. If the population increases, then the skilled labor force increases. This causes the industry to increase, which again increases the pollution. An increase in pollution causes a decrease in the population.

*B5: The effect of street vendor density on tourists*. If the number of tourists increases, it will become more attractive for unorganized street vendors to sell their goods in the city center, and the number of street vendors will thus increase. This will cause the street vendor density of the city center to increase. Too many street vendors may have a negative effect on the tourists and thus either decrease the number of tourists in the city center or dampen the growth of tourists.

*B6: The effect of tourism on garbage in the streets*. If the number of tourists increases, then there will be more people throwing garbage in the streets. This will make the city center less attractive for tourists and may make the level of tourists to decrease again or stagnate the growth in the level of tourists.

*B7: The effect of street vendors on garbage in the streets*. There is also a balancing loop describing the effect that, for example, an increase in the number of street vendors has on garbage in the streets and tourists. If the number of street vendors increases, more garbage will be thrown into the streets. This will make the city center less attractive for tourists. As the level of tourists decrease, the number of street vendors will also decrease.

*B8: Regulation of street vendors*. If the level of street vendors increases, there will be a higher density of street vendors in the city center. This will make the city center less attractive for the street vendors as it will be more crowded and more difficult for them to sell their goods. After some time, this, consequently, may lead to less street vendors, which again makes the city center more attractive for the street vendors.

### 10.2.5 Examples of Delays in the Quito model

There are several delays in the Quito model. The delays make it difficult to see the consequences of actions because it takes time before the consequences become apparent. The user interface of the future version of the Quito prototype should represent these delays and communicate their effects.

All stocks in the system represent delays, as it takes time for the levels to increase or decrease. The *perceived attractiveness* of the city for (potential and existing) citizens is represented as a stock in the system that is accumulating or deteriorating over time. This is because it takes time for the city to build up or erode its actual attractiveness and, subsequently, its reputation, based on, say as in this case, the number of crimes, the employment rate, and the level of pollution. Moreover, it takes some time before people start moving to the city, when its reputation improves, in the same way as it takes time before people start moving from the city when the reputation deteriorates.

In a similar way, the population gradually accumulates or decreases. This is based on the birth and death rates of the population and on the immigration and emigration rates (typically resulting from the perceived attractiveness described above).

There are also delays in the formal employment sector of the model as it may take years from new industry is planned until the construction work has been completed and jobs become available to the citizens. Industry also declines over time as some close down production or move to other locations each year.

The informal employment sector shows that also the number of street vendors accumulates or declines gradually. The decision to become a street vendor is not made in an instant and is in this case dependent on the number of tourists, the number of street vendors that already exist, and the lack of formal employment opportunities. In the same manner, the number of street vendors also declines gradually. This is dependent on tourism and growth in construction-related jobs. The number of street vendors may also decline over time due to law enforcement and city government policies to limit the number of unorganized workers (the educational level was not part of the model). The tourism sector contains delays in that it may take several years to make tourists choose Quito as their travel destination, and it is likely that it will take years to increase of decrease the number of tourists that visit the city each year. This is determined by the attractiveness of the city for tourists. This attractiveness changes gradually based on other factors that develop over time, such as the density of street vendors, the number of crimes, the reported quality of the historic buildings, the density of tourists, the amount of garbage in the streets, and pollution in the air. In addition, the city's attractiveness for the tourists and travel agencies is not an immediate consequence of these factors and the number of tourists that visit the city will therefore change only gradually over time.

Pollution that is produced by humans is also accumulated gradually. This is determined by public transportation, private cars, and the number of tourists. Pollution from industry is also accumulated. This is influenced by the industry which, as mentioned previously, is also accumulated over time.

It also takes time for garbage to accumulate in the streets. Levels of garbage are determined by the garbage that is already in the streets, the inflow of garbage, which is determined by the number of street vendors, and tourists and the outflow, which is determined by the capacity to remove the garbage.

The income to the city government is saved in a fund, which is used as a basis for creating a budget. The fund is gradually accumulated by taxes in amounts determined by population, tourism, and industrial units. The quality of the historic buildings is also gradually improved and/or deteriorating. Improvements are based on the funds spent on building restoration per year. Deterioration occurs over time based on the air pollution, the number of street vendors, and other deterioration factors.

The number of crimes reported in the city also changes. This is based on the unemployment rate and population density, which also change over time. Investments in law enforcement also change in a lagged response to the number of reported crimes and, subsequently, affect the crime rate.

The problem of both, understanding and representing time delays is that a change has its origin in several changes that occurred over different periods of time in the past. A change can be traced back to different sources depending on the time period in the past one chooses to study. It is determined by how much time it takes before the effect of a change in the past is seen in the future. In the next subsection I discuss some of the nonlinearity which makes the system even more difficult to understand and control.

### 10.2.6 Example of Nonlinearity in the Quito model

The Quito model contains a number of nonlinear equations. This section offers a description of nonlinearity in the model. The description is intended as an illustration of why the dynamics of the Quito model may be difficult to understand and control and serves as an example of a characteristic that must be visualized in the user interface of the enhanced prototype.

Three different possible government policies are tried out in the system in order to illustrate how nonlinearity becomes apparent in the model behavior. The example considers the effect of the number of street sellers and the level of building restoration on the quality of the historic buildings:

1. The number of street vendors is limited to 2000 and the building restoration level is set to 90%.
2. The number of street vendors allowed is 4000 while the building restoration level is reduced to 70%.
3. The number of street vendors is limited to 2000 while the building restoration level is 70%.

The resulting dynamics are shown in the graphs in figures 10-2 and 10-3 where the numbers reflect the policy tested.

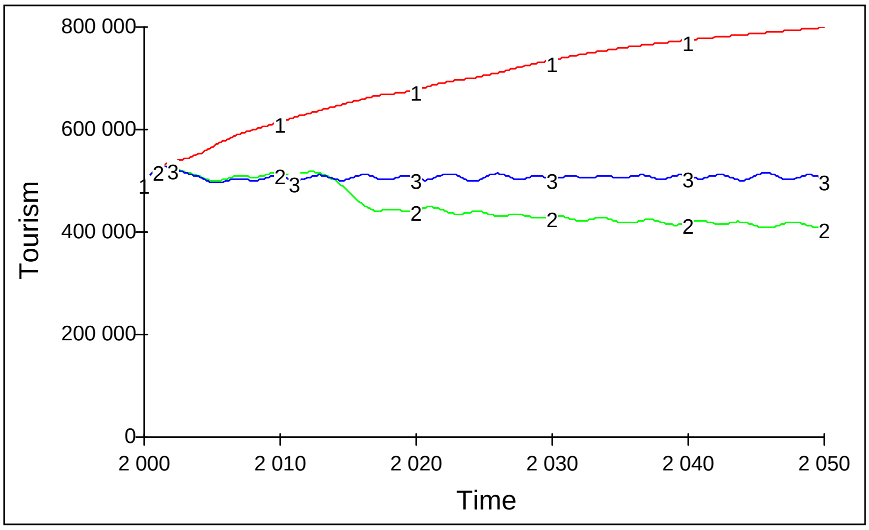


*Figure 10-2: Behavior graph of tourism in the Quito model*

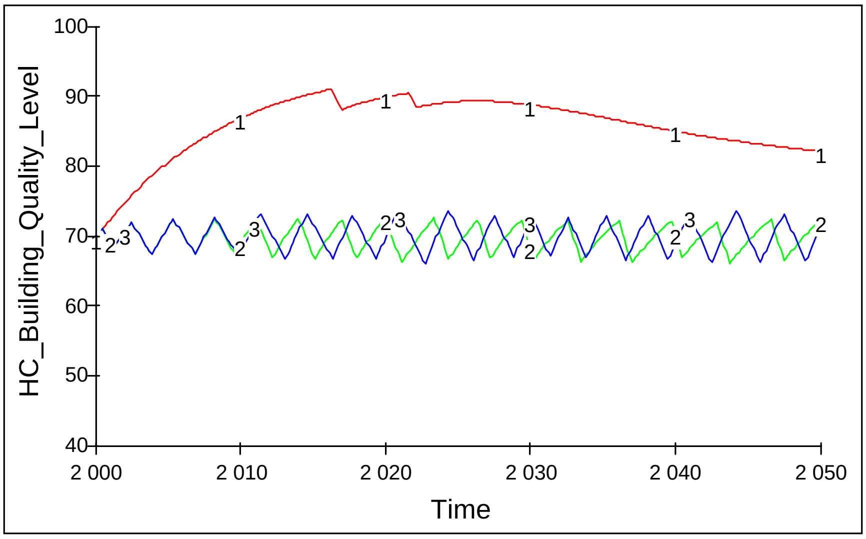


*Figure 10-3: Behavior graph of building quality in the Quito model.*

What these policies show is that a change of policy that may have a large effect on the system under certain conditions, may not have a correponding effect under other conditions. For example, if the law enforcement section of the city decides to decrease the number of street vendors allowed in the streets from 4000 to 2000 (for the purpose of decreasing the deterioration of the historic buildings and increasing the number of tourists who visit the city), the graphs show that there would be little or no effect unless more money was simultaneously spent on improving the quality of the buildings.

This section has been a description of some of the components and characteristics of the Quito model. These components and characteristics must be taken into account when developing the user interface of the prototype. The following section is a further description of the characteristics of the model, arguing that there is diversity of system elements and characteristics and that this diversity must be taken into account when developing the user interface.

## 10.3 Diversity of System Elements and Characteristics

As discussed in the previous sections, the underlying simulation model of the Quito prototype is larger and much more complex than the underlying simulation model of the Two-Shower prototype and several additional issues must be considered when developing visualizations for an audience of system dynamics students. This section considers how a model of a social system, such as the Quito model, contains diverse characteristics, such as diversity of flows, state variables, and causal relationships. The diverse factors are part of a synthesis that make the model complex and therefore difficult to understand. The diversity must be taken into account when developing the prototype and one of the aims must be to communicate the diversity of the model through the user interface. Realizing this diversity of the model characteristics is important for developing a better understanding of the system. In this section, I discuss the diversity of flows, state variables, and causal relations in the Quito model. The intention in the prototype, then, is that the user interface supports the participants in obtaining an understanding of how the various parts altogether form the behavior of the system. In the further design of the user interface of the Quito prototype one must seek to visualize the integrated diversity of the Quito model by developing visualizations that are as clear as possible and as characterizing as possible of the phenomena one is trying to explain.

### 10.3.1 Diversity of State Variables

There is diversity of state variables in the Quito model. Each state variable represents a different dimension of the system. It may therefore be misleading to represent them in the same graphs. The population and the historical buildings, for example, cannot be compared directly because they represent two very distinct aspects of a system and are controlled by different system processes. They also have different units of measure.

It is not only important to separate fundamentally different variables that have distinct units such as the population and the historical buildings. Systems representing, for example, perceptions also represent a challenge. There may be a difference between the perceived and the actual structure within a system and the perceived and actual state of a system. These variables are usually denoted by the same units of measure. In the Quito model there is a difference between the *Perceived City Attractiveness* and the *Current Attractiveness of City* (see figure 10-4). These aspects are modeled in this way so as to capture the way matters are being perceived, i.e. the perception process, including the time it takes for the population to react to changes in the city. In this case, there are two models of the attractiveness of the city, one is the actual attractiveness, while the other is the perceived attractiveness. When, for example, people discover that the city has become more attractive to live in, characteristics of the system may be taken into account when decisions are made. In this case, the population gradually becomes aware of changes in attractiveness and can consider them when deciding whether to move to or from the city.

Although measured in the same unit, the variables actual and perceived attractiveness represent two very different aspects of the system that usually take different values. Perceived attractiveness may adjust towards the real attractiveness over time as the actual value of the variable influences the perceived value of the variable. This diversity must be visualized in the Quito prototype. In such a diverse system, there may be variables that have the same unit of measure but that represent quite different dimensions of the system and whose values have an impact on the behavior of the system.

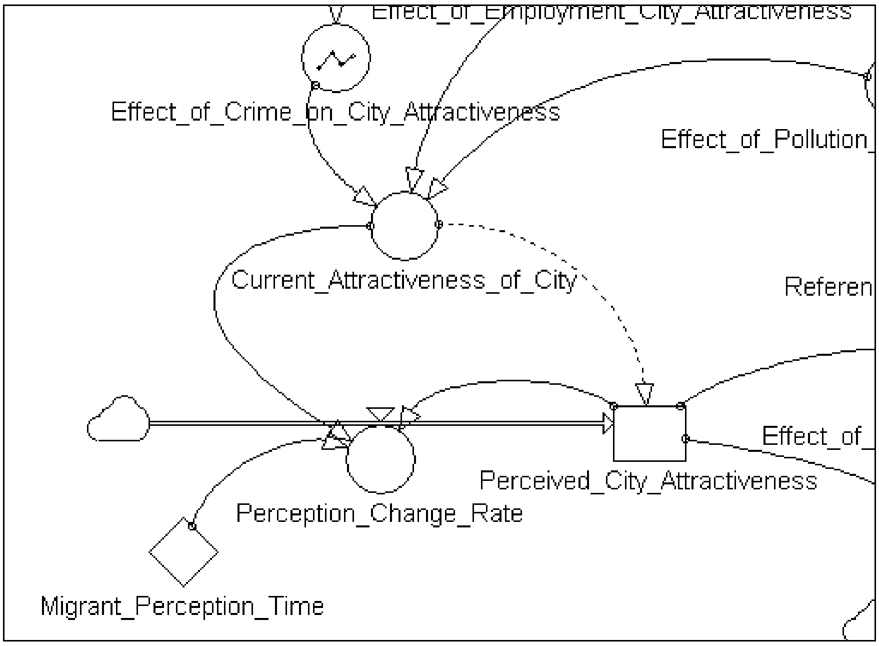


*Figure 10-4: Excerpt from the stock and flow diagram of Quito showing both the representation of the current attractiveness and perceived attractiveness of the city.*

## 10.3.2 Diversity of Flows

The integration process is the accumulation of rates into stocks that makes the stock levels and thus the state of the system change over time, - how the population grows or declines based on inflow and outflow rates or how buildings develop as a result of restoration or deterioration. In order to better comprehend the integration process, it may initially seem more effective to represent the flows as net flows where the inflows and outflows are merged into one flow (see figure 10-5). The problem of visualizing the diversity of flows was not present to the same extent for the Two-Shower prototype because the flows only consisted of water. In the Quito prototype other flows must be represented, such as population flows. In the representation of the Quito model there would then be one representation of the population of the city and one representation of the net change in the population. The intention would be to simplify the user interface through an aggregate representation and in this manner solve the problems of the large amount of information that the model contains.

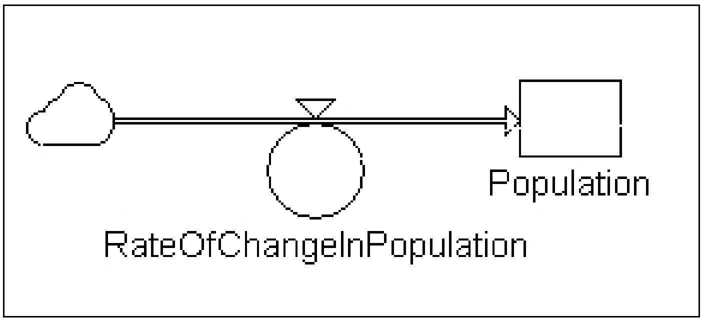


*Figure 10-5: Example of a representation of a net flow of a population in a stock and flow diagram.*

The problem is that this net flow is an abstraction of reality and does not represent the processes that occur in the real system, and thus probably does not support a better understanding of the system. The actual flows do not constitute a single type of process. There is a difference between the processes that cause an inflow and an outflow, or several inflows to or outflows from the same stock. For example, in the population of Quito there is a fundamental difference between the processes that control the birth, death, emigration, and immigration rates (see figure 10-6). Merging them into a net flow would obscure our understanding of the integration processes that occurs and thus our understanding of why the population changes over time.

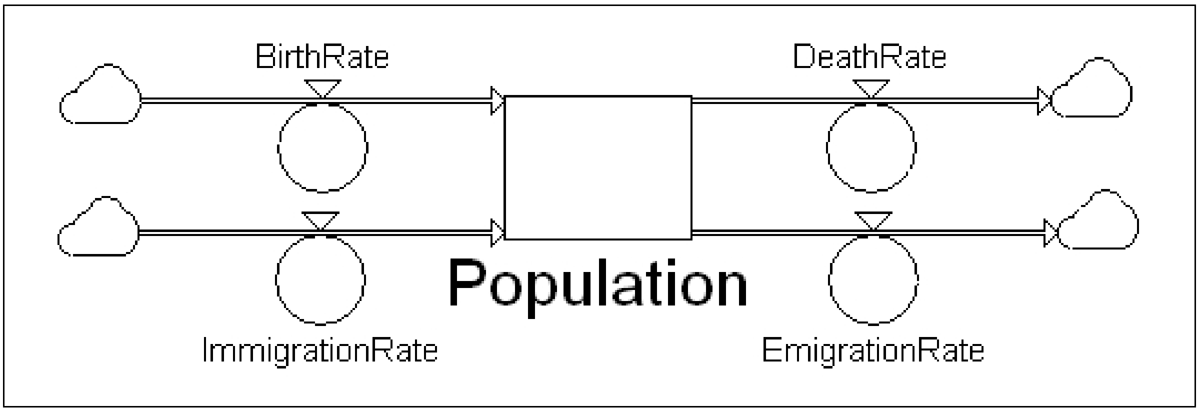


*Figure 10-6: Example of a representation of the diversity of flows of a population in a stock and flow diagram.*

### 10.3.3 Diversity of Causal Relations

There is also diversity of causal relations in the Quito model. A causal relationship is when two variables are linked and the value of one has a direct influence on the value of the other over time. A causal relationship is determined by the characteristics of the variables and physical and biological conditions in the system. In a stock and flow transition the rate causes a change in the level and the level influences the rate. For example, the death rate of the population of Quito affects the level of the population directly. Consequently, the death rate is directly affected by the level of the population.

All causal relationships involve stocks. Because the accumulation of stocks involves time, causal relationships also involve time. It takes time from a change in one variable occurs until the influenced variable reacts. In a relation that takes time the state of a variable is connected to a previous state through system feedback. This is only possible for relations that involve accumulations. In reality, all causality takes time because all causality involves accumulation. The influence that causally related variables have on each other over time must also be represented in the Quito prototype.

There is diversity of causal relationships. Each causal relationship represents a unique relation that is both qualitatively and quantitatively different from other causal relationships. Each causal relation is formed in a different manner and controlled by different system processes. For example, the death rate of the population of Quito is determined by influence from different relations that represent completely different system processes. There is a difference between the natural death rate due to old age and deaths that are influenced by pollution in the city (see figure 10-7). The death rate is here influenced by the size of the population in two different relations that represent two qualitatively different processes. Mathematically these two relationships can be merged in order to calculate the death rate, however, necessary information about the processes that occur within the system will then be lost. It is therefore important to present the diversity of such relationships in the Quito model in order to support the understanding of the diversity of processes that control the system.

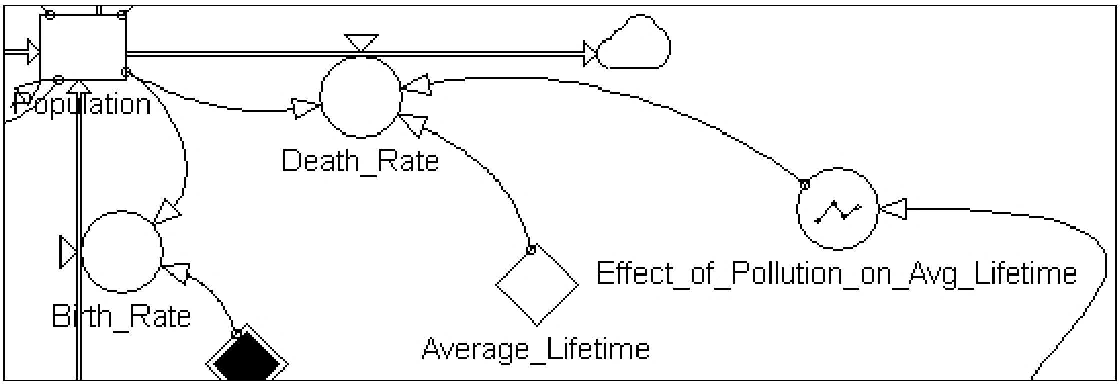


*Figure 10-7: Example of diversity of causal relations, the difference between the natural death rate due to old age and the influence of the pollution level on the death rate.*

The complexity of systems is manifested by their diversity. Diversity in this sense does not refer to the size of the system: A linear system may contain many variables, yet related through the same kind of (linear) relationships that do not allow for interaction between the variables. Rather, diversity refers to the variance in types of flows, state variables, and causal relations. In transforming the simulation model into an interactive learning environment, it therefore becomes important to seek to represent the diversity of the system and not to make simplifications that hinder an understanding of that diversity. It would be most important to explain the model at a sufficiently high level of detail so as to separate the various processes both graphically and textually.

## 10.4 The Current User Interface of the Quito Prototype

The current user interface of the Quito prototype was developed by Hanne-Lovise Skartevit and Nils Magnus Djupvik. I was involved in the initial idea and concept for the current version of the Quito prototype. The intended user groups for this version of the prototype were people with no previous experience in the formal study of complex systems. This version of the prototype was considered for use as a basis for discussion and learning about problems related to the dynamics of a city and was to be used by city planners, government officials, or citizens to discuss how problems regarding mixed use of the city may be solved.

The analyses in Sections 10.2-10.3 will be used particularly in the discussion of the planned enhanced prototype described in Section 10.6, but the model structure obviously had influence on the development of the current user interface described here. The different intended audiences affected, however, the development of the prototypes and the need for detailed understanding of the model.

The main focus of the current version of the Quito prototype is on the behavior of the model. There are behavior graphs for all the main state variables in the system. One of the main innovations in the current version of the Quito prototype is the use of video clips to illustrate model behavior. The video clips are associated with the behavior graphs and change based on the state of the model.

The first view of the current version of the Quito prototype is an introduction where the participants are informed about the problems that are faced by the citizens of and visitors to Quito and are offered some additional background information about the city (see figure 10-8). There is a text stating that the historical buildings of the city are a UNESCO World Heritage Cultural Site and that a renovation of the buildings and relocation of street vendors were made between 2000 and 2005. The next view informs the participants that they will play the role of the mayor of the city and make decisions about the city (see figure 10-9). They are informed that they must make decisions regarding building restoration, pollution prevention, garbage collection, public transportation, and law enforcement. The last introductory view presents the citizens and informs the participants that they may watch video clips of citizens who tell them how they feel about the conditions in the city at different times in the course of the simulation (see figure 10-10).

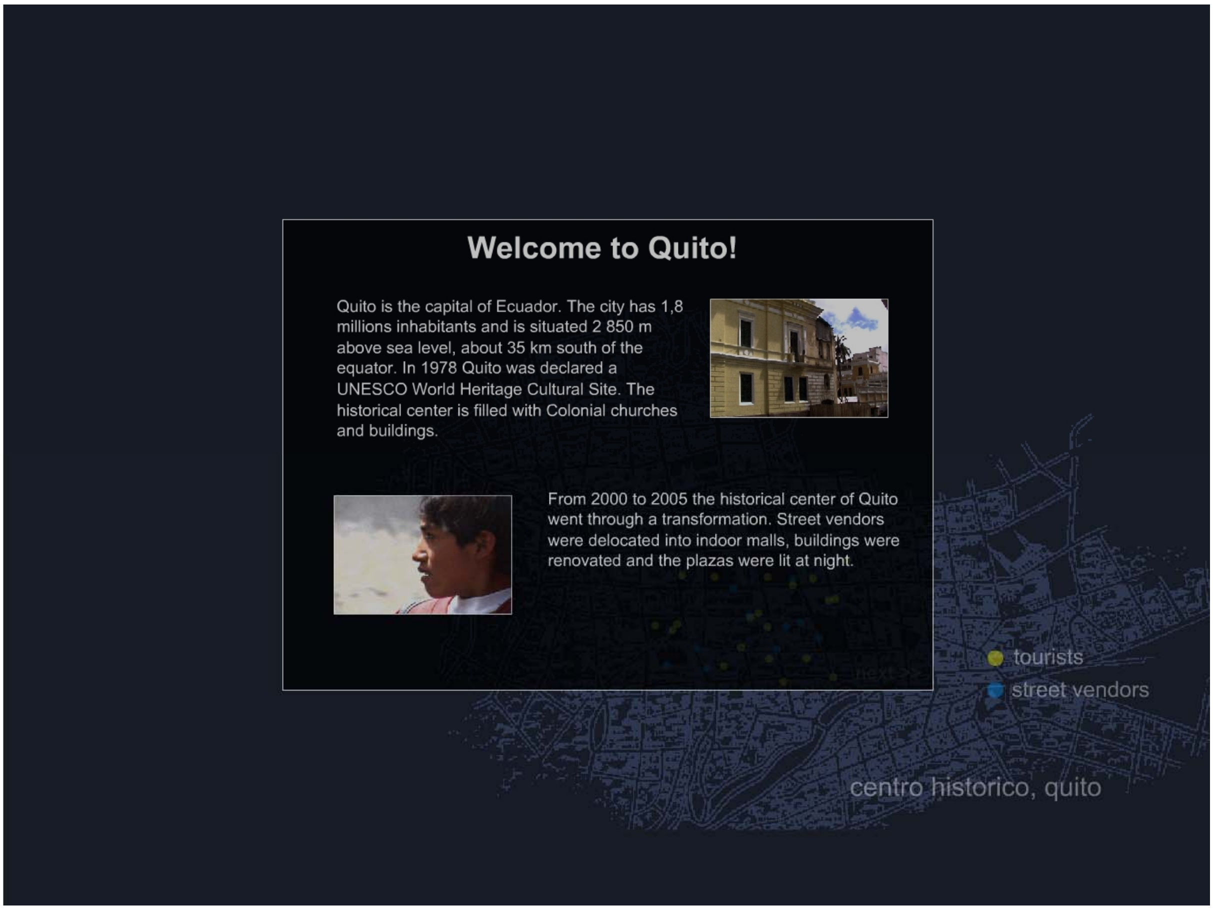


*Figure 10-8:* Welcome to Quito - *the first introductory view of current version of the Quito prototype.*

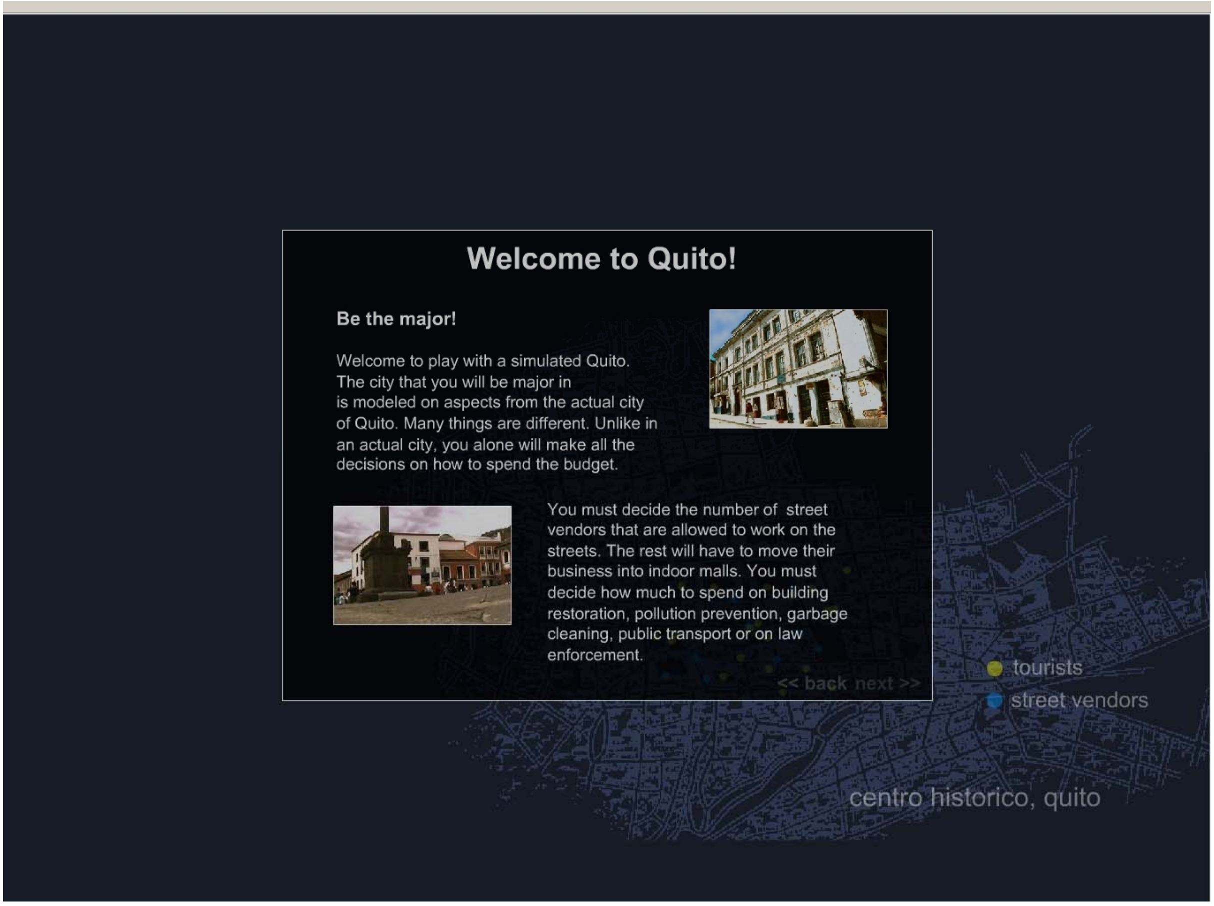


*Figure 10-9:* Be the mayor *- the second introductory view of the current version of the Quito prototype.*

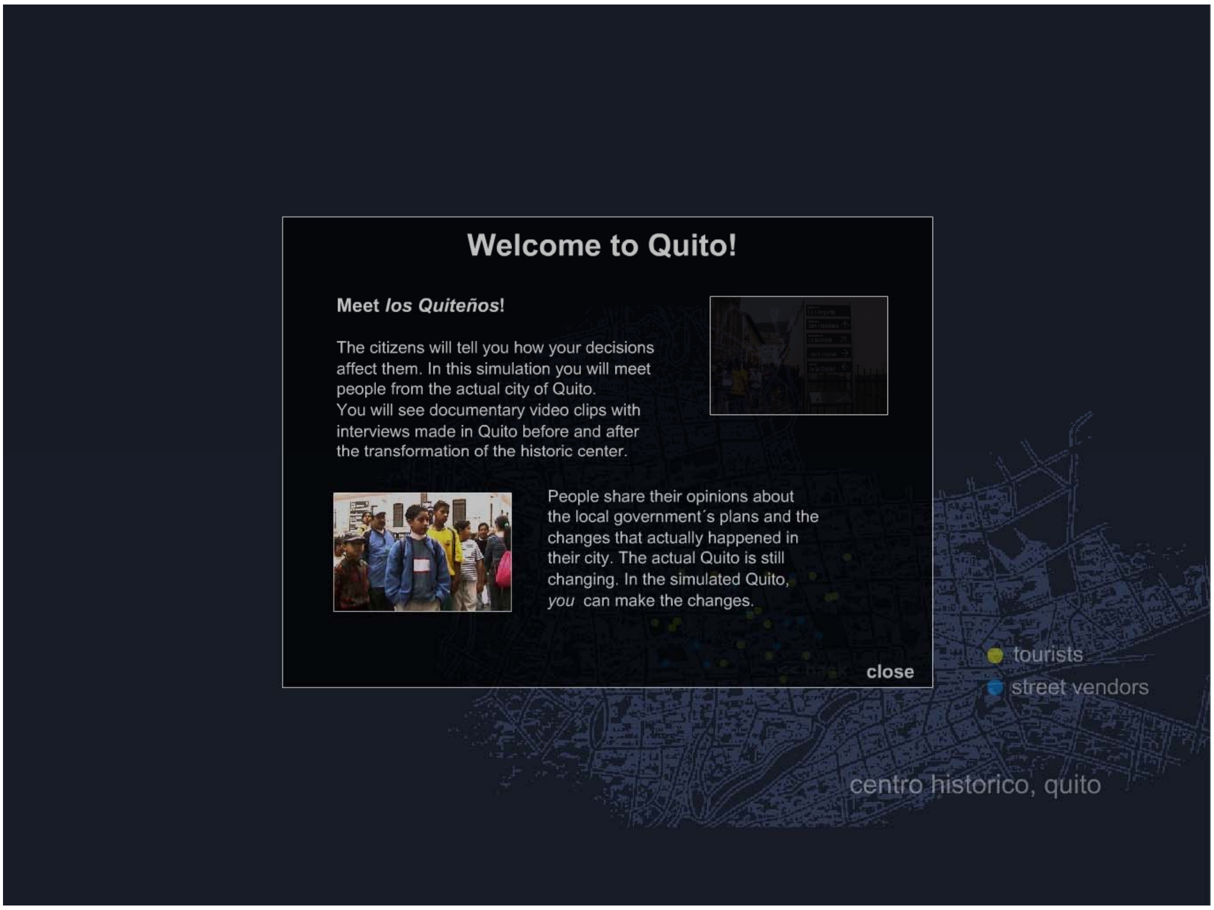


*Figure 10-10: Meet the citizens - the third introductory view of the current version of the Quito prototype.*

The current version of the Quito prototype has only one main view where the simulation model is run (see figure 10-11). The view consists of behavior graphs on the right and left hand sides of the screen. In the upper right hand corner there is a timeline. The simulation control buttons are located in the lower left hand corner, next to slider controls for the variables that the participant may control. In the lower right half of the main view there is a map of Quito with blue and yellow dots indicating the number of tourists and street vendors in the city. Attached to some of the graphs are video clips that may be pressed and watched in a larger window. They show interviews in which tourists, street sellers, car drivers, and police officers appear dependent upon the values taken by variables in the model. These video clips reflect the behavior of the system and they change based on the state of the variables that they represent. The attractiveness of the city for tourists, satisfaction of street sellers, and traffic congestion are examples of variables whose behavior are monitored and illustrated. The video clips serve as indications of the states of the variables, explanations of some causal relationships, as well as hints about the influence they may have on the behavior. The video clips are also, in part, used to represent the assumptions that are made in the underlying model. There is, for example, an interview with some tourists who like the street vendors' contribution to the city atmosphere. In this way, some of the underlying assumptions that are made in the simulation model, in this case that the street sellers have a positive effect on tourists, are represented through the video clips.

The goal and layout of the current Quito prototype resembles somewhat the game SimCity where the participants play the mayor and make various decisions regarding the city. SimCity *"…is a complex depiction of the process of urban planning, city economics, and the evolution of human community; it is a simulation game"* (Salen & Zimmerman, 2004, p. 424). SimCity is based on a much more complex model (for example, the city is divided into different geographical areas) and the goal is to make the city grow in general. The Quito prototype also has a stronger focus on a few particular problems such as those of the population, tourists, and the historic buildings. It does have, however, some of the same characteristics as simulation games (Skartveit, 2007, forthcoming).

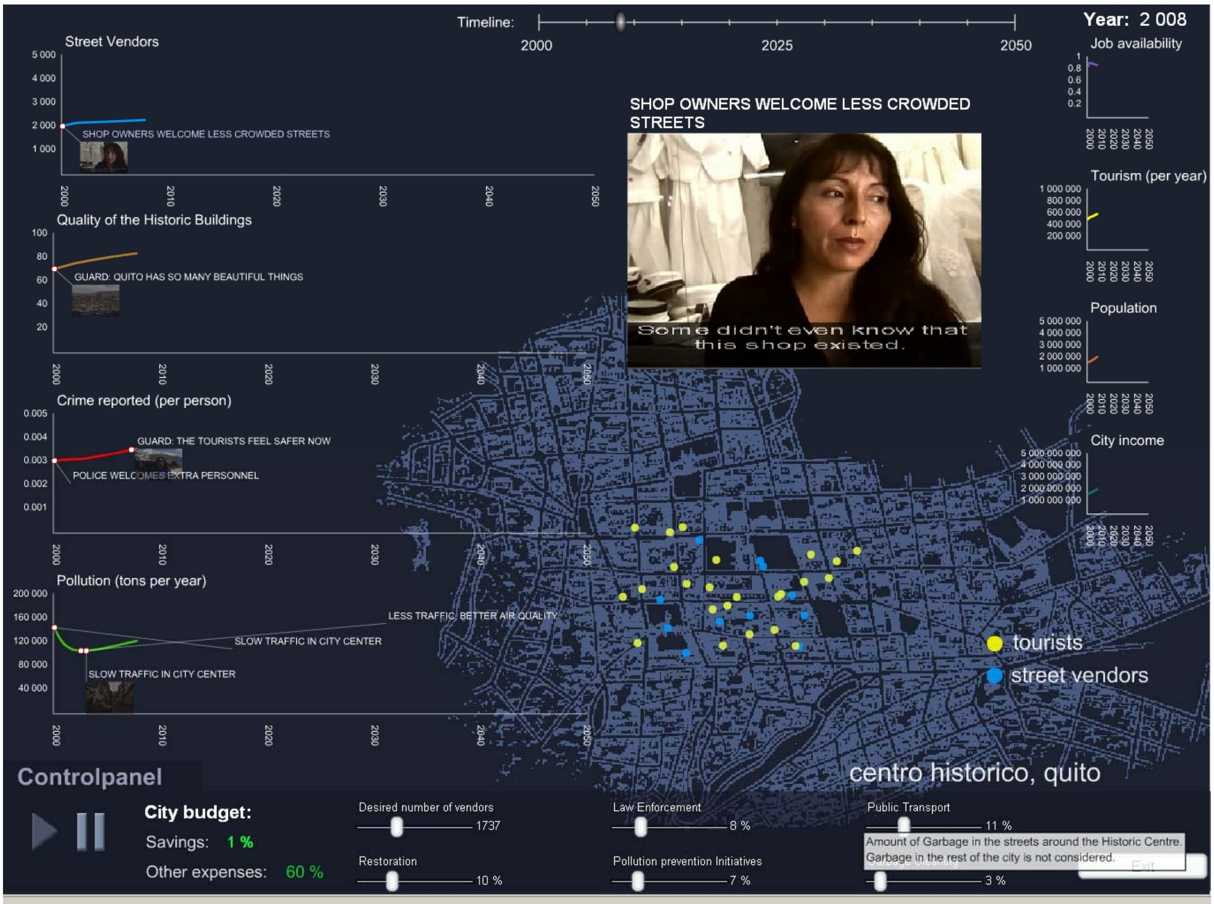


Figure 10-11: The main view of the current version of the Quito prototype.

### 10.4.1 Representation of Structure and Behavior in the Current Version of the Quito prototype

As mentioned earlier, the main focus of the current version of the Quito prototype, as it now stands, is mainly on representation of behavior and there is little focus on representation of model structure. The stocks of the model, such as street vendors, the quality of the historic buildings, crimes reported, pollution, job availability, tourism, population, and city income are represented in the behavior graphs.

Information about the state of the system is provided through the video clips. There is, for example, a video clip of tourists who complain about too much pollution in the air. Some of the headlines of the video clips also give indications about the state of the system. One video clip, for example, is called *Quito's monuments need renovation*. The participant is in this way provided with information about undesired or desired states of the system. The map in the lower right end of the screen is a map of Quito. The number of tourists and street vendors are indicated by yellow and blue dots on the map (see figure 10-11).

There is, however, little information about how the state variables are related. Some of the structural relationships may be inferred from statements made in the video clips. For example, in a video clip with a tourist guide interview, the guide refers to the increase in the police force, which has led to less crime and better conditions for the tourists. Another video clip shows an interview with a street cleaner who comments on how there is less garbage in the streets now that there are fewer street vendors.

Part of the reason the structure was not emphasized in the current user interface may be that the model was to be used as a basis for discussion. One of the intentions for the group discussions could be to address more in detail the assumptions made in the underlying structure.

### 10.4.2 Lessons Learnt from the Current User Interface

If the Quito prototype is to be used as part of a course in system dynamics, some modifications and enhancements to the current user interface must be undertaken. These modifications are not based on a test of this particular interface, but on experiences from the development process and tests of the Two-Shower prototype, as well as the considerations noted in 10.2 and 10.3. I wish to emphasize that these changes only apply if the prototype is to be used as part of a course in system dynamics. As discussed earlier, the intended use of the current version of the Quito prototype was not in the context of learning about complex systems. In this section general changes are noted. In section 10.6 I will discuss modifications called for in the creation of a more fully developed learning environment.

One of the main concerns of the current user interface is that there may be too much information in one view. In Test 1 of the Two-Shower prototype, some of the participants reported to be overwhelmed by a large amount of information (Frotjold, 2005). The current user interface of the Quito prototype only contains one view and there is little gradual introduction to the complexity of the model other than the textual explanation in the beginning before the simulation model is run and of course, the video clips. In the Two-Shower prototype gradual introduction of system elements was provided. This gradual introduction was one of the characteristics of the Two-Shower prototype that the participants of Test 2 seemed to be satisfied with. If the underlying model of the Quito prototype is to be revealed at a more detailed level, it is necessary to distribute the information across several views.

The findings from Test 2 of the Two-Shower prototype showed that some participants, who had little or no experience with system dynamics, reported to have problems reading behavior graphs. Moreover, the user interface of the current Quito model contains numerous graphs, which may make it difficult to study the graphs at the level of detail that is necessary when the aim is to learn about complex systems. In a future version of the prototype, it will therefore be reconsidered how behavior graphs may be better represented.

Some additional changes could be considered, however, not based directly on experiences with the Two-Shower prototype. The timeline of the current version of the Quito prototype could be revised so that it becomes more visible and may serve its purpose better. I consequently suggest that the timeline be moved further down and located next to the simulation controls (see figure 10-13, p. 316).

Another change that may be considered if the Quito prototype is to be used as part of a course in system dynamics is the inclusion of indications of *how well* the participants are doing. This is reflected in the videos of the current version, but additional graphical or animated indications of the level of success or satisfaction level of the various stakeholders in the system could make it easier for the participants to find out how they are doing. Such indications could also provide a better representation of the structural relationships in the model. If changes were made to the model, the effects of the change in the satisfaction level for the citizens, tourists, and street vendors would be displayed on the screen.

The previous two sections have been a short introduction to the underlying model of the Quito prototype and a description of the current user interface. In the following two sections I will discuss proposed changes to the current interface to make the prototype suitable for novice system dynamics students. Before I begin the detailed discussion of the proposed prototype, I feel that it is necessary to take a step back and discuss the view of social theory from a system dynamic perspective. This is important as the Quito prototype and its underlying model, in contrast to the metaphorical Two-Shower model, describes decisions and actions of human beings. This brings forward some additional concerns regarding design and representation of the assumptions that are made when utilizing a simulation model of a social system. The following section is based on the views presented in Lane (2001a; 2001b), which are again based on Lane (2000a; 2000b) and Lane and Oliva (1998). After this brief introduction I return to concrete design issues picking up also specific considerations that need to be made when designing a prototype of a social system. Briefly this includes representing the assumptions of the underlying model at a high enough level of detail to make the underlying assumptions explicit.

## 10.5 System Dynamics, Social Theory, and its Implications for the Quito Prototype

Lane (2001a) explores how system dynamics practice deals with three issues or assumptions concerning aspects of social theory: *"...assumptions of how human beings behave, how societies hold together and how knowledge about such processes can be acquired"* (p. 98). Lane asks the question of whether system dynamics has a social theory and whether it is necessary to place system dynamics within a social theory. *"System dynamics deals with sets of differential equations - mathematical entities manipulated in a world of almost Platonic purity"* (ibid.). He argues that the mathematics of dynamic systems theory does not equal system dynamics theory. According to Lane:

> *"System dynamics must be seen as a modeling approach that relies on assumptions about how human agents use information, how one can go about collecting empirical data to construct models, how groups of people can develop confidence in such models, and how those models can be used in a social context to address some issue" (ibid.).*

He uses this to support his argument that system dynamics deals with problems similar to that of the social sciences.

Lane bases his analysis on a two-by-two matrix developed by Burrell and Morgan (1979, see figure 10-12). The matrix represents four main disputes within sociology: whether reality is *"given or a product of the mind"*, whether experience is a necessary prerequisite for understanding, whether humans have free will or their behavior is determined by the environment, and whether the scientific method or experience is best suited for understanding something about the society.[30] They further unite these issues into social theories that are based either on the assumption that the goal of a society is *regulation* and stability or social theories where the focus is on *radical change* and revolutions. In addition, they distinguish between social theories that see the society as *objective* or *subjective*. In Burrel and Morgan's matrix, one axis ranges from *subjective* to *objective* social theories, while the other distinguishes between *radical change* views and *regulatory* views of the society. Lane (2001a) tries to define system dynamics practice into different groups, places them according to the diagram and discusses whether they belong to the category of *radical humanism*, *radical structuralism*, *interpretative sociology*, or *functionalist sociology*.

[30] http://faculty.babson.edu/krollag/org_site/org_theory/Scott_articles/burrell_morgan.html (accessed January 9, 2006).

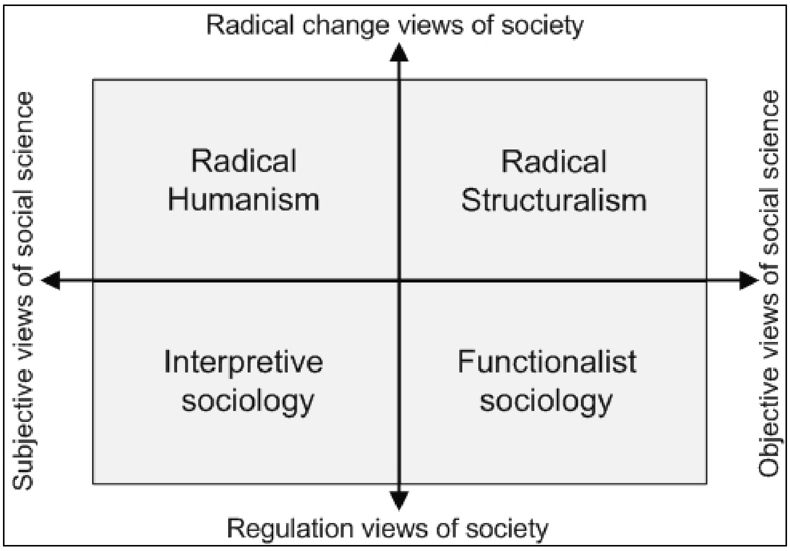


*Figure 10-12: Burrel and Morgan's (1979) matrix based on the representation in Lane (2001a).*

The problem is that system dynamics does not claim to belong to a specific social theory and their view on social theory must therefore be evaluated by looking at the system dynamics practice (ibid.). Lane divides the system dynamics practice into subgroups that operate under different assumptions about the society and knowledge about the society. With an exemption of *modeling as radical learning* which he considers a form of *radical humanism* and *holon dynamics*[31] , - a form of *interpretive sociology,* he places most of the practices within *functionalist sociology*. This view is that the structure of society is relatively objective and tends to seek for stabilization and regulation. There are several groupings of the system dynamics practice within functionalist sociology.

Lane concludes that placing system dynamics within social science theory and commenting on system dynamics in social theory terms is difficult because it has no explicit social theory. System dynamics is sometimes criticized for having a deterministic view of the social world where causal laws govern human behavior. A social theory is deterministic if it claims that the causal laws of the society exist outside subjective human decisions. The pursuit would be to find a grand theory

[31] Holon dynamics is Lane and Oliva's attempt to incorporate system dynamics and soft systems methodology (SSM) (Lane & Oliva, 1998).

constituted by the causal laws. According to Lane (ibid.) system dynamics is criticized for trying to find a *grand* theory of how the society functions and that it has an extremist view of the control that social structure has over human agency. If this were true for the Quito model, for example, the assumption would be that in theory it would be possible to develop a model of some superior, independent structure that portrayed how all citizens of the city would behave. All citizens would act according to this superior structure. Lane (2001a) argues, however, that system dynamics does not encompass a grand *theoretical* claim about society, but that phenomena may be explained by their structure. This is a grand *methodological* claim (Lane, 2001a, p. 110). For the Quito model this implies that from a system dynamic perspective it is possible to explain some of the characteristics of this society based on identifying its structure.

Lane further argues that the constructs of causal loop diagrams and stock and flow diagrams are also not part of a grand theory, but rather part of the practice of how the methodological theory may be implemented. From a system dynamics viewpoint the behavior of systems cannot be inferred by just studying a mapping of the structure. Computer simulation is necessary because the systems are too complex for humans to be able to infer the behavior based on the structure. The implications for the Quito model would be that it is necessary to study the behavior of the model in relation to its structure through computer simulation. According to Lane (2001a) this is an *empirical* claim.

In system dynamics, there is also the presumption that modeling aids the understanding of social phenomena. This presumption originates from the belief that humans have problems understanding complex systems and that computers may be used as aids to understand and infer the behavior of such systems. Lane refers to Sterman (1994) and argues that this is proved empirically through studies of the difficulties people have in determining the future behavior of complex systems and how this is aided by computer modeling. The problems of understanding complex systems were also discussed in Sections 2.2-2.3 of this thesis.

Lane (2001a) continues by asking the question of whether the causal links in system dynamics are appropriate for representing human behavior. Lane refers to Morrow and Brown (1994) and Ritzer (1996) and states that there is a debate regarding two forms of explanations: *erklären* and *verstehen*, where one represents explanation in a natural science manner and the other represents the view that explanations of individual human action requires a hermeneutical interpretation. In the social sciences the aggregation level of what is studied requires and makes possible different levels of explanation.

The intention of system dynamics is not to explain what happens at an *individual* level. *"System dynamics is concerned with aggregate social phenomena, not individual meaningful actions"* (Lane, 2001a, p. 111). Lane further concludes:

> *"...system dynamics does not involve the view that individual human decisions are explainable solely by causal laws, that subjective explanations of the Verstehen type are irrelevant. The field is simply not operating at a level of detail low enough for it to be accused of such a crude stance" (ibid., p. 112).*

In relating the model to the two categories of questions considering problems of *erklären* or *verstehen* the underlying simulation model of the Quito prototype deals with problems of the *erklären* type. It should not be interpreted as a model of individual behavior. Each citizen of Quito does not act in accordance with the representation of his or her actions in the model. The structure and behavior of the Quito model are aggregate assumptions which are used to explain and understand the aggregate behavior of certain phenomena in the city.

However, in the user interface of the current Quito prototype we have tried also to focus on the *verstehen* part by incorporating video recordings of actual Quito citizens. They talk about how they experience the development of the city, their behavior in relation to this, the conditions they live under, and what they feel about the changes that happen in their society. Part of the purpose of the video clips is to remind the participants that the aggregated population represents a number of individuals with different opinions and different intentions for their actions.

Lane (2001b) further concludes that system dynamics should emphasize *"its shift away from objective extremes"* (p. 303) but that it should not be a purely subjectivist field (ibid.). Lane (ibid.) states that system dynamics would lose its distinct characteristics if it tried to be purely subjectivist. Building simulation models is impossible if no form of generalization is allowed.

As a conclusion to the difficulty of placing system dynamics on the two-by-two grid of Burrell and Morgan (1979), Lane (2001b) suggests that system dynamics must be discussed in relation to the ongoing agency-structure debate within the social sciences. The core of the debate about agency structure concerns how the theories of structure and agency can be integrated. Lane (2001b) refers to Berger and Luckmann (1966) and their *social construction of reality thesis* that suggests that social order is continuously produced by human action and that social conduct creates and maintains social institutions. Institutions influence appropriate conduct, are maintained, and given new life as humans act according to what is seemed appropriate. Lane (2001b) discuss whether the agency structure debate has a place in system dynamics and whether system dynamics has anything to offer in the debate.

According to Lane, system dynamics is strongly influenced by its roots in engineering and control theory. *"As a result, system dynamicists frequently do not speak a language that communicates well with social scientists"* (Lane, 2001b, p. 302). Lane further states:

> *"for example, failing to articulate a clear social theory of system dynamics – treating it merely as a 're-craftable method' in which model building is as 'friendly' and socially contingent as is necessary for acceptance – is dangerously ambivalent and rootless" (ibid.).*

Lane (ibid.) further presents some reasons that justifies using system dynamics to study social systems. First, all studies have a practical relevance and focus and the aim is to increase the understanding of a problem. All variables in a model should be meaningful. The focus of the study is disequilibrium analysis. Models are not viewed as correct representations of reality and human behavior, but as a tool that incorporates different criteria and perspectives to foster debate. Finally, the modeling process may bring forward different views from the participants.

The implications of Lane's discussion for the Quito model concerns how such a model may be viewed and interpreted. From a system dynamic viewpoint it is considered adequate to construct a model such as the Quito model to seek to explain some of the social phenomena in the city of Quito, however, all variables should be meaningful, and the model should be viewed as a tool that incorporates a perspective to foster debate.

Another implication is a concern regarding design of the user interface, related to the level of aggregation and assumptions that are made in the model. Different levels of concrete versus abstract transparency and transparency in relation to the level of aggregation versus the level of detail must be considered in the design of the user interface. The user interface should be designed in such a way that it would be possible for the participants to evaluate and validate the assumptions of the underlying model. The underlying model of the Two-Shower prototype may not be as controversial and does not incorporate issues that are subject to discussions similar to those that the Quito model may trigger. Embedded in the underlying simulation model there are aggregate assumptions about how people may behave in response to changes in policies and changes in living conditions. In the current version of the user interface of the Quito prototype these aggregate assumptions are represented in video clips where people are talking about their experiences and views of problems in the city. In one way these video clips contribute to the concrete transparency, aiming at representing differing views among the stakeholders of the city. Video as a representation format should also be familiar to the participants even if they have little or no experience with the formal study of complex systems. On the other hand, these video clips may represent aggregations that do not incorporate the views of other people within the same groups in the city (for example, all tourists may not think that a large number of street vendors are good). The current version of the Quito prototype, however, may not incorporate the level of detail that is necessary to study the details of the assumptions that are made in the underlying model. In a future version of the prototype, the assumptions should also be made explicit to the participants who may wish to study them more closely at a detailed level, such as stock and flow diagrams and equations. These would represent the details of the assumptions that were made in the underlying model and that could as suggested by Lane be used to discuss the perspective represented in the prototype. The problem may, however, be that it is difficult to understand the equations. Some textual explanations of the assumptions made in the equations could be included as part of the user interface. The user interface could include in this way means for the participants to study the assumptions of the model by making them more transparent through the user interface.

It is also important that the facilitators of a session in which the Quito prototype is being used inform the participants about the aggregation level of the model. The facilitators should somehow try to mediate that the system portrays individuals that may have their own reasons for acting as they do and that they may not act according to the model. It is also important to communicate to the participants that the model represents a theory of a system and that the structure of the system is partly constructed by the agents in the system at the same time as their actions are both, constrained and enabled by the system.

In the design of an educational context for the Quito prototype it would be important to communicate that the assumptions that are made in the underlying simulation model are simplified and could not be relied on for detailed decision-making. The use of video to describe assumptions that are made in the model could lead the participants to confuse the assumptions made in the model with behavior of the real system. It would therefore be important for the facilitators to communicate that this would not necessarily be the case. The participants should be encouraged to discuss the assumptions made in the model and also to criticize them and find alternative assumptions that could influence the conditions of the real city. With these general comments about representation of the underlying model of the Quito prototype I turn to some more specific issues regarding the design of the user interface.

## 10.6 Suggestions for Design of the a Future Version of the Quito Prototype

In designing an extended prototype for use in a system dynamics course a number of issues raised in this chapter needed to be taken into consideration. These included the particularities of the underlying model as described in Section 10.2, lessons learned from the Two-Shower prototype (e.g. regarding the use of narrative, graphics, two users etc.), and the concerns regarding models of social systems. In this section I consider how the user interface may be designed with the intention to support the participants in navigating through the model, how transparency may be implemented, and how some of the model elements may be presented by use of multimedia representations and narrativity. These suggestions are not implemented as part of the current user interface. If a complete interactive learning environment were to be constructed, the simulation model would require further testing and revision, and a complete plan for design and testing of the user interface would have to be made. This is outside the scope of this project.

### 10.6.1 A Single or Multi-User Prototype and Support for Collaboration

The Two-Shower prototype was, as discussed previously, a collaborative system and the opportunity to collaborate was well received by the participants in Test 2 (see Section 8.2). The implementation of the future version of the Quito prototype as a collaborative system would have both, pros and cons. A multi-user prototype may be more engaging for the participants, but it may also be more difficult to understand the system when the actions of the others must be taken into account in the analysis of the model. This may be a problem for a large model such as the Quito model. It may be, for example, more difficult to test a policy if input to the model is made by others if one is not aware of this input and its effect on the system. Another participant may change variables and thereby change the basis on which the policy is being tested. In real systems the input of others would of course be present and this should be something that system dynamics students should learn to take into account.

If support for collaboration should be incorporated into a future version of the Quito prototype it could be done in two main ways. One would be to organize the activities around the use of the model as collaborative activities. The other is to turn the model itself into a collaborative system.

A facilitator could have a more advanced user interface with more information that could help them guide the participants. This is suggested in Davidsen (2000). The underlying simulation model would then be shared but the participants and the facilitator could have different user interfaces.

If the prototype was developed into a multi-user system, the participants could also play different roles. Examples would be chief police officer, street vendor, mayor etc. The participants could have access to different user interfaces that provided access to different information according to the information that their characters would be likely to have. They could also have access to control of different variables in the system. There could also be different methods for communicating with the other participants in the system. This, however, would require more work regarding the intended setting. Group discussions before, after, and in between running the model would have to be coordinated. The incorporation of support for collaboration in this manner may be suitable for a more general discussion about resource sharing or problems related to cities. It may be questionable, however, whether it is an adequate way to learn about system dynamics.

### 10.6.2 Narrativity in a Future version of the Quito Prototype

In the Two-Shower prototype, elements of narrative were incorporated into the user interface. To recap, the setting was at a hotel and the problem was introduced as a story of a tired person who checked into a hotel and wanted to take a shower. The two simulated persons in the showers were characterized by photos of a woman and a man in the shower.

The Quito model is also suitable for including elements of narrativity. Its characteristics make the creation of a setting and characters possible. In this case, the setting is a city where there is a problem of allocating resources to different sectors of the city. The participant plays the role of the mayor who controls the development of the city by allocating resources. In the current version of the Quito prototype the citizens and tourists of the city are characterized through video clips with actual citizens and tourists of Quito who talk about their view on some of the city government plans and decisions and how they experience changes to the city in their daily lives.

### 10.6.3 A General Discussion about Navigation in a Future Version of the Quito Prototype

As already discussed, the underlying model of the Quito prototype is a more complex and larger model than the Two-Shower model. The level of detail that is necessary for the intended audience, i.e. people who are expected to learn system dynamics, demands that the user interface of a future version of the Quito prototype displays a larger amount and more complex information than for the Two-Shower prototype.

For a future, improved version of the Two-Shower prototype, I suggested seven views that address various aspects of the model (see Section 9.2). For a small model like the Two-Shower model, this organization of the user interface may be suitable. For a larger model like the Quito model, a more advanced method for navigation and organization of views may be appropriate.

Because the Quito model is so much larger and the goal is that each view must not contain more information than each view of the Two-Shower prototype, a different system for navigation must be constructed. Tone could divide the user interface in the same manner as suggested in the improved Two-Shower model by way of a menu providing access to several views that represented, respectively, structure, behavior, a story generator (relating structure to behavior), and, possibly, the stock and flow diagram. For each view, there could be an additional menu for viewing the different sectors of the model, such as, for example, the population sector, the historic building sector etc. (see figure 10-13). An alternative representation of model structure could be to start with an aggregate view of the model structure where participants could click on parts of the model and, thereby, be represented with a more detailed version of this particular section.


*Figure 10-13: Example of organization of the user interface in a future version of the Quito prototype.*

In the discussion of the Two-Shower prototype, I addressed the necessity of revealing the model structure to the participants on the basis of the intended user group of the system. Because the future version of the Quito prototype would be intended for students of system dynamics, revealing causal loops diagrams, stock and flow diagrams or equations would be appropriate because a thorough investigation of the underlying model could be helpful for understanding the model.

Information about what lies behind the equations would also be important. In Section 10.5 I briefly emphasized that the underlying assumptions of the model should be represented in the user interface. With the division of the user interface into several views, where the participants can go deeper into the details of the underlying model, there is also room for a representation of the underlying assumption and for the users to validate the assumptions that are made in the model.

### 10.6.4 Representation of Structure in a Future Version of the Quito Prototype

As discussed previously, in the Two-Shower prototype the first view that involved simulation was the Shower Room view, which was mainly a representation of the behavior of the model. The structure was introduced in the Pipe System view and the story generator provided a link between structure and behavior. Since the Quito model has a much more complex structure than the Two-Shower model,it may be more adequate to introduce an overview of the model structure before the model is subjected to simulation. A suggestion for the Quito prototype, made in order to avoid confusing the participants, could be, prior to running the model. to introduce a causal loop diagram, that shows how the elements of the model are interelated in a feedback structure. . Some of the participants in Test 2 of the Two-Shower prototype reported that it was difficult to understand what was happening in the Shower Room view. It was commented also, however, that the advantage of the lack of structural information made them try to infer the structure based on information about the behavior (see Section 8.2).

Another suggestion could be to introduce some of the characteristics of the model as a story in the same manner that it was in the introductory story of the Two-Shower prototype.

As discussed previously, also the representation of structure must be implemented in such a way that there is a possibility for the participants to go deeper into the structure of the model and explore the details of that structure. The participants could, for example, press a representation of the population and receive a new view with a more thorough explanation of this part of the model. Text and graphics could be used to explain this part of the structure. The participant could then click on a variable to obtain additional and more detailed views of the model. In the end part of the causal loop diagram could be reached, one that gave access to the model equations. The participants should probably only gain access to these views when the model is paused or before the simulation starts to avoid a cognitive overload.

### 10.6.5 Representation of Behavior in a Future Version of the Quito Prototype

In the Two-Shower prototype, animation, changing colors, and photo slideshows were used to represent the behavior of the model. This could also be done in the case of the Quito prototype. It may be difficult for the participants to monitor the model as a whole. Figure 10-14 is an example of how the user interface may be divided into different views, each representing the various parts of the model. This is an example of an overview of the model behavior. The participants would have to enter the more detailed views in order to study the behavior thus far of the different parts of the model more closely.

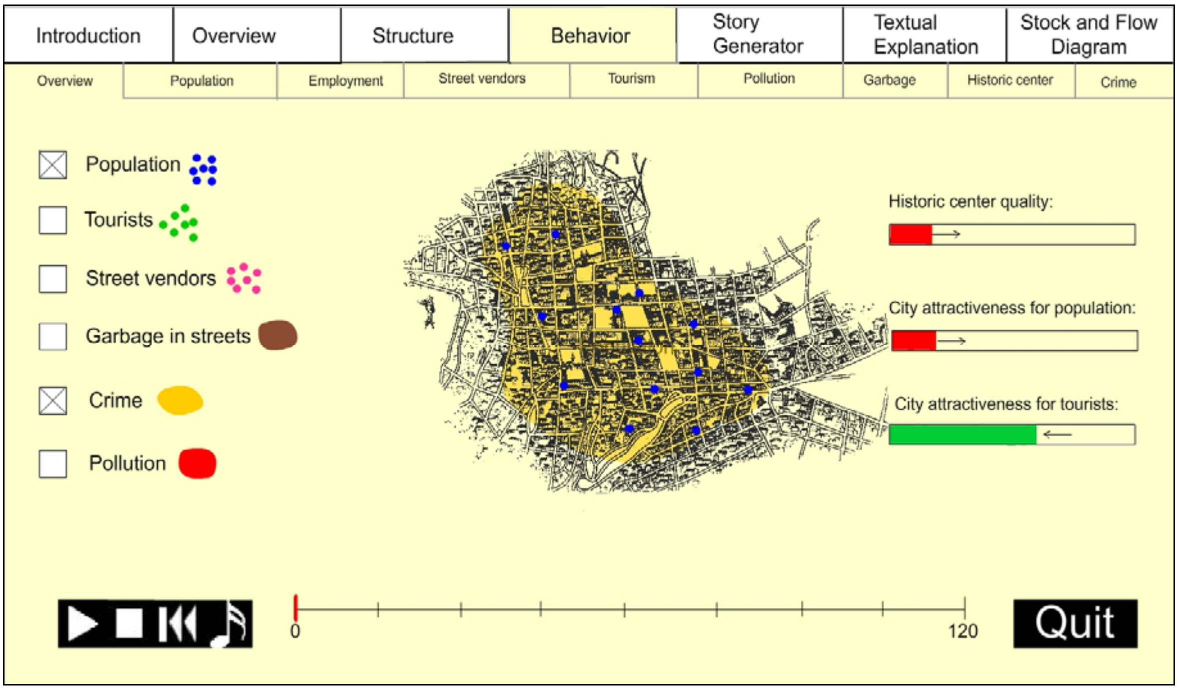


*Figure 10-14: Example of representation of an overview of the behavior in a future version of the Quito prototype.*

In the current version of the Quito prototype the map displays a number of dots to indicate its state, i.e. the number of street sellers and tourists in the city center (see figure 10-11 p. 302). Indications of the states of other variables may also be mapped in a new version of the prototype. Spence (2001) refers to Colby and Scholl (1991) who invented the *Z-thru mapping technique* (Spence, 2001, p. 122). They used a map of Boston and demonstrated how this could be covered by a number of transparent information layers, where each layer represented information about some feature of the city, such as crime statistics or traffic density. This is also utilized in SimCity where the players may choose to activate a small map of the city and choose between layers indicating, for example, the pollution level, the water or electricity coverage, or the attractiveness of various geographical areas of the city.

The map of Quito could be extended with layers, for example, with information about the current population level, number of tourists, density of street vendors, garbage in the streets, and pollution. The participants could click to view the different layers and compare them. In this way, features utilized in computer games may also be utilized in interactive learning environments.

The division of the city into different geographical areas with different levels of, for example, pollution and street vendors in the various areas has not been considered for the Quito prototype. This would make the model much more complex and it would be even more difficult for the participants to understand how the underlying structure cause the resulting behavior. For the purpose of learning about complex systems or system dynamics, this could make it unnecessary complex.

In the Two-Shower prototype the photo slideshows of the changing face expressions were used to represent the *success* of the simulation or the *performance* of the participant. Similarly, there should be some indicators of how the simulation is progressing, with the Quito prototype. This could be shown at all times and for all views so that the participant can decide to switch between the various views portraying different model characteristics in order to observe what is happening and try to control the simulation model in the desired direction. The *happy-meters* in the right hand side in figure 10-14 shows how this may be implemented in a future version of the Quito prototype. This feature is also part of SimCity.

### 10.6.6 The Story Generator in a Future Version of the Quito Prototype

For the Two-Shower prototype a story generator was developed which analyzed the structural characteristics of the model and gave feedback to the participants about input which had been made and how it affected the model. The Quito model is, as discussed previously, much more complex and there may be a need for a more complex story generator where the participant can go back and trace the structural origin of the behavior in time. A story of what went wrong, how it started, and the implications it had on the behavior of the system may be told. The effects of decisions that were made may be explained. The participant could be guided through the structure and the behavior with the purpose of developing an understanding of the relationship between them.

A more advanced story generator may be based on the eigenvalue analysis. I will not elaborate on this here but note that it is a mathematical analysis of the model which can be used to identify the main structural links at each point in time. The eigenvalues for the controllable variables may, for example, show the effect that each would have at different times. For a more indepth discussion of eigenvalues see Saleh (2002).

This chapter was a presentation of the Quito prototype and a discussion of how the experiences and findings from the Two-Shower prototype could be used in the design of a future version of the Quito prototype. These issues of transferability and future work are also the focus of the final chapter of this thesis.

# 11. Lessons Learnt: Transferability

## 11.1 Introduction

Future work for the two prototypes described in this thesis has been discussed in the last three chapters. Chapter 8 was a discussion of the results from Test 2 of the Two-Shower prototype and the indications that the feedback gave for future development of the prototype. Further development was also considered in Chapter 9. Specific suggestions for further development of the user interface and a proposed future study of the prototype were proposed. In Chapter 10, I discussed how the feedback about the Two-Shower prototype could be used in the development of a prototype based on larger model of a social system. I have chosen in my final chapter, therefore, to offer very briefly some further comments on transferability (Section 11.2). I end with Section 11.3, which is a brief comment on how I would have approached and carried out this project if I were to start again and a return to the need for further research.

## 11.2 Transferability

The discussion of using the experiences and feedback about the Two-Shower prototype in the development of the Quito prototype indicated that there were some possibilities as well as hindrances in transferring the design ideas to other interactive learning environments of complex systems. By *transferable* I mean that the theory and design ideas largely can be used to make design decisions about other system dynamic-based interactive learning environments.

Also to be considered are what further types of models should be visualized to test transferability. Examples could be other models of resource-sharing, models dealing with business-related problems, or models of environmental problems.

### 11.2.1 Transferability of Narrative Characteristics

In Chapter 3 I discussed how narrative characteristics were included in the Two-Shower prototype. A narrative setting and characters were included both in the Two-Shower prototype and in the Quito prototype, however, in different ways. In the Two-Shower prototype the participants can *control* the character while in the Quito prototype they *play* the character. For the Quito prototype two options were discussed in Chapter 10, the current single-user version where the participant takes the role of the mayor, and a multi-user version where a group of people could play different roles in the system, each controlling different variables.

Some problems of narrative characteristics were discussed in Chapter 8. Some of the participants in Test 2 of the Two-Shower prototype reported that the use of neutral photos in the slideshows was confusing. The slideshows included neutral photos, such as water running from the showerhead, in order to prevent the photos from switching directly from, for example, too hot to too cold water. These were included with principles of film editing in mind. Some of the participants found these neutral photos superfluous, as they did not add information about the conditions of the system. There may thus be a contradiction between the inclusion of certain narrative characteristics and representing information about the underlying model.

In addition, the inclusion of computer game characteristics could lead the participants to take the role of players rather than learners (Alessi, 2002). This was also observed in Test 2 of the Two-Shower prototype where some of the participants stopped after having reached a desirable shower temperature, stating that they had solved the task.

Experiences with the two prototypes show that narrative characteristics and elements of computer games should not be incorporated uncritically into the user interface but with a close consideration of how it represents the underlying model with the purpose of supporting the users' understanding. It is important to consider whether the inclusion of narrative characteristics in some ways may also limit understanding, as discussed in Chapter 8.

### 11.2.2 Transferability of the Story Generator

Technically the story generator could be used and adjusted for similar models. Frotjold (2005), however, notes in her thesis that it cannot yet be used to explain *n-ordered* delays (in contrast to a *pipeline* delay an *n-ordered* delay changes for every time step). The original Two-Shower model contains a *pipeline* delay where the delayed variable will change as a result of a change in another variable at a specific point in time. A suggestion for representing *n-ordered* delays would be to describe the delay at both the starting variable and the ending variable. This would involve a more complex analysis of the equations in the system (ibid.).

Visually and textually, the story generator could be transferable to other models. As was observed in Test 2 of the Two-Shower prototype, however, the amount of text that should be included must be considered based on the intended audience. It would also be important to provide time for the participants to read the text so that they would not get the feeling of lagging behind when they pause the system to read the text.

### 11.2.3 Transferability of Use of Multimedia and Visualization

Chapter 4 was a discussion of presentation and visualization in the Two-Shower prototype. How multimedia was utilized to enhance the characteristics of the underlying model was raised. The use of multimedia and visualization can also be transferred to some extent to other models.

Animation and the use of color could be used to illustrate the behavior and dynamics in other interactive learning environments. Different values would have to be defined with different color codes. The representation of the pipeline delay of the Two-Shower prototype could also be used to represent other pipeline delays. It could be used, for example, to represent the delay of a pipeline production. Different colors could represent the number of what was being manufactured at each stage in the production process.

Sound was utilized in the Two-Shower prototype to warn the participants of critical values in the water temperature. Sound could be utilized in a similar manner for other systems, but it is important to note that it should be used with caution and that the sound recording should fulfill a certain level of quality. The experience from the Two-Shower prototype was that when the quality was poor it rather contributed to disturb the participants. The experience was also that the background noise could be particularly annoying. This was included as part of the contribution to the narrative characteristics and show again that how it supports the representation of the underlying model must be considered when narrative characteristics are incorporated.

As discussed earlier in this thesis, photo slideshows were used in the Two-Shower prototype and video was incorporated into the Quito prototype. One of the problems of the photos was that some of the participants claimed that it was difficult to understand the difference between the hot and cold photos. Caution should also be shown when including photos, and one should ensure that the characteristics or values they represent cannot be confused with other characteristics or values of the system.

Video was used to represent *verstehen*-type questions (see Section 10.5). Video may also be used for other models but it is important to consider how the use of video is intended to support the understanding of the underlying model. Although not tested in this thesis, too extensive use of video material may not necessarily support understanding but may be useful for encouraging participants to think more widely about the problem.

### 11.2.4 Transferability of Model Transparency

The models discussed in this thesis were divided into different views, which represented different characteristics of the underlying model. The division of the learning environments into different views, each providing a different focus of the underlying model, can be tested also for other models.

In Chapter 5 I discussed model transparency for the Two-Shower prototype. The behavior was first introduced to trigger the curiosity of the participants, followed by the structure in the Pipe System view, and the link between structure and behavior in the story generator. The visualizations were first relatively aggregate but concrete (for example photos of a real person who is taking a shower), representing information more closely related to what one would normally experience in a shower. The information that was provided about the underlying model gradually became more abstract but at the same time more detailed as one would move on to the later views of the prototype. The creation of a user interface that gave a gradual presentation of the underlying model can also be used in the development of other models. In Chapter 10 I suggested that a quick overview of the main model components and an aggregation of the model structure may be used for visualizations of large models such as the Quito model.

### 11.2.5 Transferability of Support for Collaboration

As discussed in Chapter 5 support for collaboration was provided through visualization in the Two-Shower prototype. The design of a simulation-based collaborative learning environment, letting the participants each control a part of the model, is not necessarily transferable to all other models. As was discussed in Chapter 10 with a large model like the Quito Model, for example, it may be difficult to see the consequences of decisions that are made in the system because there are other participants who also make decisions and change variables. The implementation of support for collaboration could depend, however, on the aim of the interactive learning environment. If it were, for example, to learn that the decisions of other actors have an influence, support for collaboration would be an advantage. If it is to learn the details of the particular model, it may not, at least not for a first run-through of the model.

If a model was intended to be used for collaboration, however, it could be possible to test some of the theories and design decisions that were used for the Two-Shower prototype. The concept of visualizing the decisions and conditions of the other participant could also be tested for other models. In such a case, it would be important to consider how the variables controlled by the other participant and the decisions of the other participant should be introduced and at what level of detail. This would be linked to the discussion of model transparency.

### 11.2.6 Transferability of Visualization of Diversity and Aggregation Level

In Chapter 10 I raised also concern about model diversity and that one should seek to portray this diversity through the visualizations. This would also be transferable to other models, not only models of social systems. The discussion of system dynamics and its relation to social theory in Section 10.5 concluded that it would be important to inform the participants of the Quito prototype of the aggregation level of a model that portrayed a social system.

The above are brief comments on the possibilities and limitations of transferability to other models. Further development and evaluations are obviously necessary on a variety of models to test the initial conclusions. This falls, however, outside the scope of this thesis.

## 11.3 Hindsight and Lessons Learnt

I have commented extensively in the last three chapters on changes and modifications if the prototypes were re-developed. My last few comments refer to what I would do if I were to start all over again. In hindsight, I would have made some changes to the development process related to the use of time, resources, and choice of models.

I would have used more time planning the user interface and on developing the visual representations of the underlying model. I would also have spent more time planning and carrying out the evaluation test. I would have timed the test according to the university semester in order to be able to get participants who were within the target audience of the prototype.

In retrospect, there should have been more programming resources in the project, as too much time was spent on programming. (The lack of programming resources was partly due to the bankruptcy of the initial business partner). Ideally, the programmers should have been paid so that both PhD students and Master's students could spend their time on developing the concepts and the visualizations.

The design process could also have been easier if it was clearer at an earlier phase in the project when the prototype was to be used, e.g. at start or middle of a course. Even better would have been to know exactly what course material the students would/should have been introduced to before running the model.

The Quito model was developed within the VOCS project group. To save time and to improve the quality of the prototype I would have chosen a system dynamic model that had already been developed and that had gone through extensive testing.

Despite my concerns raised above, I believe that this thesis has offered a contribution to the study of visualizations of complex systems. Obviously this is a small step on the road to developing visualizations of complex systems and far more research and prototype development needs to be carried out. I hope, however, that the work done within the VOCS project, and this thesis, may contribute even in a small way to this exploration.

Magnhild Viste

September, 2007

# References


Akkermans, H. (2001). Renga, A Systems Approach to Facilitating Inter-Organizational Network Development. *System Dynamics Review, 17*(3), 179-194.

Alba, J. W., & Hasher, L. (1983). Is Memory Schematic? *Psychological Bulletin, 93*, 203-231.

Alessi, S. M. (2002). *Model Transparency in Educational System Dynamics.* Paper presented at the 20th International Conference of the System Dynamics Society, Albany, New York.

Alessi, S. M., & Trollip, S. R. (2001). *Multimedia for learning : methods and development* (3rd ed.). Boston: Allyn and Bacon.

Anderson, J. R. (1995). *Cognitive psychology and its implications* (4th ed.). New York: W.H. Freeman.

Anderson, J. R. (2005). *Cognitive psychology and its implications : John R. Anderson* (6th ed.). New York: Worth Publishers.

Arnseth, H. C., & Solheim, I. (2002). *Making Sense of Shared Knowledge.* Paper presented at the Foundations for a CSCL community, CSCL 2002, Boulder, Colorado, USA.

Arquitt, S., & Johnstone, R. (2004). A Scoping and Consensus Building Model of a Toxic Blue-Algae Bloom. *System Dynamics Review, 20*(2), 179-198.

Backus, G., Schwein, M. T., Johnson, S. T., & Walker, R. J. (2001). Comparing Expactations to Actual Events: the Post Mortem of a Y2K Analysis. *System Dynamics Review, 17*(3), 217-137.

Baker, M., Hansen, T., Joiner, R., & Traum, D. (1999). The Role of Grounding in Collaborative Learning Tasks. In P. Dillenbourg (Ed.), *Collaborative Learning: Cognitive and Computational Approaches* (pp. 31-63). Amsterdam: Pergamon.

Barlas, Y. (1996). Formal Aspects of Model Validity and Validation in System Dynamics. *System Dynamics Review, 12*(3), 183-210.

Berger, P. L., & Luckmann, T. (1966). *The social construction of reality : a treatise in the sociology of knowledge* (Anchor Books ed.). Garden City, N. Y.: Doubleday.

Bianchi, C. (2002). Opportunities and Pitfalls Related to E-Commerce Strategies in Small-Medium Firms: A system Dynamics Approach. *System Dynamics Review, 18*(3), 403-429.

Bordwell, D., & Thompson, K. (1997). *Film art : an introduction* (5th ed.). New York: McGraw-Hill.

Burrell, G., & Morgan, G. (1979). *Sociological paradigms and organisational analysis elements of the sociology of corporate life*. London ,: Heineman.

Campbell, D. (2001). The Long and Winding (and Frequently Bumpy) Road to Successful Client Engagement: One Team's Journey. *System Dynamics Review, 17*(3), 195-217.

Chaiklin, S. (2007). Modular or Integrated? - An Activity Perspective for Designing and Evaluating Computer-Based Systems. *International Journal of Human-Computer Interaction, 22*(1 & 2), 173-190.

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988, May 15-19). *Development of a Tool Measuring User Satisfaction of the Human-Computer Interface.* Paper presented at the ACM SIGCHI '88 Conference on Human Factors in Computing Systems, Washington D.C., USA.

Colby, G., & Scholl, L. (1991, March, 1991). *Transparency and blur as selective cues for complex visual information.* Paper presented at the SPIE Conference.

Craik, K. J. W. (1966). The Mechanism of Human Action. In S. L. Sherwood (Ed.), *The Nature of Psychology: A Selection of Papers, Essays and other Writings by the Late Kenneth J. W. Craik* (pp. 1-90). Cambridge: Cambridge University Press.

Davidsen, P. I. (1994). The Systems Dynamics Approach to Computer-Based Management Learning Environments: Implications and their Implementations in Powersim. In J. D. W. Morecroft & J. D. Sterman (Eds.), *Modeling for Learning Organizations* (pp. 301-316). Portland: Productivity Press.

Davidsen, P. I. (2000). Issues in the Design and Use of System-Dynamics-Based Interactive Learning Environments. *Simulation & Gaming, 31*(2), 170-177.

Davidsen, P. I., Spector, J. M., & Milrad, M. (1999, July 20-23). *Learning in and about Simple Dynamic Systems.* Paper presented at the 17th International Conference of the System Dynamics Society, Wellington, New Zealand.

de Jong, T., & van Joolingen, W. R. (1998). aDiscovery Learning with Computer Simulations of Conceptual Domains. *Review of Educational Research, 68*(2), 179-201.

Diehl, E., & Sterman, J. D. (1995). Effects of Feedback Complexity on Dynamic Decision Making. *Organizational Behavior and Human Decision Processes, 62*(2), 198-215.

Dix, A., Finlay, J., Abowd, G., & Beal, R. (1998). *Human-computer interaction* (2nd ed.). London: Prentice Hall Europe.

Djupvik, N. M. (2006). *The Simulation Visualization Lab - Understanding Complex Systems using Video Narratives.* Unpublished Master's thesis, University of Bergen, Bergen.

Djupvik, N. M., & Frotjold, L. (2003). *The Quito Model.* Unpublished term project, University of Bergen, Bergen.

Doyle, J. K., & Ford, D. N. (1998). Mental Models Concepts for System Dynamics Research. *System Dynamics Review, 14*(1), 3-29.

Doyle, J. K., & Ford, D. N. (1999). Mental Models Concepts Revisited: Some Clarifications and a Reply to Lane. *System Dynamics Review, 15*(4), 411-415.

Doyle, J. K., Ford, D. N., Radzicki, M. J., & Trees, S. W. (2001). *Mental Models of Dynamic Systems*, 2005

Doyle, J. K., Radzicki, M. J., & Trees, S. W. (1998). *Measuring Change in Mental Models of Dynamic Systems: an Exploratory Study*.Unpublished manuscript.

Dyck, J., Pinelle, D., Brown, B., & Gutwin, C. (2003). *Learning from Games: HCI Design Innovations in Entertainment Software.* Paper presented at the Graphics Interface, Halifax, Nova Scotia.

Dörner, D. (1997). *The logic of failure : recognizing and avoiding error in complex situations*. Reading, Mass.: Perseus Books.

Faust, L. J., Jackson, R., Ford, A., Earnhardt, J. M., & Thompson, S. D. (2004). Models for Management of Wildlife Populations: Lessons from Spectacled Bears in Zoos and Grizzly Bears in Yellowstone. *System Dynamics Review, 20*(2), 163-178.

Feltovich, P. J., Spiro, R. J., Coulson, R. L., & Feltovich, J. (1996). Collaboration Within and Amoung Minds: Mastering Complexity, Individually and in Groups. In T. Koschmann (Ed.), *CSCL: Theory and Practice*. Mahwah, New Jersey: Laurence Erlbaum Associates.

Ford, A. (1999). *Modeling the environment : an introduction to system dynamics models of environmental systems*. Washington, D.C.: Island Press.

Forrester, J. W. (1961). *Industrial dynamics* (Students' ed.). Cambridge, Mass.: M.I.T. Press.

Forrester, J. W. (1969). *Urban dynamics*. Cambridge, Mass.: M.I.T. Press.

Forrester, J. W. (1971a). Counterintuitive Behavior of Social Systems. *Technology Review, 73*(3), 52-68.

Forrester, J. W. (1971b). *World dynamics*. Cambridge, Mass.: Wright-Allen Press.

Forrester, J. W. (1980). System Dynamics - Future Opportunities. *System Dynamics, 14*, 7-21.
Forrester, J. W. (1994a). *Learning through System Dynamics as Preparation for the 21st Century.* Paper presented at the Systems Thinking and Dynamic Modeling Conference for K-12 Education, Concord, MA.
Forrester, J. W. (1994b). Policies, Decisions, and Information Sources for Modeling. In J. D. W. Morecroft & J. D. Sterman (Eds.), *Modeling for Learning Organizations* (pp. 51-85). Portland, OR: Productivity Press.
Forrester, J. W. (1997). *System Dynamics and K-12 Teachers* (Lecture article for lecture at the University of Virginia School of Education). Cambridge, MA.
Forrester, J. W., & Senge, P. M. (1980). Tests for Building Confidence in System Dynamics Models. *System Dynamics, 14*(1980), 209-228.
Frotjold, L. (2005). *Explanation Generation for Complex Systems.* Unpublished Master thesis, University of Bergen, Bergen.
Glaeser, E., & Sacerdote, E. (1999). Why is there more Crime in Cities? *Jourbal of Political Economy, 107*(6), 225-258.
Goldvarg, E., & Johnson-Laird, P. N. (2001). Naive Causality: a Mental Model Theory of Causal Meaning and Reasoning. *Cognitive Science, 25*(2001), 565-610.
Größler, A. (1997). *Giving the Black Box a Lid - Providing Transparency in Management Simulations.* Paper presented at the 16 th International Conference of the System Dynamics Society, Istanbul, Turkey.
Größler, A. (1998). *Structural Transparency as an Element of Business Simulators.* Paper presented at the 16th International Conference of the System Dynamics Society, Quebec City, Canada.
Größler, A., Maier, F. H., & Milling, P. M. (2000). Enhancing Learning Capabilities by Providing Transparency in Business Simulators. *Simulation & Gaming, 31*(2), 257-278.
Guevara-Chaves, P., & Perez-Bennet, A. (2002). *Planning of the Historic Centre of Quito: A System Dynamics Approach for a Persistent Issue.* Unpublished term project, University of Bergen, Bergen.
Homer, J., Hirsch, G., Minniti, M., & Pierson, M. (2004). Models for Collaboration: How System Dynamics Helped a Community Organize Cost-Effective Care for Chronic Illness. *System Dynamics Review, 20*(3), 199-222.
Howie, E., Sy, S., Ford, L., & Vicente, K. J. (2000). Human-computer interface design can reduce misperceptions of feedback. *System Dynamics Review, 16*(3), 151-171.
Hutchins, E. (1995). *Cognition in the wild.* Cambridge, Mass.: MIT Press.
IMQ. (1991). *Direccion de Planificacion* (No. 1A-1D). Quito: IMQ (Illustre Municipio de Quito).
INEC. (2001). Retrieved August 9, 2006, from http://www.inec.gov.ec/default.asp
Jensen, E., & Brehmer, B. (2003). Understanding and Control of a Simple Dynamic System. *System Dynamics Review, 19*(2), 119-138.
Jensen, J. F. (2000). Interaktivitet & Interaktive Medier - med et Postscript om Interaktivitet og Læring. In G. Haugsbakk & Y. Fritze (Eds.), *Workshop: Interaktivitet, teknologi og læring* (pp. 29-87). Oslo: Unipub, Akademia AS.
Johnson-Laird, P. N. (1983). *Mental models : towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.
Johnson-Laird, P. N. (1989). Mental Models. In M. I. Posner (Ed.), *Foundations of Cognitive Science* (pp. 469-499). Cambridge, MA: MIT Press.

Jonassen, D. H., Hernandez-Serrano, J., & Choi, I. (2000). Integrating Constructivism and Learning Technologies. In J. M. Spector & T. M. Andersen (Eds.), *Integrated and Holistic Perspectives on Learning, Instruction and Technology - Understanding Complexity* (pp. 103-128). Dordrecht Boston: Kluwer Academic Publishers.

Jones, G. A., & Bromley, R. D. F. (1996). The Relationship between Urban Conservation Programmes and Property Renovation: Evidence from Quito, Ecuador. *Cities, 13*(6), 373-385.

Jordan, B., & Henderson, A. (1995). Interaction Analysis: Foundations and Practice. *The Journal of the Learning Sciences, 4*(1), 39-103.

Juul, J. (1998, November, 1998). *A Clash between Game and Narrative.* Paper presented at the Digital Arts and Culture (DAC), Bergen, Norway.

Juul, J. (2003). *Just what is that makes computer games so different, so appealing?* Retrieved February 15, 2006, from http://www.igda.org/columns/ivorytower/ivory_Apr03.php

Kashihara, A., Kinshuk, Oppermann, R., Rashev, R., & Simm, H. (2000). A cognitive load reduction approach to exploratory learning and its application to an interactive simulation-based learning system. *Journal of Educational Multimedia and Hypermedia, 9*(4), 253-276.

Kelley, H. H. (1973). The process of causal attribution. *American Psychologist, 28*(2), 107-128.

Kirkpatrick, D. L. (1996). *Evaluating Training Programs*. San Francisco.

Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and instruction, 12*(1), 1-10.

Klabbers, J. H. G. (2000). Learning as Acquisition and Learning as Interaction. *Simulation & Gaming, 31*(3), 380-406.

Kleinmuntz, D. N. (1985). Cognitive Heuristics and Feedback in a Dynamic Decision Environment. *Management Science, 31*(6), 680-702.

Koschmann, T. (1996a). *CSCL : theory and practice of an emerging paradigm*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Koschmann, T. (1996b). Paradigm Shifts and Instructional Technology: An Introduction. In T. Koschmann (Ed.), *CSCL: Theory and Practice of an Emerging Paradigm* (pp. 1-25). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Koschmann, T., Kelson, A. C., Feltovich, P. J., & Barrows, H., S. (1996). Computer-Supported Problem-Based Learning: A Principled Approach to the Use of Computers in Collaborative Learning. In T. Koschmann (Ed.), *CSCL: Theory and Practice of an Emerging Paradigm* (pp. 83-124). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Koschmann, T., & LeBaron, C. D. (2003, 14-18 September). *Reconsidering Common Ground: Examining Clark's Contribution Theory in the OR.* Paper presented at the ECSCW 2003: the Eight European Conference on Computer Supported Cooperative Work, Helsinki, Finland.

Koschmann, T., & LeBaron, C. D. (2003, September 14-18). *Reconsidering Common Ground: Examining Clask's Contribution Theory in the OR.* Paper presented at the ECSCW 2003: the Eight European Conference on Computer Supported Cooperative Work, Helsinki, Finland.

Lane, D. C. (1999). Friendly Amendment: A Commentary on Doyle and Ford's proposed Re-definition of 'Mental Model'. *System Dynamics Review, 15*(2), 185-194.

Lane, D. C. (2000a, August 6th-10th 2000). *Rerum Cognoscere Causas: Part I - How do the Ideas of System Dynamics Relate to Traditional Social Theories and the Voluntarism/Determinism Debate?* Paper presented at the 18th International Conference of the System Dynamics Society, Bergen, Norway.

Lane, D. C. (2000b, August 6th-10th 2000). *Rerum Cognoscere Causas: Part II - Opportunities Generated by the Agency/Structure Debate and Suggestions for Clarifying the Social Theoretic Position of System Dynamics.* Paper presented at the 18th International Conference of the System Dynamics Society, Bergen, Norway.

Lane, D. C. (2001a). *Rerum Cognoscere Causas:* Part I - How Do the Ideas of System Dynamics Relate to Traditional Social Theories and the Voluntarism/Determinism Debate? *System Dynamics Review, 17*(2 summer), 97-119.

Lane, D. C. (2001b). *Rerum Cognoscere Causas:* Part ll - Opportunities Generated by the Agency/Structure Debate and Suggestions for Clarifying the Social Theoretic Position of System Dynamics. *System Dynamics Review, 17*(4), 293-311.

Lane, D. C., & Oliva, R. (1998). The Greater Whole: Towards a Synthesis of System Dynamics and Soft Systems Methodology. *European Journal of Operational Research, 107*(1), 214-135.

Lave, J., & Wenger, E. (1991). *Situated learning : legitimate peripheral participation*. Cambridge: Cambridge University Press.

Lewis, J. R. (2002). Psychometric Evaluation of the PSSUQ Using Data from Five Years of Usability Studies. *International Journal of Human-Computer Interaction, 14*(3&4), 463-488.

Liddell, W. G., & Powell, J. H. (2004). Agreeing Access Policy in a General Medical Practice: A Case Study using QPID. *System Dynamics Review, 20*(1), 49-73.

Lipponen, L. (2002, January 7th-11th 2002). *Exploring Foundations for Computer-Supported Collaborative Learning.* Paper presented at the Computer Support for Collaborative Learning: Foundations for a CSCL community. Proceedings of the Computer-supported Collaborative Learning 2002 Conference, Boulder, Colorado.

Lyneis, J. M., Cooper, K., G, & Els, S. A. (2001). Strategic Management of Complex Projects: A Case Study using System Dynamics. *System Dynamics Review, 17*(3).

Machuca, J., A. D. (2000). Transparent-Box Business Simulators: An Aid to Manage Complexity of Organizations. *Simulation & Gaming, 31*(2), 230-239.

Mack, R. L., & Nielsen, J. (1994). Executive Summary. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods* (pp. 1-23). New York: John Wiley & Sons, Inc.

Mangurian, D. (1999). *All Dressed Up and Looking for Investors*. Retrieved August 9, 2006

Manovich, L. (2001). *The language of new media*. Cambridge, Mass.: MIT Press.

Mayer, R. E. (1989). Models for Understanding. *Review of Educational Research, 59*(1), 43-64.

Mayer, R. E. (2001). *Multimedia Learning*. New York: Cambridge University Press.

Mayer, R. E. (2005). Introduction to Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 1-18). New York, NY: Cambridge University Press.

Mayer, R. E., & Anderson, R. B. (1991). Animations need Narrations: An Experimental Test of a Dual-Coding Hypothesis. *Journal of Educational Psychology, 83*(4), 484-490.

Mayo, D., D., Callaghan, M., J., & Dalton, W. J. (2001). Aiming for Restructuring Success at the London Underground. *System Dynamics Review, 17*(3), 261-291.

Meadows, M. S. (2002). *Pause & effect : the art of interactive narrative*. Indianapolis: New Riders Publishing.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and Perception*. Cambridge, MA: Harvard University Press.

Milrad, M. (2006). *Designing Interactive Learning Environments to Support Learners' Understanding in Complex Domains.* Unpublished PhD, University of Bergen, Bergen.

Milrad, M., Spector, J. M., & Davidsen, P. I. (2003). Model Facilitated Learning. In S. Naidu (Ed.), *Learning and Teaching with Technology* (pp. 13-27). London: Kogan Page Limited.

Mojtahedzadeh, M., & Andersen, D. (2001). *Digest: A New Tool for Creating Insightful System Stories*. Retrieved March 6th., 2003, from http://www.albany.edu/cpr/sdgroup/pad824/DigestF.pdf

Mojtahedzadeh, M., Andersen, D., & Richardson, G. P. (2003, July 20-24). *Using Digest to Implement the Pathway Participation Method for Detecting Influential System Structure.* Paper presented at the 21st International Conference of the System Dynamics Society, New York City, USA.

Morecroft, J. D. W. (1994). Executive Knowledge, Models, and Learning. In J. D. W. Morecroft & J. D. Sterman (Eds.), *Modeling for Learning Organizations* (pp. 3-29). Portland, OR: Productivity Press.

Morecroft, J. D. W., Larsen, E. R., Lomi, A., & Ginsberg, A. (1995). The Dynamics of Resource Sharing: a Metaphorical Model. *System Dynamics Review, 11*(4), 289-309.

Morrow, R. A., & Brown, D. D. (1994). *Critical theory and methodology*. London: Sage.

Moxnes, E. (1998a). Not only the tragedy of the commons: Misperceptions of bioeconomics. *Management Science, 44*(9), 1234-1248.

Moxnes, E. (1998b). Overexploitation of renewable resources: The role of misperceptions. *Journal of Economic Behavior & Organization, 37*(1), 107-127.

Moxnes, E. (2000). Not only the tragedy of the commons: misperceptions of feedback and policies for sustainable development. *System Dynamics Review, 16*(4), 325-348.

Moxnes, E. (2004). Misperceptions of Basic Dynamics: the Case of Renewable Resource Management. *System Dynamics Review, 20*(2), 139-160.

Naumann, J. D., & Jenkins, M. A. (1982). Prototyping: The New Paradigm for Systems Development. *MIS Quarterly, 6*(3), 29-44.

Neisser, U. (1976). *Cognition and reality : principles and implications of cognitive psychology*. San Francisco: W. H. Freeman.

Nielsen, J. (1993). *Usability engineering*. Boston, MA: Academic Press.

Nielsen, J. (1994). Heuristic Evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods* (pp. 25-62). New York, USA: John Wiley & Sons, Inc.

Nocera, J. A., & Sharp, H. (2007). An Approach to the Evaluation of Usefulness as a Social Construct Using Technological Frames. *International Journal of Human-Computer Interaction, 22*(1 & 2), 153-172.

Oreskes, N., Schrader-Frechette, K., & Belitz, K. (1994). Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science, 263*, 641-646.

Paich, M., & Sterman, J. D. (1993). Boom, Bust, and Failures to Learn in Experimental Markets. *Management Science, 39*(12), 1439-1458.

Pea, R. D. (1993). Practices of Distributed Intelligence and Designs for Education. In G. Salomon (Ed.), *Distributed Cognitions: Psychological and Educational Considerations* (pp. 47-87). Cambridge, UK: Cambridge University Press.

Pfeiffer, D. G., & Lossius, T. (2001). *Symphony of Eigenvalues: A Prototype for Sonification in System Dynamics.* Paper presented at the 19th International Conference of the System Dynamics Society, Palermo, Italy.

Piattelli, M. L., Cueno, M., A., Bianchi, N., P., & Soncin, G. (2002). The Control of Goods Transportation Growth by Modal Share Re-planning: the Role of a Carbon Tax. *System Dynamics Review, 18*(1), 47-71.

Reigeluth, C. M., & Stein, F. (1983). The elaboration theory of instruction. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status* (pp. 335-382). Hillsdale, NJ: Erlbaum.

Reynolds, R. E., Sinatra, G. M., & Jetton, T. L. (1996). Views of Knowledge Acquisition and Representation: A Continuum from Experience Centered to Mind Centered. *Educational Psychologist, 31*(2), 93-104.

Richardson, G. P. (1986). Problems with Causal Loop Diagrams. *System Dynamics Review, 2*(2), 158-170.

Richardson, G. P. (1991). *Feedback Thought in Social Science and Systems Theory*. Waltham, Massachusetts: Pegasus Communications.

Richardson, G. P. (1997). Problems with Causal Loop Diagrams Revisited. *System Dynamics Review, 13*(3), 247-252.

Richardson, G. P., Andersen, D., Maxwell, T. A., & Stewart, T. R. (1994). *Foundations of Mental Model Research.* Paper presented at the 1994 International System Dynamics Conference, Stirling, Scotland.

Richardson, G. P., & Pugh, A. L. (1981). *Introduction to system dynamics modeling*. Portland, Or.: Productivity Press.

Ritzer, G. (1996). *Sociological theory* (4th ed.). New York: The McGraw-Hill Companies.

Ruud, M. (2000). *The Application of a Hydro Electric Production Simulation Model.* Paper presented at the 18th International Conference of the System Dynamics Society, Bergen, Norway.

Sadoski, M., Paivio, A., & Goetz, E. T. (1991). Commentary: A Critique of Schema Theory in Reading and a Dual Coding Alternative. *Reading Research Quarterly, 26*(4), 463-484.

Saleh, M. (2002). *The Characterization of Model Behavior and Its Causal Foundation.* Unpublished Ph.D., University of Bergen, Bergen.

Salen, K., & Zimmerman, E. (2004). *Rules of play : game design fundamentals*. Cambridge, Mass.: MIT Press.

Salomon, G., & Perkins, D. N. (1998). Individual and Social Aspects of Learning. *Review of Research in Education, 23*, 1-24.

Seel, N. M. (2003). Model-Centered Learning and Instruction. *Technology, Instruction, Cognition and Learning, 1*, 59-85.

Senge, P. M. (1990). *The fifth discipline : the art and practice of the learning organization*. New York: Doubleday.

Sfard, A. (1998). On Two Metaphors of Learning and the Dangers of Choosing Just One. *Educational Researcher, 27*(2), 4-13.

Sherwood, S. L. (1966). *The nature of psychology : a selection of papers, essays and other writings by the late Kenneth J.W. Craik*. Cambridge ,.

Shuell, T. J. (1986). Cognitive Conceptions of Learning. *Review of Educational Research, 56*(4), 411-436.

Skartveit, H.-L. (2007, forthcoming). *Representing the Real through Play and Interaction - Changing forms of Nonfiction.* Unpublished PhD thesis, University of Bergen, Bergen.

Skartveit, H.-L., Goodnow, K., & Viste, M. (2002). *'SimHeritage': Multimedia and System Dynamics in Learning Tools for Heritage Management.* Paper presented at the United Nations Educational, Scientific and Cultural Organization, World Heritage Center 30th Anniversary Virtual Congress: World Heritage in the Digital Age, Management of Heritage Cities: Planning for Mixed Use & Social Equity, Mexico City, Mexico.

Skartveit, H.-L., Goodnow, K., & Viste, M. (2003). *Visualized System Dynamics as Information and Planning Tools.* Paper presented at the Insite Conference 2003, Finland.

Skartveit, H.-L., & Viste, M. (2003). *System Dynamics as a Story Engine for Interactive Video.* Paper presented at the Digital Arts and Culture Conference, Melbourne Australia.

Southgate, D., & Lach, L. (1995). *Air Pollution*. Retrieved August 9, 2006

Spector, J. M. (2000). System Dynamics and Interactive Learning Environments: Lessons Learned and Implications for the Future. *Simulation & Gaming, 31*(3), 457-464.

Spector, J. M., & Davidsen, P. I. (1997a). *Constructing effective interactive learning environments using system dynamics methods and tools : interim report*. Bergen: Educational Information Science and Technology Research Program.

Spector, J. M., & Davidsen, P. I. (1997b). Creating engaging courseware using system dynamics. *Computers in Human Behavior, 13*(2), 127-146.

Spector, J. M., & Davidsen, P. I. (1998). Constructing Learning Environments Using System Dynamics. *Journal of Courseware Engineering, 1*, 5-11.

Spector, J. M., & Davidsen, P. I. (1999, April 19-23, 1999). *Transparency and Interaction in System Dynamics Based Learning Environments.* Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Spence, R. (2001). *Information visualization*. Harlow: Addison-Wesley.

Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. (1988). *Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains.* Paper presented at the 10th Annual Conference of the Cognitive Science Society, Mahwah, NJ.

Sridharan, S., & Hunt, K. (2000). *Long-Term Individual and Community-Level Effects of Incarceration of Drug Offenders: A System Dynamics Approach.* Paper presented at the 18th International Conference of the System Dynamics Society, Bergen, Norway.

Stahl, G. (2002). *Meaning and Interpretation in Collaboration.* Paper presented at the CSCL 2002, Bergen, Norway.

Sterman, J. D. (1994). Learning in and about Complex Systems. *System Dynamics Review, 10*(2-3), 291-330.

Sterman, J. D. (2000). *Business dynamics : systems thinking and modeling for a complex world*. Boston: Irwin McGraw-Hill.

Sterman, J. D., & Sweeney, L. B. (2002). Cloudy Skies: Assessing Public Understanding of Global Warming. *System Dynamics Review, 18*(2), 207-241.

Sudhir, V., Srinivasan, G., & Muraleedharan, V. R. (1997). Planning for Sustainable Solid Waste Management in Urban India. *System Dynamics Review, 13*(3), 223-246.

Sweeney, L. B., & Sterman, J. D. (2000). Bathtub Dynamics: Initial Results of a Systems Thinking Inventory. *System Dynamics Review, 16*(4), 249-287.

Tufte, E. R. (1997). *Visual explanations : images and quantities, evidence and narrative*. Cheshire, Conn.: Graphics Press.

Vennix, J. A. (1990). *Mental Models and Computer Models: Design and Evaluation of a Computer-Based Learning Environment for Policy-making.* University of Nijmegen, the Netherlands.

Viste, M., & Skartveit, H.-L. (2003, June 14 - 18, 2003). *Visualization to Support Grounding.* Paper presented at the CSCL 2003, Bergen, Norway.

Viste, M., & Skartveit, H.-L. (2004). Visualization of Complex Systems - The Two-Shower Model. *Psychology, 1*(5).

Weller, H. G., Repman, J., Lan, W., & Rooze, G. (1995). Improving the Effectiveness of Learning through Hypermedia-Based Instruction: The Importance of Learner Characteristics. *Computers in Human Behavior, 11*, 451-465.
White, M., & Caldwell, T. (1998). *Exemplars: a Practical, Extensible Framework for Dynamic Text Generation.* Paper presented at the 9th International Workshop on Natural Language Generation, Niagara-on-the-Lake, Canada.
Wichansky, A. M. (2000). Usability Testing in 2000 and Beyond. *Ergonomics, 43*(7), 998-1006.
Wilson, S., Galliers, J., & Fone, J. (2007). Cognitive Artifacts in Support of Medical Shift Handover: An In Use, In Situ Evaluation. *International Journal of Human-Computer Interaction, 22*(1 & 2), 59-80.
Zapp, A. (2002). Computerisation and Film Language Convergence. In M. Rieser & A. Zapp (Eds.), *New Screen Media: Cinema/Art/Narrative*. London: BFI (British Film Institute) Publishing.

# Appendix A

## Information Presented Before Running the Prototype

The following information was presented (orally in Norwegian) to the participants of Test 2 at the beginning of each session (the text was used as a guide to the introduction):

You will now participate in a test of an interactive learning environment. The system consists of a simulation prototype with a multimedia interface.

The purpose of the study is to test some design principles for visualization of simulation models. This is a prototype, not a fully functioning system, so you may run into some problems on the way.

One of the goals is that you should collaborate on controlling the system an in trying to understand the underlying model. It would be nice if you discuss what you do, problems you run into, and how you understand the model. This will make it easier for me to understand potential problems with the system and the user interface.

You will not be able to ask questions as the simulation is running, but you are welcome to ask questions before and after.

Video recordings will be made while you are working with the prototype and I will take some notes.

Before we start I will ask you to fill out a questionnaire with some background information.

After the simulation you will get another questionnaire where we ask you to provide feedback about the user interface and explain how you have understood the underlying simulation model.

# Pre-Questionnaire

Dear participant. This form should be filled out before you participate in the test and will be used to map the background of the participants. We ask you to answer as accurate as possible.

Thank you in advance!

Participant number:

**Gender:**

Female ☐ Male ☐

**Age:**

**What is your highest level of education?**

| | | | | | |
|---|---|---|---|---|---|
| 9 years of schooling | High school | Bachelor | Masters degree | PhD | Other |

**Which educational program are/have you attending/attended?**

**At what level is your most advanced, completed course in mathematics?**

| | | | | | |
|---|---|---|---|---|---|
| $9^{th}$ grade or less | $10^{th}$ grade | $11^{th}$ grade | $12^{th}$ grade | University level | Other |

**How often do you use computer programs other than a text editor or a web browser?**

| | | | | | |
|---|---|---|---|---|---|
| Never | 1 to 3 days a month | Once a week | Many days a week but not every day | Daily | Other |

**What is your experience with system dynamics?**

| | | | | | |
|---|---|---|---|---|---|
| Have never heard of it | Have heard of it but does not know what it is about | Know what it is about | Have taken a course in system dynamics | Have relatively long experience with system dynamics | Other |

# Post-Questionnaire

You are to answer these questions when the simulation is finished. You are welcome to ask questions!

Participant number:

Part 1

In this part we would like to find out a little about what you think of *the Two-Shower prototype.*

**1. What do you think of *the Two-Shower prototype*?**

**2. I think the underlying principles of *the Two-Shower prototype* were:**

| | | | | |
|---|---|---|---|---|
| Difficult | A little difficult | Neither difficult nor easy | A little easy | Easy |

to understand.

3. *The Two-Shower prototype* was divided into three views: one *introduction view* with graphics of a hotel and a textual introductory story, a *shower room view* that represented the shower, and a *pipe system view* where the pipes and the hot water tank were represented.

**3 a. The Introduction view was:**

| | | | | |
|---|---|---|---|---|
| Difficult | A little difficult | Neither difficult nor easy | A little easy | Easy |

to understand.

**3 b. Comments about the *Introduction view*:**

**3 c. The *Shower Room view* was:**

| | | | | |
|---|---|---|---|---|
| Difficult | A little difficult | Neither difficult nor easy | A little easy | Easy |

to understand.

**3 d. Comments about the *Shower Room view*:**

**3 e. The Pipe System view was:**

| | | | | |
|---|---|---|---|---|
| Difficult | A little difficult | Neither difficult nor easy | A little easy | Easy |

to understand.

**3 f. Comments about *the Pipe System view*:**

**4 a. I think the graphics made the model:**

| | | | | |
|---|---|---|---|---|
| More difficult | A little more difficult | Neither more difficult nor easier | A little easier | Easier |

to understand.

**4 b. Comments about the graphics:**

**5 a. I think the photo slideshows made the model:**

| | | | | |
|---|---|---|---|---|
| More difficult | A little more difficult | Neither more difficult nor easier | A little easier | Easier |

to understand.

**5 b. Comments about the photo slideshows:**

**6 a. I think the animations (such as for example the change in the colors of the pipes, the change in the color of the background, and the flow of water from the hot water tank) made the model:**

| | | | | |
|---|---|---|---|---|
| More difficult | A little more difficult | Neither more difficult nor easier | A little easier | Easier |

to understand.

**6 b. Comments about the animations:**

**7 a. In the upper right hand corner there were some graphs that showed the behavior of the model over time. I think the graphs made the model:**

| | | | | |
|---|---|---|---|---|
| More difficult | A little more difficult | Neither more difficult nor easier | A little easier | Easier |

to understand.

**7 b. Comments about the graphs:**

**8 a. In the model you could press a button to receive a textual explanation of the system. Did you use this opportunity?**

Yes No

**8 b. If yes: I think the text made the model:**

| | | | | |
|---|---|---|---|---|
| More difficult | A little more difficult | Neither more difficult nor easier | A little easier | Easier |

to understand

**8 c. Comments about the textual description:**

**9 a. You could pause the model and by moving the mouse pointer over the timeline you could receive an explanation of what had happened earlier during the simulation. Did you use this pause button?**

Yes No

**9 b. If yes: I think this opportunity made the model:**

| | | | | |
|---|---|---|---|---|
| More difficult | A little more difficult | Neither more difficult nor easier | A little easier | Easier |

to understand.

**9 c. Comments about the explanations one received by pressing the *pause* button:**

**10 a. I think the organization of the user interface (with the gradual introduction of complexity) made the underlying model:**

| | | | | |
|---|---|---|---|---|
| More difficult | A little more difficult | Neither more difficult nor easier | A little easier | Easier |

to understand.

**10 b. Comments about the organization of the user interface:**

**11 a. I think navigation in the system was:**

| | | | | |
|---|---|---|---|---|
| Difficult | A little difficult | Neither difficult nor easy | A little easy | Easy |

to understand.

**11 b. Comments about navigation:**

**12 a. I think collaboration with another person made the underlying simulation model:**

| | | | | |
|---|---|---|---|---|
| More difficult | A little more difficult | Neither more difficult nor easier | A little easier | Easier |

to understand.

**12 b. Comments about collaboration:**

Part 2

We will now ask you some questions about the model (you are welcome to look at the prototype while you answer these questions).

**14. Can you explain why the temperature oscillated?**

**15. If you did not change the tap setting for a while you could still experience changes in the water temperature. Can you explain why?**

Thank you for your help!!