Simulation in University Chemistry Education
Cognitive and affective aspects

Mikael Winberg

Department of Chemistry, Chemistry Education Research
Umeå University
Umeå 2006
# Table of contents

1 List of papers ................................................................. - 1 -  
2 Preface ............................................................................... - 2 -  
3 Introduction ........................................................................ - 5 -  
  3.1 Outline of the main areas of the study ................................. - 6 -  
  3.2 Brief definitions and outline of variables .......................... - 6 -  
4 Models ............................................................................... - 10 -  
  4.1 Attitudes Toward Learning ............................................... - 10 -  
  4.2 Cognitive Load ................................................................. - 12 -  
  4.3 Simulations ....................................................................... - 15 -  
    4.3.1 Learning with conceptual simulations ......................... - 15 -  
  4.4 Flow ............................................................................... - 18 -  
  4.5 Engagement modes .......................................................... - 19 -  
5 Summary of Papers ............................................................ - 20 -  
  5.1 Paper I ............................................................................ - 20 -  
  5.2 Paper II ............................................................................ - 22 -  
  5.3 Paper III .......................................................................... - 23 -  
  5.4 Paper IV .......................................................................... - 24 -  
6 Methods ............................................................................... - 25 -  
  6.1 Interviews ......................................................................... - 28 -  
    6.1.1 Purpose ...................................................................... - 28 -  
    6.1.2 Selection for interviews ............................................... - 29 -  
    6.1.3 Implementation .......................................................... - 29 -  
    6.1.4 Documentation and analysis ........................................ - 31 -  
    6.1.5 Reporting data from interviews ................................... - 32 -  
  6.2 Questionnaires .................................................................. - 34 -  
    6.2.1 Attitudes Toward Learning ........................................... - 35 -  
    6.2.2 Flow and perceived learning outcome ......................... - 36 -  
    6.2.3 Engagement modes ..................................................... - 37 -  
  6.3 Discourse analysis ............................................................. - 37 -  
7 Main conclusions and discussion ........................................... - 39 -  
References ............................................................................. - 44 -  
Appendices ............................................................................. - 49 -  
Appendix 1 ............................................................................ - 49 -  
Appendix 2 ............................................................................ - 53 -
1 List of papers

This thesis is based on the following papers, from here on referred to by their roman numbers.


IV. **Winberg T.M., Hedman L.** (2006) Student attitudes toward learning, level of pre-knowledge and instruction type in a computer simulation: effects on flow experiences and perceived learning outcome. *Instructional Science (In Progress)*
2 Preface

The central theme in this thesis is a computer simulation based on acid-base titrations and buffers, which is used in chemistry courses at Umeå University. This might seem somewhat narrowly focussed, since there are many other simulations in existence. Thus, to enable generalization, it could be argued that it would have been better to investigate more than one simulation.

However, although each separate paper listed above focuses on a few distinct aspects of the exercise, the intention of the present work is to present a multiple perspective on what influenced the use and effects of this particular simulation. The rationale behind this is illustrated by the western version of the Buddhist canon, Udana 68-69, by the poet John Godfrey Saxe (1816-1887):

It was six men of Indostan
   To learning much inclined,
Who went to see the Elephant
   (Though all of them were blind),
That each by observation
   Might satisfy his mind.

The First approached the Elephant,
   And happening to fall
Against his broad and sturdy side,
   At once began to bawl:
   “God bless me! but the Elephant
   Is very like a wall!”

The Second, feeling of the tusk,
   Cried, “Ho! what have we here
So very round and smooth and sharp?
   To me ’tis mighty clear
This wonder of an Elephant
   Is very like a spear!”

The Third approached the animal,
   And happening to take
The squirming trunk within his hands,
   Thus boldly up and spake:
   “I see,” quoth he, “the Elephant
   Is very like a tree!”

The Fourth reached out an eager hand,
   And felt about the knee.
“What most this wondrous beast is like
   Is mighty plain,” quoth he;
   “ ’Tis clear enough the Elephant
   Is very like a tree!”
The Fifth, who chanced to touch the ear,
   Said: “E’en the blindest man
Can tell what this resembles most;
   Deny the fact who can
This marvel of an Elephant
   Is very like a fan!”

The Sixth no sooner had begun
   About the beast to grope,
Than, seizing on the swinging tail
   That fell within his scope,
“ I see,” quoth he, “the Elephant
   Is very like a rope!”

And so these men of Indostan
   Disputed loud and long,
Each in his own opinion
   Exceeding stiff and strong,
Though each was partly in the right,
   And all were in the wrong!

So oft in theologic wars,
   The disputants, I ween,
Rail on in utter ignorance
   Of what each other mean,
And prate about an Elephant
   Not one of them has seen

Thus, the advantage of this multiple perspective is that it provides an opportunity to get a glimpse of what this particular elephant, i.e., the learning within this simulation exercise, might look like.

To do this, I have studied several variables which all make up different aspects of the learning, namely context, process, and outcome. However, although the variables examined in this thesis might represent a good starting point, much is left to do. Not only are there several other variables and interactions that need to be studied, but also the limited number of student participants in the papers also calls for additional validating studies in other contexts and with larger numbers of participants.

Much effort has been invested in trying to find appropriate methods to assess the different aspects of learning outcomes and performance. The instruments that I have used to describe these aspects, for example, the quality of reasoning, often comprised a qualitative interpretation followed by quantitative analysis of the interpretations. From some of the reactions that I have received to this approach, I conclude that the parable of the blind men and the elephant is still
applicable. It seems that either a quantitative or qualitative approach will do, aligned to the established norms of the respective traditions; but that a combination of the two meets criticism from both “camps”.

This has been very encouraging since it gives me the impression that the methods so far developed contribute something substantial to research in this area of investigation. Although it would be presumptuous to believe that this thesis will revolutionize educational research in this area, I hope that it will help to close the gap between “qualitative” and “quantitative” research that is apparent today.

To be able to make claims of pragmatic validity (Kvale, 1997), educational research benefits from being situated in an authentic educational context. However, as the conscientious reader will notice, authentic settings sometimes cause problems with experimental design. For example, the “subjects” (students) do not always show up in the place or time they are expected; access to “subjects” may be restricted due to heavy workload; unforeseen technical problems may make students be reluctant to participate in the study or call for last-minute rearrangements of the experimental design. This is clearly a disadvantage of “contextualized” research in contrast to clinical, de-contextualized research, where variables can be controlled to a greater extent.

The development of the instruments for assessing the different aspects of the simulation exercise has sometimes been a tedious process and does not lack examples of dead ends. Latent Semantic Analysis of log-files to characterize different types of user behaviour provides one example of these. Nevertheless, I have experienced that persistence is often rewarded. The intense sense of Flow when solving a problem or discovering something interesting is what keeps research going (besides money...).

Although this thesis represents my efforts to learn the craft of educational research rather than a “once in a lifetime peak research performance” (there is hopefully more to come), I hope that the interested teacher as well as the researcher finds something in this thesis worthwhile to contemplate or use.
3 Introduction

Like most modern educational researchers, I agree with the general ideas of constructivism in the sense that learning is an active construction of knowledge rather than a passive receiving of information, and that instruction has to support students’ knowledge construction rather than simply presenting information (P. A. Kirschner, Martens, & Strijbos, 2004).

This thesis undertakes two main tasks: first, to explore, within an authentic educational context, some of the variables that influence the quality and outcomes of a knowledge-constructing activity during a simulation exercise; and, second, to find appropriate instruments and methods to measure these variables, processes and outcomes.

Although the purpose of the simulation was to help students improve their knowledge in the subject area of acid-base reactions, perspectives from other disciplines suggested that there were other interesting outcomes from the simulation, in addition to the production of pure declarative knowledge. For example, research on working memory and cognitive load were central to understanding the changed behavior of students during laboratory work and their improved ability to apply chemical knowledge when interviewed. Educational psychology, in particular research on cooperative learning and cognitive conflict, and discourse-analysis within the social sciences, influenced the work in papers II and III, while research on Flow and attitudes were salient for papers IV and I, III-IV, respectively.

It is my view that chemical education research is essentially about providing a bridge between fundamental research (often clinical studies) and educational practice. Hence, the different theoretical models have not usually been exploited in every detail but have been used at the level of complexity necessary for understanding the variable, process or outcome of interest.

In the section below, an outline of the main areas that have been studied in this thesis is presented, followed by brief definitions of the variables that have been examined and an outline of how these relate to the different phases of the simulation exercises.
3.1 Outline of the main areas of the study

There are three main areas that have been targeted in this thesis:

1) the learning environment, comprising both what the student brings into the situation, e.g., attitudes toward learning and prior knowledge, and factors that relate to how the situation is organized, e.g., the format of instruction, media used for communication, and group configurations

2) the learning process, involving interaction between students, cognitive focus, quality of reasoning, and motivation

3) the learning outcomes in terms of cognitive and affective effects; e.g., usability, or accessibility, of knowledge, and perceived conceptual change.

3.2 Brief definitions and outline of variables

Students’ attitudes toward learning (Perry1998), more thoroughly discussed in section 4.1, concern beliefs about the nature of knowledge and how it is to be acquired. In particular, the role of the student, peers, teacher, and assessment in learning are in focus.

Group configurations refer to how dyads were composed with respect to the participants’ attitudes toward learning.

Prior knowledge was either measured or inferred from discussions or final examination results and referred only to declarative knowledge.

Engagement modes (EM) (Hedman, 2001) describe different ways of interacting with information technology. In this thesis, engagement modes were used to investigate differences in students’ predispositions to engaging in the use of information technology for learning.

In paper III, two different media for communication were compared with respect to their efficiency in promoting learning, namely face-to-face and web-based, written, threaded discussions.

Two different instruction types, comprising open- and closed- ended tasks were compared with respect to the type of affective and cognitive responses that they generated.

Motivation refers to the flow variables that describe the degree of concentration, challenge and enjoyment students felt during the simulation exercise. The learners’ perceptions with respect to the extent of in-depth
discussions, as well as their reported level of engagement and interest were also interpreted as indicators of motivation.

*Locus of control*, does not have a single meaning; it is used to describe whether the student knew what to do (procedural and manipulative issues) as well as whether the student was confident about their own reasoning being correct (cognitive issues), and the extent of stress associated with not understanding the chemistry.

*Discussion quality*, e.g. degree of co-constructive activity and level of chemical knowledge expressed in the discussions, was measured by methods of discourse analysis, described in depth in section 6.3, and to some extent in 6.1.5.

Students’ *Cognitive focus* during laboratory work was assessed in terms of what questions students posed to the assistant teacher. Our primary interest was to investigate whether the questions (and the students) were focused on the theoretical or the practical issues associated with the laboratory exercise, and whether the questions were spontaneous or thoughtful.

*Knowledge accessibility*, or usability, was judged by analysing the way (accuracy and complexity) students used their chemical knowledge in situations that required relatively fast responses (interviews). This term, as used in this thesis, is equivalent to *intuitive knowledge*, i.e., the ability to apply knowledge in new and complex situations. Somewhat confusingly, both are used in paper I.

*Perceived learning* is the extent to which students felt that they had gained new knowledge or discovered and rectified deficiencies in their previous knowledge while performing the simulation exercise.

Figure 1, below, gives an overview of how these variables relate to the simulation exercise, distinguishing between factors that are assumed to be formed (e.g. attitudes toward learning) or designed (e.g. group configuration) prior to the learning activity, those that occur during the learning activity, and those that are regarded as the “products” of learning.
Figure 1. Overview of the variables that were examined. Experiments that were conducted, but have not been reported yet have been excluded.

The distinction implied by figure 1 is, however, not all that easy to make. Most of the process variables, e.g., locus of control, discussion quality or motivation, could be regarded as outcomes of the media used for communication or the student's attitudes toward learning. Furthermore, the learning environment does not only include the predefined factors but in addition the discussions that are held and the students' affective reactions. Hence, Figure 1 is not a validated conceptual model of the relationships between variables. As mentioned in the preface, the learning environment also includes many other variables that are not mentioned in this thesis.
An overview of how the variables occur in the papers is presented in Table 1, which also illustrates how the chase of the elephant is manifested by the use of multiple dependent and independent variables in each study.

Table 1. Overview of the variables measured or discussed in the papers. Variables that were only discussed are indicated by an asterisk

<table>
<thead>
<tr>
<th>Variable</th>
<th>Paper</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudes to learning</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Group configuration</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instruction format</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-knowledge</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Discussion media</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engagement modes</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived value</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive load*</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cooperative activity</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive focus</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motivation 1</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locus of control</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conceptual change 2</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemistry reasoning</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

1 Includes indicators of Engagement, and the flow components: Concentration, Challenge and Enjoyment.
2 Perceived conceptual change

In the following section I describe a selection of the theoretical models that have been used either to select a relevant set of dependent and independent variables, to explain observations, or provide arguments for the validity of instruments and results. More information about the models can, in most cases, be found in the respective papers (see Table 1).
4 Models

4.1 Attitudes Toward Learning

As seen in Table 1, students’ attitudes towards learning chemistry have been used in the majority of the papers contained in this thesis. There are good reasons for this: it has been claimed that students’ beliefs about knowledge and learning influence motivation (Windschitl, 1998), strategy selection and socio-cognitive engagement when working with peers (Hogan, Nastasi, & Pressley, 1999), and ultimately, the learning outcome.

In paper I, there are indications that students’ attitudes toward learning also influence the way that they perceive a task. This not only has implications for what and how they learn, but also for the cognitive load the students perceive during problem solving, for example. This proposal is discussed under section 4.2. For a comprehensive review of how students’ beliefs might affect learning see Hofer (1997). For the reasons mentioned above, student attitudes towards knowledge have been used frequently as an independent variable in my papers. Since there is sometimes a confusion in the literature about how to define ‘attitudes’, I present below some influential views and indicate how I have defined the term in this thesis.

Secord and Backman (1969, as cited in Arnold, Cooper, & Robertson, 1998) defined attitudes as: “certain regularities of an individual’s feelings, thoughts and predispositions to act toward some aspect of his environment”. Thus, they proposed three attitudinal components; affective, cognitive, and behavioural, directed toward a specific object. The requirement of a specific object is an important attribute of attitude, distinguishing it from personality, which reflects a person’s predispositions across a range of situations (Arnold et al., 1998).

In this thesis, the object the attitudes refer to is “learning chemistry”. Thus, the attitudes that have been expressed by students in the papers are valid for this object and only in this particular higher education context.

In my papers, as well as in contemporary attitude research, behaviour is interpreted as an outcome of an attitude rather than a part of it (Arnold et al.1998). Although being positive to learning chemistry in general might influence behaviour, we assumed that the cognitive aspects, i.e., student beliefs about the process of learning chemistry, would be more clearly linked to their actual learning behaviour and performance in the present learning contexts. Hence, we chose to assess the cognitive attitude component, illustrated in Table 2, which describes the features of Perrys’ (1970) scheme of intellectual development.
Table 2. Overview of student perceptions of educational characteristics implied by the Perry scheme. Adapted from Finster (1991), Moore (1994), and Cornfeld and Knefelkamp (as cited in Peggy Fitch & Richard S Culver, 1984).

<table>
<thead>
<tr>
<th>Nature of knowledge</th>
<th>Dualism Position 1 &amp; 2</th>
<th>Early multiplicity Position 3</th>
<th>Late multiplicity Position 4</th>
<th>Contextual relativism Position 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role of the instructor</td>
<td>Source of knowledge. Role is to transmit knowledge</td>
<td>Guide to the truth. Role is to teach how to find truth</td>
<td>Only one of many authorities. Modelling the way to think and use evidence.</td>
<td>Source of expertise but critically evaluated since mutuality of learning is sought</td>
</tr>
<tr>
<td>Role of the student</td>
<td>Role is to receive information and to demonstrate having that information in tests</td>
<td>Find right ways to find the right answers. To work hard</td>
<td>Hard work not sufficient. Think independently to make sense. Challenge instructor as equal</td>
<td>Evaluate information in relation to context. To use the intellect. Legitimate source of knowledge</td>
</tr>
<tr>
<td>Role of peers</td>
<td>Not legitimate sources of knowledge</td>
<td>Not authorities but can help in finding the truth</td>
<td>Sources of diversity. One is just as good as the other.</td>
<td>Legitimate sources of learning and diversity</td>
</tr>
<tr>
<td>Assessment</td>
<td>Evaluation is related to sense of self. Tests should be clear-cut and objective</td>
<td>There must be right answers to the questions. Quantity counts</td>
<td>Anything goes. What counts is what the teacher expects = Independent thought</td>
<td>Evaluation can be separated from self and seen as part of learning</td>
</tr>
<tr>
<td>Primary intellectual task</td>
<td>Learn basic information. separate right from wrong</td>
<td>Compare and contrast the content</td>
<td>Analysis. Use supportive evidence. Learn to think abstractly</td>
<td>Modify and expand concepts. Determine adequacy of arguments in context</td>
</tr>
</tbody>
</table>

In the field of research into epistemological beliefs (Hofer & Pintrich, 1997), which is generally seen as a continuation of the work of Perry, the central themes in Perry’s scheme have been further refined.

Hofer and Pintrich (op. cit.) suggested that the core of epistemological belief consists of two main themes. The first concerns the nature of knowledge which comprises beliefs about whether knowledge is certain or if it changes with time,
and whether knowledge consists of discrete facts or interrelated concepts that are context-dependent (simplicity of knowledge). The second theme deals with the process of knowing. Here the two central questions are:

a) Does knowledge reside in authorities or is it constructed by the learner?

b) Can and should knowledge be justified.

If knowledge is certain and simple, there is no reason to justify it, but if knowledge is constructed and context-dependent, it should be critically evaluated. These core epistemological beliefs can also be identified in the descriptions of students’ perceptions of educational characteristics describe for students in the different positions in Perry’s scheme (Table 2). Hence, ‘attitude toward learning’, as used in this thesis, is strongly related to epistemological beliefs. This strong link is acknowledged in particular in paper IV, which is partly based on findings from the field of research on epistemological beliefs.

4.2 Cognitive Load

In paper I, students’ cognitive focus, in terms of the aspects of a laboratory exercise they chose to direct their attention towards, and the degree of reflection on these aspects, was examined during a practical laboratory exercise. It has been shown that students tend to focus on manipulative details and other procedural issues rather than considering the underlying theory and linking it to the exercise (Hofstein & Lunetta, 2004; Johnstone & Wham, 1982). Johnstone (op. cit.) argued that this represents natural behaviour when the input of information exceeds the students’ information processing capacity, i.e., cognitive overload. Cognitive overload is followed by a response from the learner, which aims to reduce the load to a manageable level. Since completing the practical part of the exercise is not optional, this is usually what the learners focus on, while theoretical considerations that are not crucial for the manipulative performance in the laboratory are saved until after the exercise (Lundberg, 1998). Fortunately, however, the learners’ information-processing capacity can be influenced by instructional design. This provides an opportunity for teachers to help students to allocate memory resources to higher level thinking during problem solving or laboratory work. How this is possible is illustrated in Figure 2, which shows a simple model of human information processing.
According to this model, working memory has a central position in an individual’s perception of the world as well as in storing and recalling information from long term memory. Conscious thinking is assumed to occur in the working memory, where new information is processed, compared with the information in long term memory and eventually combined with previous knowledge, i.e., expanding what we know. While long term memory can contain a virtually unlimited amount of information, the working memory capacity is limited. On average, working memory can handle about seven items of information simultaneously, and even less if there are interactions between the items that need to be processed at the same time as they are kept in mind (John Sweller, van Merrienboer, & Paas, 1998). Several mechanisms have evolved to compensate for this limitation in processing capacity.

a) The human brain does not perceive all information that is “emitted” by the environment. In the light of our previous experiences, knowledge, and interests, the perception filter separates “noise” from the information that we need or want to know. Thus, perception helps us to allocate the limited working memory resources only to information that we consider important. (Johnstone, 1997)

b) New information can be stored in at least two distinctly different ways in the long term memory: either connected to previous knowledge or as separate pieces of information. By connecting related items of information to each other, we construct schemata, which in turn can be activated by different situations or
Schemata facilitate retrieval from long term memory as activation of one part of a schema triggers activation of the other items that are linked to it. Schemata also reduce working memory load, since one schema (regardless its complexity) is treated as one information item in the working memory. Thus, information that is learned as a separate item, with very few or no connections to previous knowledge, is difficult to retrieve (remember) and occupies a relatively large proportion of working memory capacity when used (Sweller et al. 1998).

c) Extensive training within a subject might lead to the automation of schemata, i.e., they are processed unconsciously, further reducing cognitive load (Paas & Van Merriënboer 1994). A simple example of an automated schema could be the “experienced” chemist applying the rule $\text{pH} = -\log [\text{H}_3\text{O}^+]$ to find the pH in a solution where $[\text{H}_3\text{O}^+] = 10^{-2.5}$ M. This would probably require very limited working memory resources, compared to the novice trying to do the same conversion. The novice, having little experience of making this conversion or applying logarithmic rules, would have to allocate significant working memory resources to determine which rule is appropriate and to perform the calculation.

d) In addition to the mechanisms mentioned above, the working memory is assumed to function as a coordinating processor for two sub-systems, namely, the visuo-spatial scratchpad and the phonological loop, allowing for simultaneous processing of visual and acoustic information (Baddeley, 1986).

There is a debate in the field of instructional design as to whether multi-modal representations in simulations, by parallel use of both sub-systems, decrease, or, for example, by split-attention or redundancy effects, increase, cognitive load. Although interesting, and a promising area for future research, this particular aspect of the cognitive architecture and its implications for instructional design has not been the focus of this thesis.

More detailed information on the different sources of cognitive load and their implications for teaching can be found in the introduction of paper I. For the interested reader, I also recommend the very comprehensive papers by Sweller et al. (1998) and Paas et al. (2003).

Hence, what we know and how we know it affects not only what we perceive, but also the accessibility of knowledge and the cognitive load when we use it. In paper I, I focused on the significance of appropriately constructed prior knowledge (schemas) to help students to identify what is important and to make information processing less cognitively demanding, thus allowing for higher level thinking during laboratory work. I believe that, given the limited
time available for training in university chemistry education, simulations might be a means of providing the intensive interaction with a topic that is necessary in order to construct functional schemata.

4.3 Simulations

The development of expert-like schemata, i.e., with a high degree of complexity and automation (Sweller et al.1998), that facilitate higher level thinking in complex situations requires extensive experience of the domain (J. Sweller, 1988). Simulations have developed to provide intensive, cheap, and safe experiences in handling complex situations. Well known examples are pilot- or air traffic control training, which could be considered as examples of operational simulations in which training is often focused on cognitive or non-cognitive sequences (procedures). Conceptual simulations, in contrast, focus on developing students’ knowledge about the scientific models that underlie the simulation and will be further discussed below.

Conceptual simulations often provide some kind of experimental situation where learners can propose hypotheses, design and perform experiments, and analyse and infer knowledge about the phenomenon from the responses of the system. Thus, a flight simulation could be either conceptual, e.g., if the student is supposed to infer knowledge about the laws of aerodynamics, or operational if the focus is on learning an airport approach procedure, i.e., a predefined sequence of actions.

In this thesis, the purpose of the simulation was not training in manipulative or procedural skills in acid-base titrations but rather to contextualise and refine students’ theoretical knowledge, thus providing an opportunity for students to integrate facts and concepts that might previously have been learned in an isolated fashion (which several students did report in interviews).

4.3.1 Learning with conceptual simulations

There is no consensus on what the primary learning outcomes of conceptual simulations should be and how these can be assessed. As implied in the previous section, learning outcomes and assessment methods are intimately related to the purpose and design of the simulation. Part of the ambiguity might stem from researchers not being clear about the purpose and the features of their simulation and, thus, what outcomes to expect. Thomas and Hooper (1991) suggested a four-category system for classifying simulations based on their
purpose. They also provide a short description of common central features and outcomes of the different types of simulation:

a) Experiencing – aims at familiarizing a novice learner with the domain and bringing about an understanding of the learning goals. An experiencing simulation might provide concrete examples, motivation, an organizing structure, or reveal deficiencies in student knowledge. Learning outcomes seem to appear primarily in application and transfer abilities rather than declarative knowledge.

b) Informing – the primary task is to transmit information. Simulations are not commonly used for this purpose. When they are used, their role is often confused with experiencing simulations. In general, the effects on performance in “knowledge tests” (not defined by the authors) have been negative or absent. In addition, since feedback on student conclusions is not usually given by the program (no interaction), any alternative conceptions are likely to persist.

c) Reinforcing – previous knowledge is applied in the same environment as it was learned. Feedback mechanisms are built-in to hone skills and knowledge. This type of simulation usually provides a high degree of structure i.e., a sequence of predefined exercises with predefined correct responses. Due to the feedback function, these simulations sometimes tend to have a more or less pronounced informing character. Effects have been shown primarily in applications associated with complex situations and problem solving.

d) Integrating – aims at integration of aspects of knowledge that were originally learned separately. Previous knowledge is applied in unique situations, i.e., not the same as those situations in which it was learned, and the system responses are used to infer knowledge about the relationships between concepts. Such programs usually have no built-in feedback functions. Effects are primarily on transfer and applications in complex situations.

Hence, the general conclusion is that simulations are most useful for experiencing and integrating purposes and that the effects are difficult to demonstrate with tests that assess retrieval of facts or procedures. Instead, the main effects of this type of simulation seem to be on increased ability to apply knowledge in unique and complex situations, also known as intuitive
knowledge (Swaak & de Jong, 1996). When intuitive knowledge is used as a measure of learning outcomes, the results are less ambiguous. Simulations with integrating purposes prove superior to expository teaching. Such simulations also prove more advantageous when student work is supported with instruction; compared to when there is no instruction (de Jong & van Joolingen, 1998).

The simulation that has been studied in this thesis aimed at situating students’ prior theoretical knowledge in a new context and making students actively review and elaborate on central concepts. By offering a number of ways for students to infer knowledge about the underlying theoretical models of the system, and requiring them to explain or discuss results from a number of perspectives, the simulation exercise was assumed to lead to integration of any previously independently learned knowledge elements into useful schemata. In all the situations that were studied in this thesis, instruction was provided to enable structuring of students’ work during the simulation exercise. The instructions provided guidance on what experiments to perform as well as what issues might be important to discuss. In paper IV, a “guiding” version of instruction was developed that was identical to previous instructions with respect to what experiments to perform but the tasks contained explicit questions as opposed to the open-ended questions in the “open” instructions.

In most cases, a cooperative setting was used. This forced students to structure their knowledge to be able to articulate it. Cooperative settings have also been said to increase the probability that students become aware of cognitive conflict, i.e., a mismatch between new information and their own conceptions of the subject, and so facilitate the knowledge-constructing activity (C. Chan, Burtis, & Bereiter, 1997; C. K. K. Chan, 2001; Crook, 1995; Limon, 2001). Results in paper I, where a significant improvement in the ability to apply chemical knowledge in a complex situation was found, together with the design-features of the simulation, suggest that the simulation studied in this thesis should be classified as Integrating.

However, the simulation could have fulfilled different purposes, depending on the degree of the students’ pre-knowledge. In papers I and IV, the chemistry students had relatively good prior knowledge while the pharmacy students in papers II and III had significantly less experience of the domain. Thus, for the latter, the simulation might have had an experiencing function by providing a more or less concrete context and helping students to identify deficiencies in their knowledge. For the chemistry students, the simulation would have had primarily integrating, and perhaps also reinforcing, functions. In all cases however, effects on intuitive rather than declarative knowledge were expected.
4.4 Flow

In paper IV, we suggested that students’ flow experiences could be used to evaluate the learning-process during the simulation exercise.

It has been suggested that ‘flow’ is an intensely rewarding experience when a person is deeply focused on performing an activity (Csikszentmihalyi, 1990). Flow is a rare condition, experienced only when a person is performing to the limit of his or her capacity, with clear goals and immediate feedback and while doing something that is in line with intrinsic motives (Csíkszentmihályi, 1999). Nevertheless, concentrating hard on a challenging task would be beneficial for learning, provided that the cognitive load imposed by the challenge is relevant to schema acquisition. Ghani and Deshpande (1994) proposed five fundamental components of flow as related to the use of computers: Pleasure, Concentration, Control, Exploration, and Challenge. Although we did not expect flow in its pure form to occur, we were interested in the extent to which the different instruction formats would “push” students in the direction of flow, i.e., the degree of Challenge, Pleasure, Concentration, Exploration, and sense of Control that was experienced by students during the simulation exercise.

As shown in Figure 3, “Control” is experienced when skill is high while the challenge is below the limit of that which the learner can perform.

Figure 3. The balance between skill and challenge affects the nature of the experience. From Csíkszentmihályi (1999)
Although a positive state, “control” is associated with difficulties in concentration and low intrinsic motivation. “Excitement”, in contrast, occurs when the challenge exceeds the skills of the learner. Although the learner may be mentally focused and active, they are likely to feel vulnerable and not in control of the situation. When flow is within reach, as for the learner in the “excitement” and “control” sectors, it might act as a positive force for learning since it motivates the learner to either improve their skills or to seek challenges in order to reach flow. If the learner is too far from the flow sector, e.g., in “anxiety” or “apathy”, the step to achieving flow is too large and often the solution is to lower the challenge or to seek stimulation from other activities (Csíkszentmihályi, 1999).

For these reasons, we found the flow-components of Ghani and Deshpandes (1994) useful for evaluation of the learning process in paper IV.

4.5 Engagement modes

‘Engagement modes’ (EM) is a construct used by Sharafi, Hedman and Mongomery (2006) to describe how a person interacts with an object, in this case information technology, to search, receive, handle, or distribute information.

According to their EM-model, there are five separate modes of engagement: Ambition/Curiosity, Enjoyment/Acceptance, Efficiency/Productivity, Avoidance/Hesitation and Frustration/Anxiety. Factor analysis has shown that these load into three basic dimensions of interaction: evaluation (whether the object is positive or negative), locus of control (whether the learner controls the object or vice versa), and focus of motivation (whether the goals are inherent in the activity or are external).

Sharafi et al. (2006) showed that the basic dimensions of the EM-model are related to the degree of intrinsic interest, competence, and autonomy perceived by the learner. As these are known to be the cornerstones of motivation (Ryan & Deci, 2000), the EM-model was found to be useful when investigating students’ predisposition to learn by using computers for simulation and communication, prior to the simulation exercise in paper III.

As in a previous study by Hedman and Sharafi (2004), we chose to study a combination of all the positive engagement modes: Enjoyment/Acceptance, Efficiency/Productivity, and Curiosity/Ambition, and all the negative engagement modes: Frustration/Anxiety and Hesitation/Avoidance.
5 Summary of Papers

5.1 Paper I

*Computer-simulated Prelab and students’ Attitudes Toward Knowledge: effects on students’ Cognitive Focus and Knowledge Usability*

In chemistry education, particularly at university level, laboratory work is seen as an opportunity for students to elaborate on theory in a meaningful context as well as to practice the manipulative skills, which are crucial for a chemist to master. However, studies have shown that students’ cognitive focus is mostly directed toward the manipulative aspects, in order to achieve good data for later elaboration (Hofstein & Lunetta, 2004; Lundberg, 1998).

In this study, Cognitive Load Theory (CLT) (Johnstone & Wham, 1982; John Sweller et al., 1998) was used as a framework to explain the mainly practical focus that was observed during a practical laboratory exercise on buffers and acid-base titrations. CLT suggests that there are three different types of cognitive load:

a) **Intrinsic**: depending on the number of “information-chunks” a learner has to keep in working memory and the number of interrelations between these chunks that have to be processed.

b) **Extraneous**: arising from issues that are not central to the task e.g., problems of interpreting a diagram due to its poor design, or difficulties in finding equipment in a poorly organised laboratory.

c) **Germane**: directly related to the revision or construction of schemata, i.e., learning.

Working memory (Baddeley, 1986), where conscious thinking occurs, can only handle a limited number of information elements. Since the three different types of cognitive load are additive, one type of cognitive load increases at the expense of the others. For example, if extraneous cognitive load is very high, there is little capacity left for information processing or learning.

Intrinsic cognitive load can be reduced by automation of problem-solving strategies and procedures or by increasing the size and complexity of schemata since a schema, regardless of its size, is treated as one “information-chunk” in the working memory.
A computer simulated acid–base titration exercise was developed and used as a pre-lab exercise to help students capitalise on their relevant prior knowledge. It was assumed that the simulation would lead to refined and appropriate schemata, and thereby reduce intrinsic and, to some extent, extraneous cognitive load (Sweller et al, op. cit.). This would allow for deeper processing and a focus on theoretical issues.

To assess whether the simulation affected students’ schemata, 28 students were selected for interviews immediately following the laboratory exercise. Since there were indications that the students’ attitudes toward learning (ATL) would influence their approaches to the simulation tasks (Finster, 1989, 1991; Hofer & Pintrich, 1997; Moore, 1994), the selection of students for interview was guided by their ATL which had been determined at the beginning of the course.

The data collected in the interviews suggested a more frequent, complex, and correct use of chemical concepts for the students who had participated in the simulation. Students with a more relativistic ATL generally performed better than those with dualistic ATL, confirming that ATL influences students’ approaches to the simulation exercise and ultimately, the functionality of their schemata. However, students’ prior knowledge seemed to play a decisive role in how they benefited from the simulation. This relationship needs to be explored in more depth.

The students’ cognitive focus during the laboratory exercise (that followed the day after the simulation) was assessed by registering the type of questions that the students posed to the assistant teachers during the laboratory work. Results show that the students’ questions were more theoretical and thoughtful in the case of those who had undertaken the simulation prior to the laboratory exercise.

Since there was an improved ability to use chemical knowledge among the students that had undertaken the simulation, we argue that the shift in focus during the laboratory exercise for those students was due to improved schemata and thus, lowered cognitive load.
5.2 *Paper II*

*Assessing discourse quality in threaded discussion groups: a comparison of methodologies.*

There is a dichotomy between research traditions using qualitative and quantitative research methods for analysis of discourses (Wegerif & Mercer, 1997). In addition, there is little or no consensus on how quality of discourse should be defined. Qualitative methods are purported to be well suited to explaining causal relationships between events in a discourse. However, they have been criticized for relying on more or less arbitrarily selected excerpts from the original material, making the comparison of results from different studies difficult or impossible, and compromising their reliability and validation.

However, quantitative methods, according to Wegerif and Mercer, have been criticized for their inability to capture the process by which learning unfolds and, when using exclusive categories, for failing to capture ambiguities and multiple meanings in statements by learners. They have also been criticized for providing only statistical measures of the outcomes produced by the group activities and summaries of individual perceptions, without analysing in-depth the process required to reach a group outcome (Benbunan-Fich et al. 2003).

In this study, three different methods were compared in order to investigate whether the perceived quality of a discourse depends on the method of analysis, whether qualitative or quantitative. The study also aimed to find the criteria that seem most important for judging quality in the context described below.

Twenty-two students in their first term on a distributed pharmacists program at Umeå University worked in dyads on a simulation about acid–base titrations, while communicating through written threaded discussion groups. Their discourses were collected and the quality was analysed using:

- a) subjective ranking by experienced chemistry teachers
- b) quantitative coding of student interactions with their partners and the content of the simulation, using categories that emerged during the analytical process
- c) quantitative coding of the accuracy and complexity of students’ chemistry reasoning, using predefined SOLO-like categories (Biggs & Collis, 1982).

Results show that both the qualitative and quantitative methods were well correlated in terms of the judged/measured quality of the different discourses. Important features shared by the qualitative and the quantitative methods
(where the categories were generated by a qualitative analytical process) were the degree of co-operative activity and reference to theory. The qualitative method also generated criteria that described persistence with the task, mutuality, and whether cognitive conflicts were realised and resolved or not. However, results from the quantitative method suggested that the sheer volume of discussion is also an important quality indicator, possibly reflecting a general level of motivation in groups. In addition, the quantitative method showed that the occurrence of incorrect reasoning might indicate not only the presence of alternative conceptions, but also that they become apparent and are revised. This was often the case in the group that was ranked by all methods as showing the highest discourse-quality. A feature that was found only in the quantitative method was metacognitive activity, e.g., monitoring of the process or reflection on the accuracy/certainty of the student’s own knowledge.

5.3 Paper III.

A comparative analysis of discussion quality in Face-to-Face and threaded discussion groups.

The extent and nature of educational discussions and interactions have changed since web-based distance learning has become increasingly prevalent. Therefore, there is a need to clarify what effects new technologies may have on educational discussions. In recent years, there has been considerable interest in comparing traditional Face-to-Face groups with computer-supported groups (Benbunan-Fich et al.2003; Ellis et al.2004; Fjermestad & Hiltz, 1998; Newman et al.1997). However, according to a review by Fjermestad & Hiltz (1998) most research concerns comparisons between synchronous computer-mediated and Face-to-Face discussions, while there have been very few studies on asynchronous technologies, such as e-mail or written threaded discussions. Furthermore, results from the studies that have been conducted on asynchronous written threaded discussions are ambiguous.

In this study, learning discussions in Face-to-Face and written asynchronous threaded discussion groups were compared with respect to a) the students’ interactions with each other and the content of the simulation, and b) the accuracy and complexity of their chemistry reasoning.

Thirty female students participated in the study, working in dyads while performing the simulation. Student attitudes towards learning and level of prior knowledge were either measured or estimated. Average level and difference in ATL within dyads was calculated. Eleven dyads communicated via threaded
discussion groups and four dyads, whose discussions were video-recorded and transcribed, conducted their discussions Face-to-Face. The methods that were used to assess the two aspects of the discussions, mentioned above, were described in paper II as Methods B and C, respectively. To facilitate the coding process, Method B was modified by using a subset of the original categories.

Results indicate that the threaded discussion promoted higher level chemistry reasoning while Face-to-Face discussions facilitated co-operative activity, demonstrated by longer and more intense discussions, frequent questions, and consideration of partners’ statements. Although results were difficult to interpret, it was suggested that group configuration with respect to differences and average level of Attitudes toward learning (ATL) in dyads, did not seem to have had an influence on performance to any great extent. Although a high average level of ATL in dyads was negatively correlated with faulty reasoning, no correlation could be found between average ATL and higher level reasoning. Since assessors perceived that the general level of chemical knowledge was low in threaded discussion groups as well as in Face-to-Face groups, it was suggested that the poor chemical knowledge may have inhibited the higher level chemistry reasoning that was expected in groups with high level of ATL.

5.4 Paper IV

Student attitudes toward learning, level of pre-knowledge and instruction type in a computer simulation: effects on flow experiences and perceived learning outcome

Attitudes toward learning (ATL) are closely related to epistemological beliefs, which have been shown to influence students’ approaches to learning, as well as the outcomes of learning (Hofer & Pintrich, 1997). There are, however, very few studies on how students’ beliefs or attitudes interact with the learning situation to influence student motivation and learning outcomes. In this study there were two groups of students: one group used an open instruction format and the other a guided instruction format. Each group comprised students with Attitudes toward learning that spanned from low to high level ATL. There was no significant difference in average level of ATL between the groups. We examined the effects of instructional format on students’ perceived cognitive change, level of engagement, and the flow variables (Csikszentmihalyi, 1990; Hedman, 2001): Control, Pleasure, Challenge and Concentration.
The guiding instruction format was correlated with perceived conceptual change, relatively high levels of Engagement, Concentration, Challenge and Pleasure (motivation). Although it was a less important variable in the model, the sense of control was low in the group that had used the guiding instructions. For the open instructions, the reverse was true. Thus, the guiding instructions seem to have been more motivating and productive for students, but at the expense of perceived Control. As learning requires cognitive conflict, and students with open instructions had the opportunity to choose what to discuss, it was suggested that students with open instructions avoided cognitive conflict and therefore reduced uncertainty - and conceptual change. The finding that perceived conceptual change was correlated with motivation was rather uncontroversial since it fits most models of motivation.

Attitudes towards learning were only weakly positively correlated with motivation and conceptual change, and weakly negatively correlated with sense of control. It was suggested that the low impact of ATL on students’ perceptions of the two instructional formats was because there was too little difference between either ATL or the character of the instructions.

6 Methods

There is a debate as to the nature and extent to which qualitative data should be quantified (Sandelowski, 2000, 2001; Wegerif & Mercer, 1997). According to Kvale (1997), it is surprisingly difficult to reach consensus. As Wegerif and Mercer (op.cit) point out, it is claimed that qualitative methods are better for explaining causal relations and are more context-sensitive. Quantitative methods, in contrast, are supposed to be more suitable for handling large amounts of data, and producing results that are more objective and testable with respect to validity and reliability. As the results of paper II indicate, the ongoing debate seems to be unnecessary since different approaches can provide complementary pictures of a phenomenon and may even arrive at the same conclusions.

In this thesis, mixed-methods were used extensively. Typically, the interviews and discussions were characterized using a categorisation scheme. The development of the schemes was either preceded or ran in parallel to a qualitative, iterative approach to discern the categories that could be used for describing the discussions and interviews. In paper I, categories were sought that would fit into a scheme, describing a general theme in the shape of different levels of students’ chemistry reasoning quality. In this process, with a
predefined theme as a goal, the categories and the scheme were developed in parallel. In papers II and III, the categories covered a broader aspect of the discussion. There was no apparent underlying “hierarchy” between categories, since the purpose was only to describe as much as possible of what was observable with respect to the students’ interactions with their partners and the content of the simulation. Instead, the underlying general themes, e.g., “pays attention to partner”, “motivated behaviour”, or “metacognitive activity”, emerged when the distributions of scores in categories were analysed after the coding. Then the scheme was constructed by selecting categories that were important for describing these themes.

The quantitative coding of utterances, i.e., assigning a numerical value (score) to a specific category, was based on qualitative judgments of the meaning of student’s statements which often required the consideration of a wider context than the isolated comment. The units of analysis were small, e.g., a single comment or short discussion on a subtopic. From this approach it followed that each discussion was not only described from several aspects (categories), but also generated a relatively large number of “observations”; e.g., statements, contributions, or sub-discussions, which, to some extent, compensated for the small sample sizes. For example, in paper I, student’s statements in interviews were coded using seven different categories, representing different ways to use chemical knowledge, resulting in between 35-154 scores in each sub-group of interest, e.g., experiment or control group, or students with relativistic or dualistic ATL. In paper II, 26 categories were used to characterize the discussions and to describe differences and similarities between them. Up to 787 scores in a single discussion were recorded (average 278), generating a total of 3053 scores for all groups together. In paper III, discussions were scored using 13 categories, resulting in up to 246 scores per dyad. Although the small sample sizes imply that generalizations outside this specific context should be made with some caution, the high number of “observations” supports the proposition that the observed similarities and differences are real and that the application of statistical methods to estimate the significance of results and to test hypotheses was appropriate.

Given the limitations of working memory and the number of variables and observations, results were very difficult to interpret without simplifying them. Therefore, principal component analysis (PCA) was applied to distinguish and make the latent structures in the data comprehensible. This strategy is frequently used in sensory science, where statistical methods are routinely employed for analysing qualitative data (Naess & Risvik, 1996). In both cases, a major task is to find ways to describe commonalities and differences in how individuals talk about or perceive an object e.g., a concept in chemistry, a simulation exercise, or the taste of a wine.
An individual’s experience of a wine, comprising more than 800 aroma components, “…consists of chemical components in interaction with our senses, and the perceived entity by the individual” (Risvik, 1996). Thus, the perceived nature of a specific wine is likely to differ between individuals, depending on, for example, their physical ability to sense the chemical components, and the way they mentally construct the composite picture of the wine, e.g., based on previous experiences, and personal preferences. Risvik argues, the perceived entity is not the sum of all attributes of that object, but rather a latent structure resulting from the aggregation of information. It is not known whether we focus on key attributes, aggregates of attributes (concepts), or holistic forms when inferring these mental latent structures from the attributes. Nor is it known whether this aggregation takes place at the sensory stage or in the brain. Theories of working memory and schema construction (Baddeley, 1986; John Sweller et al., 1998), imply that “mental latent structures” could be similar to “schemata”, providing means to circumvent the limitations of working memory, as also suggested by Risvik (1996).

Not only are these ideas central to our understanding of human information processing, they have also influenced the choices of methods used in this thesis. As in the wine-tasting example above, a judgment of the quality of a discussion between two students could be influenced by the assessor’s ability to identify (sense) the important attributes that make up the quality, any assumptions regarding what components are most important for describing quality, and the way the different attributes or aggregates of attributes are weighed together into a perceived entity, i.e., the mental latent structures that describe the quality. Thus, the influence of the researcher on the final results may be significant and must be handled appropriately.

The methods used in this thesis have one common theme; they provide means to disconnect the researcher from the final latent structures that describe the studied phenomenon. I do not claim, however, that it is possible to completely remove the influence of the researcher. For example, every decision that was made during a categorization of a discussion or interview was based on my previous experiences of similar situations or my perception of how the context had led to a specific statement. However, the final judgment of the quality of a student’s reasoning was not under the immediate influence of the researcher since it was not based on a single categorisation, and in papers II and III neither was it based on the sum of all categorizations. Instead, the latent structures that summarized the meanings of combinations of categories emerged for the first time when the PCA was applied to the data. Since the most influential category is not necessarily the most frequent, but rather the one that contributes most to the variation between groups, and since coding was recorded for one group at a time, it was very difficult for the researcher to predict or influence the outcome.
of the PCA in any specific direction. Hence, objectivity of analysis can be improved using a “semi-blind” approach like this.

There are other benefits attendant on mixing qualitative and quantitative methods as in this thesis. For example:

a) the correlation pattern between categories provides information on the validity of the construct, e.g., coding scheme or questionnaire, and may assist in the refinement of instruments (Gardner, 1996)
b) quantifying qualitative data enables comparisons across different experiments
c) data mining could be employed to find and explore unexpected relationships in the data and to detect and describe bias caused by background variables, as in my case, the behaviour of interviewers or the format of the interview-guide.
d) data can easily be accumulated over time to increase model reliability.

Points b and d imply that mixed method approaches could facilitate surmounting the problem with the small sample sizes inherent in qualitative research methods, because data from repeated studies may be used to expand the sample and check the consistency of the model.

6.1 Interviews

Interviews constituting the research for paper I were conducted to assess the students’ ability to handle chemical knowledge as an indirect measure of the accessibility and complexity of their schemata. However, in particular section 6.1.5 “reporting data from interviews” also has a bearing on discourse analysis.

6.1.1 Purpose

The most commonly mentioned strength of interviews is the opportunity for exploring the range of different possible perspectives among interviewees on a subject. Kvale (1997) divided the purpose of the research interview into:

a) Empirical aspects, aiming at exploring or describing a phenomenon
b) Theoretical aspects, to test or develop a theory (grounded theory)
I am not sure whether this division is relevant because, in the present thesis, interviews have been used for both these purposes simultaneously.

Following the advice of Oppenheim (1992), that qualitative exploration of a domain is a necessary step before constructing quantitative tools, the coding schema that was used in paper I was based on a qualitative analysis of the different ways in which students expressed their chemical knowledge in the interviews (empirical purpose). Subsequently, results from the coding of the same interviews were used to test the hypothesis that the simulation lead to intuitive knowledge of the subject of acid–base chemistry, and whether students’ attitudes towards learning influenced that characteristic of their knowledge (theoretical purpose).

Hence, the purpose of the interviews was not so much to investigate the different conceptions of acids and bases among students as to test students’ abilities to apply chemical knowledge in a complex and cognitively demanding situation. Recently, van Bruggen, Kirschner and Jochems (2002) provided support for the idea that interviews, i.e., synchronous communication, might impose the necessary amount of cognitive load to discriminate between different levels of functionality of knowledge (schemata).

6.1.2 Selection for interviews

Given the tedious and time-consuming task of analysing the interviews, only a limited number of interviews could be performed. In order to achieve valid results, purposeful sampling (Seidman, 1998) was employed. Thus, given the focus of paper I, contrasting groups (Oppenheim, 1992; Seidman, 1998) with respect to students’ attitudes towards knowledge were identified for interviews.

6.1.3 Implementation

An interview is an interaction between the interviewer and the interviewee. Thus it is not possible to perform an interview without influencing the interviewee. The structure of the interview and the behaviour of the interviewer, i.e., the questions the interviewer asks and how they are put, are potential sources of bias (Kvale, 1997), threatening the validity of the interview data. Bias could arise both due to subtle “planting” of ideas by the interviewer (leading questions or wording), and the interviewees’ own perceptions of how a statement would make them appear in the eyes of the interviewer or themselves, i.e., prestige-bias. The latter was of particular concern when the interviews in paper I were conducted.
The main question was whether we should let the subjects know what information we wished to elicit, i.e., their level of chemical knowledge, or not. Telling them what aspects we were interested in would make students focus on figuring out what we might want to hear, and the possibility of them withholding important information about the level of their spontaneous use of chemical knowledge. Since the focus of the interview was to examine the accuracy and complexity of students’ spontaneous use of chemical knowledge, this was obviously not an option. Instead, much effort was spent avoiding explicitly examining chemistry questions, but, at the same time, providing students with rich opportunities to use chemical knowledge in their stories.

In practice, we asked students to tell us about their experiences and perceptions of a practical laboratory exercise that they had done immediately before the interview. Our impression was that, as long as we did not put pressure on them to explain chemical concepts or phenomena, students were relaxed, answered openly and used chemical knowledge spontaneously. In the second sub-study in paper I, in an attempt to gain even more information about the students’ chemistry reasoning abilities, the interview-guide was revised and new items were introduced that explicitly asked students to perform interpretations of diagrams at a certain level, e.g., connecting the equivalence point and the speciation of the weak acid in the analyte. In these interviews both interviewers noticed that students seemed to feel uneasy when explicit questions were put. This was manifested by, among other things, the students making wordy excuses for not having studied the subject as thoroughly as they obviously thought they should have.

For the kind of research-interviews we wanted to perform, this methodology has at least two serious drawbacks:

a) prestige-bias, as mentioned above, is likely to occur when interviewees feel that their knowledge and indirectly their performance as students is being examined

b) since student responses are no longer a result of their own choices, assumed to be based on the accessibility of their chemical knowledge, the ability of the interview to assess the accessibility of students’ knowledge is lost or impaired

To gain rich information, we tried to let the student talk as much as possible while the interviewer tried to keep all questions short, using silence wisely to increase the pressure on the interviewee. For the sake of validity and to facilitate analysis, interviewers asked follow-up questions and carefully probing questions whenever required, in order to clarify whether their interpretation was
correct. For a more thorough review of different question types and approaches that might be useful in interviews, see Kvale (1997).

Interview guides were semi-structured, i.e., the order of the questions could be adjusted according to the progress of the interview and allowed for additional, clarification or probing questions where needed.

As a general reflection, it was clear that, to be able to make relevant on-line interpretations and follow-up questions, and to keep the interview moving in a productive direction, the interviewer had to have relevant knowledge in the domain, as well as being a skilled interviewer.

6.1.4 Documentation and analysis

There seems to be a strong preference for the use of transcriptions of interviews in the literature. Since the ability to consider the context in which statements are made is considered to be a major strength of qualitative research, and given the available technical solutions for handling sound and video, the seemingly habitual transcription of interviews is surprising.

As Kvale (1997) pointed out, there are several problems with transcriptions, pertaining to the loss or distortion of the original situation that may occur in the transcription process. Since translation of talk into written text requires an interpretation of what is said, this step adds a source of error that might threaten the reliability and validity of the transcript. The transformation, or de-contextualization, of the interview makes it a poor representation of the original dynamic discourse. In a worst-case scenario it could change the meaning or strength of statements and words. Furthermore, due to the differences in the “conventions” of written and spoken conversation respectively, a direct transcription of talk often results in texts that may lead the reader to question the intelligence of the participants, raising issues about ethics and students’ motivation to participate in future studies.

The ideal interview would, of course, be the one that is already analysed when completed. However, owing to the detailed and systematic analyses that were employed in this thesis, and given the limited memory capacity of the researcher, this was not an option. Instead, as a compromise between preserving as much as possible from the original situation and facilitating data handling, interviews were recorded on minidisc, copied on to a computer hard-disc, and imported into the “Qualitative Media Analyzer” software (QMA; Skou, 2001).

QMA allows the analyst to code passages directly in the sound file, jump between coded passages for re-listening and re-coding, and to compile and
replay passages within a certain category from all interviews, thus allowing for control of the consistency of coding between interviews. As the data, e.g., labels, comments and timestamps, that are associated with each passage can be exported for statistical analysis, this is a very convenient way to code and analyse interviews.

Listening to hours of interviews, and systematically and consistently interpreting the meaning of what the interviewee says, is tiresome enough without having to deal with background noise such as scraping from chairs, hissing, distortion due to poorly calibrated recording volume, or the interviewee (or yourself) shuffling around keys on the table. Since fatigue threatens reliability, a quiet environment and good sound quality is essential for analysing, and preferably for recording, interviews.

In subsequent interviews, not described in this thesis, stereo rather than mono recordings markedly increased sound quality. Also, elevating the (high quality) microphone somewhat, i.e., not placing it directly on the table, reduced unwanted background noise. Most of my recordings were digitalized with 22 kb/sec and 8 bit compression, which gave a technically acceptable sound quality.

6.1.5 Reporting data from interviews

There still seems to be a discussion in the literature about how interviews should be analysed and reported. The method that was used for visualizing the qualitative data in paper I has been criticized for not considering “the established norms of qualitative research” (McGinnis, J.R. & Swanson, E., pers.com.). Given the growing number of educational researchers advocating mixed method approaches (Hmelo-Silver, 2003; Naess & Risvik, 1996; Sandelowski, 2000; Wegerif & Mercer, 1997), comments of this type probably emerge rather from the epistemological and ontological beliefs of the reader than from scientific evidence.

Furthermore, the boundary between qualitative and quantitative research is not always clear, if existing at all at the data level. For example, a qualitative analysis of what themes (categories) occur in different interviews or discussions might be directly translated into counts of how often and in which interviews these themes occur, thus allowing for additional statistical exploration of data, without necessarily limiting the possibility of making qualitative interpretations from the results (Trochim, 2004). In these cases, and provided that the description of the coding procedure and categories allow the reader to judge the
validity of the scheme and the coding procedure, the application of appropriate statistical tests improves rather than reduces the credibility of the conclusions.

However, as Kvale (1997) points out, the frequency of a statement does not always convey information on what the statement actually means. In Kvale’s example, vivid denial could mean the opposite, that is, the interviewee is aware of, or agrees with something, but denies it because he or she is embarrassed to be connected with the subject. The question is then how frequently a denial has to be made in order to turn it into a confirmation. It is, of course, not possible to give a general answer to this question. Just as has been done in the analysis of the interviews in paper I and the discussions in papers II and III, the meaning of a statement has to be judged by considering the context in which the statement is made, for example, previous statements by the interviewees and/or the atmosphere of the interview.

There is a strong relationship between methods of analysis, format of report and the research question. Since there seems not to exist any consensus on methods for analysing and reporting results from interviews or discourses (Kvale, 1997), the choice of methods should be influenced by the research question, the type of interview, the researchers’ competencies and epistemological beliefs, and, not least, assumptions regarding the readership of the research report.

Since the aim of the interviews was to estimate the magnitude of the impact of the simulation on students’ reasoning abilities as well as making comparisons between groups, the mixed method described above was judged appropriate.

Since chemistry education research, in the eyes of the author of this thesis, partly aims at bridging the gap between pedagogical and psychological research on one hand and, specifically, teachers in natural sciences on the other, the methods of analysis and format of reports were to some extent chosen in order to align with the research paradigms that are common in natural sciences. I am, however, aware that my choice of methods imposes a “subjectivity of perspective” on the analysis (Kvale, 1997), i.e., the interviews were interpreted from only one perspective – the accuracy and complexity of the chemical knowledge that was expressed by students in the interviews. There is no reason to deny there could be other, equally valid, perspectives on the performance of students in the interviews. However, I am quite convinced that the methods I have chosen are appropriate for their particular purposes, given my own competencies and epistemological beliefs.
6.2 Questionnaires

The validity of self-reports for retrospective assessment of students’ mental processes during an exercise can be questioned because students’ reconstructions of what happened could be biased, e.g., by prestige or simply by an unclear memory of the mental processes that were involved (Richardson, 2004).

Part of the solution to this is to collect data as soon as possible after the activity, to design questions that reinstate the psychological context in which the activity took place, and to phrase the questions so that all alternatives are equally acceptable for the respondent. For these reasons I do not believe that questionnaires designed to be generally applicable over a wide range of disciplines, cultures, situations, or students, are valid in all these contexts. Response bias, for example, where the respondent tends to score “to the right” or “to the left”, is another threat to the validity of questionnaires. The most obvious solution to avoid response bias is to use questions with different polarities.

Consequently, the questionnaires used in this thesis were either directly constructed for a specific context (e.g., learning chemistry in higher education in Sweden, as in the attitude questionnaire) or complemented by additional items that provide the link to the specific situation (e.g., the flow questionnaire in paper IV). Furthermore, the questionnaires were administered in situ or, in the case of the flow-questionnaire (which assessed student responses to a particular exercise), immediately after the exercise. Mixed-polarity items were used in the attitude questionnaire and the second part of the flow questionnaire, although with a slight emphasis on favourable items in the latter.

Typically, questionnaires are validated according to one or several of the following criteria:

a) Concurrent validity, i.e., correlation between outcomes of different instruments measuring the same feature.

b) Predictive or discriminative validity, i.e., correlation between outcome of the instrument and some aspect that is assumed to be related to the feature measured by the instrument, e.g., age, gender, educational level, cognitive development, or problem solving ability.

c) Perceived validity, i.e., subjectively judged by knowledgeable assessors in the field.

d) Construct validity, meaning that statistical analysis of the latent structures in responses to the questionnaire yield an expected number of
components, describing an appreciable amount of variation, and with intelligible patterns of correlation between items in the different dimensions. Since alphas (such as Chronbach’s) are functions of the number of items, they might yield high values of internal consistency when average inter-correlation is low and might even erroneously indicate single dimensionality for scales that are not uni-dimensional (Eagly & Chaiken, 1993; Gardner, 1996). Hence, for all questionnaires except the engagement mode questionnaire, principal component analysis, revealing co-linearity among individual items, i.e., using all the information in the material rather that averages, was applied to examine dimensionality and internal consistency of constructs.

6.2.1 Attitudes Toward Learning

Initially, the purpose with of this questionnaire was to find contrasting groups with respect to students’ attitudes towards learning. In paper I, these groups were called HiPos and LoPos, with HiPos expressing attitudes toward learning that were more relativistic than in the LoPos group, who could be described as having a more dualistic perspective. In the later papers, no contrasting groups were used. Instead, the numerical values of the students’ PCA scores were used as a measure of their attitude toward learning and the term ATL (Attitude Toward Learning) replaced the HiPos and LoPos groupings.

As mentioned under section 4.1, student attitudes toward learning might influence student motivation and approaches to learning. Hence measuring student attitudes is important, not only for research purposes but also for the design of effective learning situations. Since the very laborious and time consuming interviews performed by Perry (1970) in the early 1950s, several paper and pencil tests have been developed, allowing for relatively rapid assessment of students’ epistemological beliefs, and their beliefs about learning, teaching, and intelligence, i.e., attitudes toward learning as defined in this thesis.

Ten of the most influential questionnaires in epistemological research have been reviewed by Duell and Schommer-Aikins (2001). Interestingly, the diverse focuses of those questionnaires reveal a seeming lack of consensus of what constitutes epistemological beliefs, though Hofer and Pintrich (1997) tried to draw acceptable boundaries.

Generally, the analytical instruments focus on one or several of the aspects of knowledge and the process of acquiring knowledge shown in Table 1, but some
also include beliefs about the speed of learning ("successful students learn quickly"), the role of the teacher, and intelligence (fixed ability versus something that can be developed by challenging your knowledge).

In this thesis, the 34-item questionnaire (Appendix 1) used in papers I-IV was developed by Berg et al. (2003), originating from previous questionnaires by Henderleiter and Pringle (1999), Reid (2003), and the adaptations of Perry’s (1970) model of intellectual development to the domain of chemistry by Finster (1989,1991).

As mentioned in section 4.1, the attitude positions measured by the questionnaire are closely related to epistemological beliefs, but also include the role of peers, assessment, and instruction. Principal component analysis has been applied to analyse the dimensionality of the questionnaire, and to predict students’ attitude positions, relative to each other, somewhere in the area between Perry’s positions 2-5. To ensure that the predicted relative attitude positions were comparable between the studies in this thesis, the same PCA model has been used for predictions throughout all papers.

The PCA of the attitude questionnaire used in this thesis produced a single component model which did not show any significant changes in the relationships between items from one test-occasion to another. The validity of the questionnaire was also supported by its ability to predict learning outcomes from the simulation in paper I and concurrent results from interviews assessing similar aspects (Berg, 2005).

6.2.2 Flow and perceived learning outcome

This process-oriented questionnaire (Appendix 2) was used in paper IV in order to complement the summative evaluation in paper I. The questionnaire was based on a translated full version of a previously validated questionnaire constructed by Ghani and Despande (1994), focusing on students’ Flow Experiences (Csikszentmihalyi, 1990). This version contained 15 items, assessing the flow components, Pleasure, Concentration, Control, Exploration, and Challenge.

To gain additional information on students’ perceptions of the learning outcome of the simulation, and to strengthen the analysis of the relationship between flow and the independent variables, 36 additional, Likert-type items were constructed. The 30 new items, that were designed to provide similar information as the items in the original construct, differed from them by referring less explicitly to the flow components and being more adapted to the
specific situation. To examine the dimensionality and internal consistency of the flow components, PCA was applied to the students’ responses to the questionnaire. Besides high internal consistency within the flow components, the model showed a high degree of positive correlation between the flow components Pleasure, Challenge, and Concentration which in turn were negatively correlated with Control. Since the simulation exercise focused on cognitive rather than operational training, it did not give much room for Exploration, i.e., testing different ways of performing operations. This flow component did not emerge as important in the model and items in the Exploration component were therefore excluded together with other poorly described items (19 items in total).

6.2.3 Engagement modes

Since engagement modes (Hedman, 2001; Montgomery, Sharafi, & Hedman, 2004) were assumed to influence students’ performance when using computers for simulation and communication, they were assessed as background variables in paper III. The five Engagement Modes assessed by the questionnaire were: Efficiency / Productivity, Enjoyment / Acceptance, Ambition / Curiosity, Avoidance / Hesitation, and Frustration / Anxiety. Each engagement mode comprised 4-5 items, 23 in total. The last two modes are considered to be negative while the first three are positive engagement modes.

The responses to this questionnaire were evaluated by summarizing the individual’s average scores on items in the positive and negative engagement modes respectively. Analysis of variance (ANOVA) was then applied to detect differences between experimental groups for the sums of averages in the positive and negative modes respectively. Since this procedure had been validated in previous studies, we did not perform any further examination of internal consistency and construct validity.

6.3 Discourse analysis

There is little consensus on what discourse analysis is and how it should be performed. As used in this thesis, “discourse analysis” refers to an analysis of an observed cognitive behaviour rather than to a conversational or rhetorical analysis. Again, it seems that the choice of method is more a question of the researcher’s own epistemological beliefs than which method is necessarily the most appropriate. There are, however, some “don’ts” in discourse analysis that
might apply. Antaki et al. (2003) listed six typical errors that could threaten the validity of such analysis:

1. Just summarizing data leads to loss of information, and subjective summarizing may distort the picture of what really happened during the discourse.

2. By taking sides, the researcher may subjectively emphasize or censure features in the discussion, leading to simplification or distortion of data.

3. Lists of selected quotations, which are often used in “profiling” of discussions, with little commenting from the researcher, disconnect the quotations from their context, which procedure might change their meaning. Not considering how the meaning of statements depends on the context may threaten the validity of the analysis.

4. Using quotations to infer the presence of a feature, e.g., “cooperative activity”, and then explaining the quotations in terms of cooperative activity is circular argument rather than analysis.

5. Over-generalization, based on a small sample.

6. Simply spotting features is not analysis if they are not also examined with respect to how they contribute to the character, or development, of the discourse.

As mentioned previously, the choices of methods in this thesis were based on the assumption that they would reduce the influence of subjectivity in the researcher. Quotes were used merely to exemplify how they were analysed, i.e., categorized, in the light of the context. Hence, “summaries” like the interview-profiles in paper I represent summaries of analyses of statements rather than piles of unprocessed quotes. In papers II and III, the PCA used categorizations (analyses) of every statement that was made, covering virtually whole discussions rather than selected samples. In addition, when interpreting the possible meaning of the categories, or clusters of categories that the PCA revealed to be important for describing different aspects of the discourses, previous research in the field as well as triangulation was used in order to validate interpretations. Although I believe that the results from the many observations in the papers forming this thesis, yield a reliable and valid picture of the reasoning abilities of these groups and individuals; they are not
necessarily true for other groups of students working under different circumstances.

7 Main conclusions and discussion

The simulation had substantial effects on students’ knowledge, demonstrated by an increased spontaneous use of chemical knowledge and higher levels of chemistry reasoning in complex situations. Considering the research on intuitive knowledge (Swaak & de Jong, 1996), cognitive load (John Sweller et al., 1998), and cognitive schema theory (Derry, 1996), I suggest that these effects were due to the students’ development of more coherent and complex schemata in the acid-base reaction domain, leading not only to the freeing of working memory capacity, but also to facilitating knowledge retrieval from long-term memory.

Tasks that support students’ regulation of scientific discovery seem to have been beneficial for both motivation and learning outcomes, concurrent with the findings of de Jong et al. (1999), who found that using a mix of open- and closed-ended tasks lead to improved intuitive knowledge in the domain of physics.

The results in this thesis are ambiguous regarding whether open-ended or closed-ended tasks are to be preferred since there is evidence that both types produce positive results. When considering perceived improvement of declarative knowledge and student experiences of flow/motivation as measures of the outcome of the simulation, the closed-ended tasks seem to be more efficient. No comparison was made between the two types of tasks when intuitive knowledge was used as a measure of efficiency. Only open-ended tasks were used in paper I, where significant effects on student intuitive knowledge were found. However, in this case the comparison was between students who had used the simulation and those who had not. Thus, it was not possible to specifically attribute the improvement to the open-ended tasks. Sweller (1988) showed that closed-ended tasks, with an explicitly expressed aim, imposed higher levels of cognitive load on the novice learner than open-ended tasks. He argued that this cognitive load is mainly extrinsic or intrinsic, i.e., not directly related to the process of evaluating and refining the student’s own knowledge. Hence, since the working memory capacity is limited, “means-ends-analysis” strategies may hamper schema acquisition, i.e., learning.

In paper I, it was suggested that the simulation provided a more holistic view of the subject area and provided an opportunity for intense interaction with the content of the simulation, i.e., experiential learning. In addition, the open-ended
tasks were supposed to lead to more autonomous forms of extrinsic motivation (Ryan & Deci, 2000) that would support learning (Marton & Säljö, 1984) In retrospect, considering the findings of Sweller (1988), it seems possible that the open character of the tasks also led to a more efficient working memory usage during the simulation, making allowances for germane cognitive load, adding to the effects on students’ intuitive knowledge seen in paper I.

Given the arguments in favour of open-ended tasks, the results from the comparison between open- and closed-ended tasks in paper IV might seem somewhat surprising since students perceived the closed-ended tasks as being more efficient in promoting conceptual change than the open-ended ones.

There are, however, strong indications that the effects of (scientific) discovery learning with simulations are primarily on the intuitive quality of knowledge (de Jong & van Joolingen, 1998; Swaak, Van Joolingen, & de Jong, 1998). This type of knowledge, or skill, is very difficult to verbalize and is mostly manifested only as an improved ability to apply knowledge in complex situations (Swaak & de Jong, 1996). Thus, the questionnaire that was used to assess perceived learning in paper IV might not have been ideal as an indicator of learning outcomes, since it was not optimized to measure intuitive knowledge and because students might not have been aware of any improvement in the intuitive aspect of their knowledge. Thus there might have been a difference between groups with regard to intuitive knowledge that was not detected in paper IV.

Although it is tempting to speculate that the open-ended tasks in the open instruction would support a more efficient schema acquisition, additional experiments would be required to elucidate this question.

The attempt to describe the quality of the learning discourse during the simulation revealed several components of quality: co-operative activity, correctness and complexity of chemistry reasoning, engagement / motivation, ability to realize cognitive conflict, and reference to theory while reasoning. Three qualitatively different methods, focusing on slightly different aspects of student discourses, were used in paper II, showing high consistency in their final judgment of the relative quality of the discourses. This indicates that “quality” is an underlying feature that permeates many aspects of the discourse and, consequently, could be targeted in different ways, e.g., focusing on quantitative as well as qualitative aspects.

As the use of IT for communication within distributed learning becomes increasingly common, the question of whether the distributed learning situation influences the learning discourse is becoming highly relevant. In paper III, dyads that communicated via written threaded discussion forums (PP groups)
were compared with dyads that communicated face-to-face (FF groups) while performing the simulation exercises. The main conclusion from this study was that the distributed learning situation supported higher-level chemistry reasoning (accuracy and complexity) and frequent references to theory while the face-to-face situation allowed for longer and more intense discussions and a higher degree of co-operative activity. Not surprisingly, high quality discussions were characterized by high scores in all these components. The reason why FF groups did not perform better on the chemistry reasoning aspects might be explained by the low level of pre-knowledge that identified in the PP as well as the FF groups. Lacking fundamental knowledge in the subject area hampers successful co-constructive learning, since students might not have the information and schemata that are necessary to evaluate and bridge the gap between their own knowledge and new information to explain experimental results. In this situation the PP groups might have benefited from the extra time available to acquire and consider relevant information prior to contributing to the discussion.

Thus, there are indications that novices might benefit from a slower mode of communication while relatively good prior knowledge is required in order to benefit fully from the face-to face situation, which would then support efficient co-constructive learning. I therefore suggest that future research should investigate how the level of prior knowledge, combined with the mode of communication, influences learning in computer-stimulated collaborative learning.

In paper IV, the flow experiences questionnaire (FEQ) was able to detect differences between groups working with instructions with open- and closed-ended tasks. The correlation-pattern, with positive correlations between the flow components Pleasure, Challenge, and Concentration which in turn were negatively or uncorrelated with Control, was understandable. These findings are concurrent with the results of Ghani and Deshpande (1994), who found that Control was positively correlated with flow only in situations that involve the use of a variety of skills and require a high degree of autonomy. Thus, the correlation-pattern found in this study indicates that the simulation exercise was not challenging enough to impose a need for control, strong enough for control to emerge as a component of flow.

Our results support the discriminative validity of the construct and indicate that these types of variables might be useful for evaluation of learning processes.

Csikszentmihalyi (1999) argued that flow, being a state where skills and challenges are matched, serves as a positive influence on learning by motivating the learner to either increase skills or seek new challenges, provided that the
differences between the degree of skill and the degree of challenge are not too large. Thus, measuring the degree to which learners experience the suggested components of flow should indicate the potential for learning in an activity.

Today, the simulation that was studied is used on regular B-level courses in aquatic chemistry at Umeå University. It is gratifying that results from chemistry education research have influenced everyday practice. Additionally, this is thanks in part to the efforts of perceptive and progressive teachers who appreciate and take cognisance of other forms of knowledge than the purely declarative, despite such forms of knowledge being more difficult to assess.
Acknowledgements

Anna – för att du är – helt underbar... (och för att du trixat med dina arbetstider så jag kunnat göra ”bryggrycket”)


Mamma Leena och Pappa Pentti – För att ni alltid trott på era ”ungar”. Jag är stolt (och lycklig) över att ha sådana föräldrar. (Men jag är faktisk 42 och kan själv bestämma när jag ska ha mössa på mig...och ja, jag aktar fingrarna när jag sägar ..)

Lillebror Per– Du har betytt mer än du tror. Du har har bevisat att intresse, integritet och engagemang (m.a.o. envishet som en äsna....) är viktigare än universitetspoäng när det gäller att skaffa sig spetskunskaper. Jag ger dock inte upp än....

Karl-Gunnar och Ing-Britt – Utan all hjälp med barnpassning hade nog inte denna avhandling blivit klar i tid. (Ni är dessutom himla trevliga...)

Bruno: För att du startade kemididaktisk forskning i Umeå och gav mig en positiv bild av vad forskning innebär.

Leif: För att du botat min psykologskräck (Jag tror inte längre att du kan se vad jag tänker....). Glada skratt och milda ord av allvar i en salig blandning gjorde handledningsträffarna till angenäma upplevelser

Anders: I didn’t go for the golden bullet
but for the elephant
It wasn’t simple, but relevant (?).

P.S: Två hjärnor tänker inte alltid bättre än en – men det är definitivt roligare!

Christina – För mentalt stöd och för avkopplande, livsnödvändiga skratt.
Jag kommer inte att glömma dig.

Alla vänner – För njutbara middagar, bad-, fiske-, seglings- och paddel-utflykter och allting annat som hör till ett rikt liv.

MaTNv för er entusiasm och framätanda– jag hoppas att jag får mer med er att göra

”Kemi gemensam” för att ni alltid är positiva och hjälpsamma, det underlättar

Natur och Kultur – för att ni alltid varit trevliga att ha att göra med och för att ni gav finansiella möjligheter att genomföra detta projekt.

Editors and anonymous reviewers – thanks for contributing to making my skin thicker and providing ideas on what a high-quality report should look like.

Henry Heikkinen – Teaching us about “the American way” to play the game of publishing.
References


Appendices

Appendix I
Appendix 2

Part 1
The following questions ask about your feelings while performing the simulation. Please describe the simulation by placing check marks on the scales given below.

Enjoyment
1. Interesting    ☐ ☐ ☐ ☐ ☐ Not interesting
2. Fun    ☐ ☐ ☐ ☐ ☐ Not fun
3. Exciting    ☐ ☐ ☐ ☐ ☐ Not exciting
4. Enjoyable    ☐ ☐ ☐ ☐ ☐ Not enjoyable

Concentration
5. Deeply engrossed in the activity ☐ ☐ ☐ ☐ ☐ Not deeply engrossed in the activity
6. Intensely absorbed by the activity ☐ ☐ ☐ ☐ ☐ Not intensely absorbed by the activity
7. Attention was focused on the activity ☐ ☐ ☐ ☐ ☐ Attention was not focused on the activity
8. Concentrated fully on the activity ☐ ☐ ☐ ☐ ☐ Did not concentrate fully on the activity

Control
9. I clearly knew what to do ☐ ☐ ☐ ☐ ☐ I felt confused about what to do
10. I felt calm ☐ ☐ ☐ ☐ ☐ I felt agitated
11. I felt in control ☐ ☐ ☐ ☐ ☐ I did not feel in control

Challenge
15. In general, how challenging did you find the activity?
    Very much ☐ ☐ ☐ ☐ ☐ Low
Part Two 2

Indicate how well the statements below describe your own perceptions of the simulation activity. The closer to the statement you mark, the more you agree with it.

16. I had no problems to understand what to do

17. The exercises stimulated to deep discussions about chemistry

18. The simulation was meaningful

22. I was engaged by the exercises

23. The discussions made me engage in the activity

24. I was not sure of what to do

25. I was not sure whether my reasoning was correct or not

26. I was very engaged in trying to understand the underlying chemistry in the simulation

27. The chemistry-discussions in my group were very interesting

28. I got several own questions during the simulation

30. The exercises made me interested in learning chemistry
31. I became stressed when I did not understand
32. Several own questions were raised during the simulation
36. The exercises were instructive
37. The simulation was a good way to test your knowledge
38. The simulation helped me to repair deficiencies in my knowledge
39. The simulation helped me find answers to my own questions
41. The simulation gave new knowledge in the acid-base and buffers domain
42. I discovered deficiencies in my knowledge in the acid-base and buffers domain
44. I felt that the chemistry that was treated in the simulation has become more comprehensible