ABSTRACT
Based on the model of Educational reconstruction that was proposed by Kattmann et al. for Science Education in 1997, we derived a specific framework for the design and development of lessons in Computer Science (CS). We present and explain this framework that is constructed according to the specific goals, knowledge structures, and teaching methods of Computer Science at schools. This framework is arranged around suitable “real world phenomena” in the sense of a teaching context. Thus, it may be applied as a road map for the “CS in context” approach as well as to other teaching approaches for CS at school. In order to validate this framework, we have conducted some empirical research in order to propose and test suitable methods for the empirical investigation and validation of the main working steps of our framework. For this purpose, we have conducted a survey that investigated which IT-phenomena would be particular interesting for students. Additionally, we asked about 100 students to describe their conceptions and beliefs about the functions and principles of computer viruses as an example for students’ conceptions of a phenomenon.

Categories and Subject Descriptors
K.3.2 [Computer and Information Science Education]: Computer Science Education, Curriculum.

General Terms
Experimentation, Human Factors, Theory.

Keywords

1. INTRODUCTION
Currently, one of the major concerns of researchers in traditional school subjects is to motivate teachers to take advantage of empirical results for their everyday work.

As stated by Duit [8], p.5: “It is common sense among science educators that improving practice is the primary aim of science education research”.

In order to support this aim, Kattmann, Duit, Gropengiesser and Komorek [17] have presented their framework of Educational Reconstruction that was designed to support the transfer of knowledge from research to the classroom: “It draws on the need to bring science content related issues and educational issues into balance when teaching and learning sequences are designed that aim at the improvement of understanding science and hence may foster the development of sufficient levels of scientific literacy.”[8], p.5.

Computer Science (CS) is a very young subject area in schools, compared to (Foreign) Languages, Mathematics or Physics. As for every “adolescent”, it is advisable for Computer Science Education (CSE) to look at the “adults”, in order to learn as much as possible from them, particularly from closely related subjects like mathematics or science. As in any other school subject, everyday lessons in Computer Science should be developed based on the results of solid research. Nevertheless, despite the similarities e.g. with mathematics or science, CS is different from all other traditional subjects in some respects, e.g. in curriculum issues or teacher education, see section 3.3, [7] or [15]. Therefore it might be problematic to apply educational models from other disciplines in CSE without taking these differences into account.

In this paper we propose a specific model for Educational Reconstruction for CSE as a framework that was derived from the established model of Duit [8] for science education. Our model was published in German language already [6].

Reflecting the specific situation and needs of CSE, our model arranges four issues that are relevant for the design and arrangement of CS lessons and courses around the central sector of CS Phenomena:

- the clarification and analysis of the science content,
- the analysis of the social demands on the subject,
- the investigation of the students’ perspectives on a topic,
- the investigation of the teachers’ perspective.

In this article we describe the derivation of the model, starting from the model for (natural) sciences of [8]. We discuss the model and identify the need for research in each part of it. In order to demonstrate exemplarily, how research results might be utilized by teachers to improve their lessons, we present several surveys. To illustrate the investigation of teacher perspectives, we refer to our study among 400 teachers from 2009. Concerning the choice of suitable phenomena, we have conducted a survey that was carried out by
students in the context of a teacher education lecture. The same students have interviewed about 100 persons in order to investigate students’ conceptions of computer viruses.

2. THEORETICAL BACKGROUND

2.1 Educational Reconstruction for Science Education

The starting point of our deliberations was set by the model of Kattmann et al. who had published it originally in German [20], but later also in English [18]. They had developed their concept of Educational Reconstruction (originally in German: “Didaktische Rekonstruktion”) as a framework for research and development in Science education. The key idea was that the science content of a teaching unit is not given, but has to be reconstructed according to the perspectives of the students as well as correspondent to the structure of the science content, which was represented by a triangle of three components that are displayed in fig. 1.

Additionally Duit et al. connect this core element with Research on teaching and learning that should reveal the perspectives of the learners as well as the teachers’ views and conceptions and also investigate the teaching and learning processes. The third element of the model is formed by the development and evaluation of (pilot) instruction, covering issues of real teaching and learning environments.

Duit et al. confirm [8], p. 6: “It is a key claim of the Didaktik tradition that both processes ‘elementarization’ and ‘construction of the content structure for instruction’ are intimately interrelated to decisions on the aims of teaching the content and the students’ cognitive and affective perspectives. These perspectives include students’ pre-instructional conceptions and their general cognitive abilities on the one hand and their interests, self-concepts and their attitudes on the other.”

In section 3 we will derive our specific framework for CSE starting from this model.

Figure 1: The Model of Educational Reconstruction [18]

As proposed by Kattmann et al. [17], the teachers should perform the “design of learning environments” after iterating the two first steps “investigation into students perspectives” and “clarification of science content”, aiming to adopt subject matter knowledge as presented in textbooks or other scientific publications to the perspectives of the students in such a way that suitable teaching content could be constructed. Each of the three components was regarded as equally important.

2.2 Combining Analysis, Research and Development

The original model of Kattmann et al. [17] was enriched and refined by Duit et al. in 2005 [9] and described very comprehensively and extensively by the same author [8] (see Figure 2). We will refer to this model shortly as the “Duit Model” in the following sections. The core element of this refined model is formed by the analysis of the content structure that comprises the subject matter clarification and the analysis of educational significance. It should be performed in two steps:

1. the elementarization, resulting in the identification of the relevant elementary ideas of the content and

2. the construction of content structure for instruction.

The outcome of the second step is not only a simplification of the science content, but also an enrichment of this by putting it into contexts that make sense for the learners [8], p. 6.

Figure 2. Educational Reconstruction according Duit et al. [8]

2.3 Context and Phenomena

Following similar approaches in other school subjects, e.g. for Chemical Education [25], [11], the “CS in Context” approach is discussed currently in the international CSE community [3]. Particularly in Germany, many colleagues are involved in the elaboration of this approach (e.g. www.informatik-im-kontext.de, [5], [7]). Nevertheless, it is not the intention of this paper to introduce this approach. But, as the terms Phenomenon and Context are applied quite generally in everyday speech and thus might have very different meanings, we have to determine how we understand these terms in this article. For that purpose, we refer to the Oxford Dictionary, which defines (oxforddictionaries.com):

- Context: “the circumstances that form the setting for an event, statement, or idea, and in terms of which it can be fully understood.”
- **Phenomenon**: “a fact or situation that is observed to exist or happen, especially one whose cause or explanation is in question.”

For chemical education, Gilbert [11], pp. 96 ff presented four models of “Context as the direct application of concepts, as reciprocity between concepts and applications, as provided by personal mental activity and as the social circumstances”.

Concerning the relationship between Context and Phenomenon, we join Purchmann, who (implicitly) suggested [25] that Phenomena could be regarded as parts of a specific Context: “For the phenomena taught during the context-based unit”.

Already in 2004, Humbert and Puhlmann had defined Phenomena in CSE as “occurrences of informatics in the world” [16], p. 66. Additionally, they had proposed a very useful categorization of Phenomena with regard to CS (synonymous to “Informatics” here) [16], p. 68: “We therefore distinguish the following types of phenomena of informatics:

1. Phenomena that are directly related to informatics systems. They occur when a person consciously uses an informatics system, such as a word processor [...].

2. Phenomena that are indirectly linked with informatics systems. They occur in everyday situations whenever informatics systems are involved without being apparent at first glance [...].

3. Phenomena that are not connected to informatics systems but have an inherent informatical structure or suggest informational reasoning [...].”

We will refer to this categorization in the discussion of our research results in section 4.

### 3. Educational Reconstruction for CSE

At this point we have to recall that the two models of Educational Reconstruction described above (see section 2) were developed for (natural) science education, where it is self-evident to arrange lessons around experiments and phenomena. It is also accepted that biology, chemistry and physics should be taught in compulsory courses for over a century now. Additionally, there is a strong consensus about the core content of the curricula of these subjects. The teachers are educated according to a tradition and to curricula that are at least similar among different countries or states. For CSE all this does not apply. So we have to take a closer look at the discussed models in order to adopt them for CSE.

In our opinion, in the Duit Model [8] several issues that are particularly important for CSE are missing, respectively not represented in a way that meets this importance. Therefore we present a specific model for CSE that was derived by reification, rearrangement and extension from the Duit Model (see fig. 2). In the following we will present the rationale for these enhancements that result in the model that is displayed in fig. 3.

#### 3.1 Phenomena and Educational Significance

Regarding the overall structure of our model, we preferred the view of Kattmann et al. [17], who had constructed their model as a triangle of planning and analyzing activities (see section 2.1), to the view of Duit at al. [8] as a combination of analysis, research and development (see fig. 2). Nevertheless, the model of Kattmann et al. missed a very important step in our opinion, as there is a broad agreement currently within CSE that teaching units should start from a “real-world” context or phenomenon, aiming to motivate the students, to open connections to prior knowledge or to show application situations of the intended knowledge. Duit [8] states on p. 6: “The science content structure for a certain topic (like the force concept) […] has not only to be simplified (in order to make it accessible for students) but also enriched by putting it into contexts that make sense for the learners”. Nevertheless, he doesn’t display neither context nor phenomena in the graphical representation of the model (see fig. 2). This requirement could be founded by the postulation of authenticity by constructivist approaches (for a discussion of constructivism in CSE see Ben-Ari [19]) as well as by the computer science in context approach that was discussed recently very intensely [3], particularly in Germany (see e.g. www.informatik-im-kontext.de, [5], [7]).

![Figure 3. Educational Reconstruction for Computer Science Education](image-url)

Already in 1969 Klafki pointed out in his five key questions of educational analysis: “What is the more general idea that is represented by the content of interest? What is the phenomena or basic principle, …” and: “What are particular cases, phenomena, situations, experiments that allow making the structure of the referring content interesting, worth questioning, accessible, and understandable for students?” [8], p. 7. Students have contact with phenomena that are related to information technology every day, oartly without even apprehending those. In our opinion it is one of the main challenges to CSE to reveal the relevance of CS for these phenomena and to explain them by concepts and principles of CS. Thus we decided to place the process Choice of Phenomenon in the center of our model as shown in fig. 3.

Another issue that we regarded very important particularly for CSE was the educational significance of the intended learning content. The Duit Model comprises this aspect in its part 1 (see fig. 2), which has to be founded by normative demands that are set by the society that represents the “funder”, the “customer” and the “principal” of educational processes at schools. As these demands might differ substantially among different countries, states or even school types in the same country, the educational significance of a certain topic might also vary heavily under different circumstances. Therefore, it is a very important step of planning lessons to assess the educational significance of a certain topic respecting the specific circumstances of this lesson. Consequently, we placed as second process Analysis of Social Demands beside the first one.

#### 3.2 Science and Students Perspectives

Following this, we included two more nodes that respected the two activities that have been well-founded by the original model of
Kattmann [17]. The first of these, the Clarification of Science Content Structure, contains the process of elementization that was described [8] and therefore might identify relevant elementary ideas of CS that are discussed currently by the “CC Principles Approach” [1]. The clarification and analysis of science content reveals how the phenomenon could be explained scientifically and which subject domain knowledge is required to understand the phenomenon. The result determines which concepts of CS have to be taught in the lesson.

The second activity that we adopted from the model of Kattman et al. [17] was Investigation of Students’ Perspectives, referred as “Perspectives of the learners (views and perspectives & affective variables)” by [8]. If at all, those perspectives are often regarded as a motivational factor only, instead of being utilized for the clarification of the science content also. Our framework intends to overcome the gap between students’ perspectives on a certain topic and the respective CS lessons, respecting them during all steps of the planning process. Investigation of students’ perspectives means for us to answer the questions:

1. How do the learners experience the chosen phenomena in everyday life or in class? Are the phenomena attractive, accessible and important for all of them, especially equally for boys and girls or for students of many different situations at home?
2. Which concepts do the students use to explain those phenomena?

As a consequence of respecting the students’ perspective we cannot talk of misconceptions, but only about students’ conceptions. The term “mis” already suggests that the students’ conceptions would be wrong or not worthwhile and thus have to be replaced in their minds by something that is more “correct” in any respect. In contrary, according to our model the students’ views have to be respected always. They are regarded as comparably important as the “official” scientific view, fitting it correctly or not.

3.3 Teachers Views and Lesson Design

Another very special issue in CSE is teacher education. As there is no compulsory subject of CS in many countries or states, the educational background of the CS teachers might be very different. Some of them might have a regular university degree in CS (M.Sc.), being hired after a short introduction course to teaching practice. Some others have a specific teachers’ degree (M.Ed.), including courses in pedagogy and psychology. The latter might have been acquired by regular studies or by specific in-ervice training courses on very different demand levels, from the pure application of standard software to a solid theoretical background. Some of the teachers might not have attended any specific courses in CS at all [15], [7]. Consequently, the teachers will differ substantially regarding their subject domain knowledge, their pedagogical knowledge and their pedagogical content knowledge (see [26]). Therefore they might work on very different abstraction levels or applying very different teaching methods for the same topic of the curriculum. Thus, from our point of view, the individual abilities of the respective teacher and in consequence, his or her perspectives on a certain topic have to be taken into account very closely. We regard the teachers’ perspective as a key factor for the design of lessons as well as for educational research, which we represented by the node Investigation of Teachers’ Perspectives. In detail this means to answer the following questions:

1. Which conceptions do the teachers have regarding lesson design and teaching CS in general?
2. What do they think which conceptions their students might have to explain the chosen phenomena?
3. Which conceptions do the teachers actually apply to explain the chosen phenomena themselves?

If respected properly, the answers to these questions might also support the acceptance of results of proposals, are suggested by researchers or other educational professionals, by the average teacher.

Finally, we included the resulting process of the design of the lesson. But, as we follow the constructivist approach as far as possible, we did not want to restrict this design to the learning content, but wanted to stress the need to plan also teaching methods, media or the organization of the lesson. Therefore we called this part Design and Arrangement of CS Lessons and Courses. We propose to apply the well-approved Berlin Model of Heimann at this point in order to have a checklist of the most important aspects of lesson design. According to this model the planning process has to cover four different decision areas: intentions (objectives, competencies, outcomes), content (topics, knowledge), teaching methods (organization, cooperation, sequencing) and media (textbooks, tools, hard- and software, programming languages). A close discussion of this model with regard to CSE is found in [15].

3.4 How to read the Resulting Model

In summary, our proposed framework of Educational Reconstruction for Computer Science Education consists of 6 parts which are displayed in fig. 3.

The arrows symbolize the direction of the influences between the parts of the model. For example: The clarification of the social demands on the subject has consequences for the design of CS lessons, but not vice versa. The choice of CS phenomena that forms the teaching context determines the clarification of science content (including the explanation of the phenomena) as well as the clarification of social demands (regarding the chosen phenomenon).

4. RESEARCH EXAMPLES

In this section we present some exemplary research projects that are suggested by our framework. Please note that we are not aiming to do any systematic validation of our model at this point, rather to illustrate some of its crucial components.

Regarding the six components of our model, we will cover only three of them in detail here. We will not discuss research aspects regarding the remaining three parts out of different reasons:

- The clarification and analysis of the science content is covered by many publications that deal with specific phenomena, applying different methodologies and leading to a variety of results.
- The analysis of the social demands on the subject depends heavily from the specific society and educational context and will differ from country to country. The analysis could include interviews with persons that represent the “customers” of educational processes (e.g. university lecturers, supervisors of vocational training or industry representatives) as well as the investigation of constitutions, laws or other legal prescriptions or preambles of curricula. Additionally, standards that are proposed or already set could also very relevant for this aspect, e.g. the proposed CSE standards by the German Gesellschaft für Informatik [10] or the CSTA standards for K-12 in the US [27].
4.1 Contexts and Phenomena

Aiming to stimulate possible research projects and to test suitable methodologies, we have conducted two surveys that were performed by the 28 students of an introductory teacher education course (called “interviewers” in the following) at our university. This was due to our opinion that teacher students should participate in “real” research as soon as possible, as postulated also by Hassanz et al. [12]. Please note that data were gathered in German language.

4.1.1 Data Gathering

The first survey intended to find out, which phenomena would be particularly interesting. The interviewers should ask three persons each to answer some questions, which asked about the personal and CSE background first and then posed the following central question: “Please give three phenomena, processes or operations relating to the everyday use of information and communication technology, which you regard as particularly complicated, incomprehensible or difficult to understand”.

The interviewers were asked to look for school students preferably and, as far as possible, for persons without any background in CS. At the end they had interviewed 86 persons, 39.5% of them being female, 59.3% being male (1 person without statement to gender). We categorized the test persons according

1) gender in order to assess how representative the survey was,
2) professional category: school students (S-students), higher education students (HE-students), workers, employees, employees of the state (EoS) and others,
3) educational background in CS: none or informal courses (G1), regular CS or IT courses in school (G2) and regular CS courses in higher education (G3).

This categorizing had three objectives: Firstly, we wanted to estimate roughly how representative our sample had been. Secondly, we aimed to assess, whether the answers of those groups would differ in any relevant way. Thirdly, we wanted to find out, which test persons would be the most similar to the students that are attending the computer science courses in our secondary schools [14]. Some results of the categorizing are displayed in table 1.

The overall gender ration was not very good in the survey (39.5% male, 59.3% female, 1.2% unspecified). Unfortunately, the original target group of school students was met only by less than a quarter of the test persons. Nevertheless, we regarded this fact as not very important for our research question, because the attractiveness of a phenomenon might not necessarily depend from that attribute. At the end the phenomena for the lessons should be chosen in such a way that they are regarded as attractive by the students even when they will have finished school and started their professional live.

Table 1. Distribution according educational background

<table>
<thead>
<tr>
<th>Category</th>
<th>G1</th>
<th>G2</th>
<th>G3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-students</td>
<td>7.0%</td>
<td>15.1%</td>
<td>0.0%</td>
<td>22.1%</td>
</tr>
<tr>
<td>HS-students</td>
<td>25.6%</td>
<td>5.8%</td>
<td>3.5%</td>
<td>34.9%</td>
</tr>
<tr>
<td>Workers</td>
<td>2.3%</td>
<td>2.3%</td>
<td>0.0%</td>
<td>4.7%</td>
</tr>
<tr>
<td>Employees</td>
<td>16.3%</td>
<td>7.0%</td>
<td>0.0%</td>
<td>23.3%</td>
</tr>
<tr>
<td>EoS</td>
<td>4.7%</td>
<td>0.0%</td>
<td>1.2%</td>
<td>5.8%</td>
</tr>
<tr>
<td>Others</td>
<td>8.1%</td>
<td>1.2%</td>
<td>0.0%</td>
<td>9.3%</td>
</tr>
<tr>
<td>Total</td>
<td>64.0%</td>
<td>31.4%</td>
<td>4.7%</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Analyzing the educational background in CS (see table 1), we observed that more then 2/3 of the school students had a formal background in CS, which was plausible in the light of the fact that two of the three types of secondary schools in our state comprise a compulsory subject of CS (Gymnasium) respectively IT (Real-schule) and as these two school types are attended by about 70% of each age group (neglecting the fact that the Gymnasium take 2 more years). In contrary only a small part of the persons in professional life should have such a background currently, because the compulsory subject had not been introduced before 2004. In total, about 2/3 of the interviewed persons have no or only an informal background in CS [15].

The answers of the test persons were very different. The length differed from 0-40 words, typically the length was about 10-15 words. We performed a qualitative text analysis according to the method of Mayring [22]. At first we coded all interviews. Afterwards we categorized the codes in four abstraction levels from level 1 up to the top level. As some students gave very abstract answers (e.g. “Internet” or “Programming”), thus even the top level categories had several direct codings. At the end we had produced 259 codings that used 198 hierarchically arranged categories.

4.1.2 Interesting Contexts

In order to find the most interesting fields of phenomena, we had to analyse the number of codings of the top level categories (see table 2). Due to their high level of abstraction, these categories represent sets of Phenomena and therefore could be regarded as exemplary Contexts (see section 2.3). On the top level we found that the most interesting top categories for the test persons were: Software Systems, Computer Systems and Internet. Table 2 lists all top level categories that had more than 10 codings over all interviews and illustrates those by giving the subcategories also.

Table 2. Codings of the most interesting top level categories

<table>
<thead>
<tr>
<th>Top level categories with subcategories</th>
<th>Codings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software systems: symbol representation, user control, special functionality, video, text processors, copy and paste, visual effects, drag &amp; drop, office software, presentation software, spreadsheet programs, help systems, file format conversion, auto features, special software systems, browsers, robotics, GPS, online banking, photo software, search engines, auto-completion, parallel processing, version control, updates, installation</td>
<td>63</td>
</tr>
<tr>
<td>Computer systems: server configuration, operation, networking, malfunctions, excessive demands, crashes, peripherals, printing, formatting of hard drives, removable media, interaction with printers, scanners, beamers, video ports, screen, touch screen</td>
<td>34</td>
</tr>
</tbody>
</table>
storage without electricity, slowdown and configuration of PCs.

<table>
<thead>
<tr>
<th>Category</th>
<th>Code</th>
<th>Level</th>
<th>Codings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet: search engines, Internet TV, connection setup, operation, social networks, disconnection, online shopping, exchange of data, chat, download, browser, creation of Web pages, Web site programming, Web sites for mobile phone, legality of websites, data transmission on the Internet, cookies, access to Web sites.</td>
<td>Programming</td>
<td>Top</td>
<td>7</td>
</tr>
<tr>
<td>Data transfer: communication over large distances, FTP, router, wireless data transmission, mobile Internet, Wi-Fi, W-LAN, Backup, Bluetooth, speed of data transfer.</td>
<td>Wireless data transfer</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Operating systems: Processes, crashes, shut down, compatibility, languages, accounts, error messages, complexity, installing programs, speed of startup, administrator role.</td>
<td>MS Excel</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Mobile phones: SMS (data loss, speed, transfer time), complex operating instructions, iPhone, tariff systems, computer update, radio holes, crashes, mobile services, software.</td>
<td>Slow-down of PCs</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Programming: test programming, syntax, hacking, GUI programming, programming languages, games programming, classes, methods.</td>
<td>Internet</td>
<td>Top</td>
<td>4</td>
</tr>
<tr>
<td>Data protection and data security: access to data, storage of personal data, use of personal information, personalized advertising, cryptography, viruses (function of antivirus programs, protection), security of Internet platforms, privacy policy</td>
<td>Mobile phone usage</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Special devices: MP3 player, ticket tellers, telephones, time control machines, digital and analog television, 3D TV, car radio, headset</td>
<td>Software updates</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

Please note that the figures in table 3 give the number of direct codings of these categories, e.g. Internet, while those in table 2 represent the sum of all direct codings together with all codings of the respective sub-categories, e.g. Internet, search engines, Internet TV etc. Despite the direct coding, it is apparent that most of these categories (except MS Excel and Slow-down of PCs) are still very abstract. On the other hand, MS Excel is an office software and not a phenomenon at all. Therefore, the only concrete Phenomenon that the test persons had addressed was Slow-down of PCs.

In Summary, due to the approximate equal distribution of the direct codings over the categories, the high abstraction level of the given answers and the fact that most of the referred phenomena belong to type 1 according Humbert and Puhlmann [16], we have to conclude that our interview method was not optimal for this purpose. We still continue our research activities by gathering open interviews according to the Grounded theory [4], therefore.

4.2 Students’ Perspectives

Originally we had intended to investigate the students perspectives regarding a phenomenon that would be particularly interesting, according to the study we have presented in the preceding section. Unfortunately, the only result of this study that was concrete enough to be called a phenomenon (Slow-down of PCs, see above) seemed not suitable for the circumstances of this second survey. Similar to the first study, it should be carried out by the students of a teacher education course and, additionally, evaluated by a students’ thesis. Thus we regarded the “most interesting” phenomenon (Slow-down of PCs) as too complex to be investigated under these conditions. Therefore, we went down the list of phenomena that were directly addressed by the test persons in the first survey and ended up at computer virus, which we had found in three interviews. This was selected, because it is a specific, not too complex phenomenon and most people might probably have an idea of it.

Thus we chose the two research questions:

1) Which conception of the phenomenon Computer Virus do the test persons show?

2) How does the educational background in CS influence the conception on computer viruses?

The objective of the second question was to assess, how precise the respective educational context has to be taken into account when investigating the students perspectives. In other words, could research results be relevant for the preparation of a certain lesson in any way if they were obtained in a different context?

For this purpose we asked our 28 students to interview up to 3 people each (preferably school students and persons without background in CS). The interviewees should answer our second questionnaire that looked exactly like the first one except the central question: What do you think how computer viruses work?

The students collected a total of 130 answers to this question in January of 2012, partly by asking friends via Facebook, partly by personal interviews. Again we classified the educational background in CS of the test persons by the categories G1, G2, G3 (see section 4.1), having the same objectives as described there.

As shown in table 4, among the respondents were 50% students in higher or vocational education (HE) and only 10% school (S-) Students, which was quite disappointing. Nevertheless, we regarded this deficit as not too problematic, because at the end the perspective on viruses should depend strongly from the specific personal knowledge, which we would not be able to assess anyway, S-student or not.
Unfortunately, the female test persons had a weaker background in CS compared to the males. While there were about double as many females with no or no formal background compared to males, the male/female ratio in G3 (HE/vocational background) was 4:2.

Table 4. Distribution of respondents among the three groups.

<table>
<thead>
<tr>
<th>Educational background in CS</th>
<th>total</th>
<th>male</th>
<th>female</th>
<th>S-students</th>
<th>HS-students</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1 none / informal</td>
<td>50.8%</td>
<td>18.5%</td>
<td>32.3%</td>
<td>3.8%</td>
<td>21.5%</td>
</tr>
<tr>
<td>G2 school</td>
<td>29.2%</td>
<td>15.4%</td>
<td>13.8%</td>
<td>7.7%</td>
<td>13.8%</td>
</tr>
<tr>
<td>G3 HE studies / voc. Training</td>
<td>20.0%</td>
<td>16.2%</td>
<td>3.8%</td>
<td>0.0%</td>
<td>16.2%</td>
</tr>
<tr>
<td>total</td>
<td>100.0%</td>
<td>50.0%</td>
<td>50.0%</td>
<td>11.5%</td>
<td>51.5%</td>
</tr>
</tbody>
</table>

Subsequently, all interviews were coded according to the method of qualitative content analysis described by Philip Mayring [22]. For this purpose, the software MAXQDA was used (see www.maxqda.com). The coding resulted in 17 (partly overlapping) categories to which all interview statements could be assigned (see table 6).

Table 6. The coded categories.

<table>
<thead>
<tr>
<th>Category</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 Spying</td>
<td>virus passes data (e.g. passwords) to third parties</td>
</tr>
<tr>
<td>C2 Hardware damage</td>
<td>virus damages the hardware</td>
</tr>
<tr>
<td>C3 Infection</td>
<td>virus enters the computer from &quot;the outer world&quot;</td>
</tr>
<tr>
<td>C4 Infection by e-mail</td>
<td>virus enters the computer via e-mail</td>
</tr>
<tr>
<td>C5 Infection via the Internet</td>
<td>virus enters the computer through the Internet</td>
</tr>
<tr>
<td>C6 no idea</td>
<td>used only when no other codable content was available</td>
</tr>
<tr>
<td>C7 Control</td>
<td>virus allows third parties to control the computer</td>
</tr>
<tr>
<td>C8 Disease</td>
<td>analogy with biological or viral diseases</td>
</tr>
<tr>
<td>C9 Delete</td>
<td>virus deletes data or makes it unusable in another way</td>
</tr>
<tr>
<td>C10 Program</td>
<td>virus is also a computer program</td>
</tr>
<tr>
<td>C11 Risk Factors</td>
<td>factors that facilitate infection or damage by the virus (e.g. &quot;ignorance of the user&quot;)</td>
</tr>
<tr>
<td>C12 Damage</td>
<td>virus damages the computer in an abstract way</td>
</tr>
<tr>
<td>C13 Protection</td>
<td>there are ways to protect themselves against infection or injury caused by the virus</td>
</tr>
<tr>
<td>C14 Protection software</td>
<td>Software exists to protect against infection or injury caused by the virus</td>
</tr>
<tr>
<td>C15 Camouflage</td>
<td>the virus is not easily identifiable (or threat of damage can not be recognized)</td>
</tr>
<tr>
<td>C16 different types</td>
<td>there are various kinds of malicious programs (e.g. various damage scenarios, but also differentiation of virus, trojans, etc.)</td>
</tr>
<tr>
<td>C17 Proliferation</td>
<td>the virus passes (independently) from an infected machine to another</td>
</tr>
</tbody>
</table>

Table 5. Relative and absolute frequencies of codings of the categories.

<table>
<thead>
<tr>
<th>Category</th>
<th>Relative Frequencies %</th>
<th>Number of Codings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damage</td>
<td>G1 57.58</td>
<td>52.63 46.15 53.8</td>
</tr>
<tr>
<td>Infection</td>
<td>G2 45.45</td>
<td>63.16 42.31 50.0</td>
</tr>
<tr>
<td>Program</td>
<td>G3 22.73</td>
<td>31.58 53.85 31.5</td>
</tr>
<tr>
<td>Spying</td>
<td>total 75.64</td>
<td>41.24 25.4 33</td>
</tr>
<tr>
<td>Infection by e-mail</td>
<td>G1 16.67</td>
<td>18.42 3.85 14.6</td>
</tr>
<tr>
<td>Delete</td>
<td>G2 16.67</td>
<td>15.79 7.69 14.6</td>
</tr>
<tr>
<td>Proliferation</td>
<td>G3 7.58</td>
<td>21.05 19.23 13.8</td>
</tr>
<tr>
<td>Control</td>
<td>total 12.12</td>
<td>7.89 15.38 11.5</td>
</tr>
<tr>
<td>Risk Factors</td>
<td>G1 3.03</td>
<td>21.05 11.54 10.0</td>
</tr>
<tr>
<td>Disease</td>
<td>G2 7.58</td>
<td>10.53 7.69 8.5</td>
</tr>
<tr>
<td>Protection</td>
<td>G3 7.58</td>
<td>10.53 7.69 8.5</td>
</tr>
<tr>
<td>different types</td>
<td>total 4.55</td>
<td>10.53 15.38 8.5</td>
</tr>
<tr>
<td>Infection via the Internet</td>
<td>G1 6.06</td>
<td>13.16 3.85 7.7</td>
</tr>
<tr>
<td>Protection software</td>
<td>G2 7.58</td>
<td>7.89 7.69 7.7</td>
</tr>
<tr>
<td>no idea</td>
<td>G3 10.61</td>
<td>5.26 6.9 9</td>
</tr>
<tr>
<td>Camouflage</td>
<td>total 6.06</td>
<td>7.89 7.69 6.9</td>
</tr>
<tr>
<td>Hardware damage</td>
<td>G1 1.52</td>
<td>0 7.69 2.3</td>
</tr>
</tbody>
</table>

The first apparent result of the data analysis was that the most addressed categories were Damage and Infection, addressed in more than 50% of the interviews, followed by Program and Spying (more than 25%). Apparently there is a quite strong awareness of the risks, but a weak understanding of the working principles of viruses.

In order to compare the conceptions of the different groups G1, G2, G3 visually, we constructed a specific type of modal concept maps. The coded categories should represent the nodes, and an edge between two nodes should be drawn if a sufficient number of interviews has addressed the categories of those nodes together. The assumption was that by addressing both of these categories, the interviewee would suggest an association between those. The more interviews would address both nodes, the thicker the edge should be drawn. Finally, the relative frequency of codings (see table 6) should define the weight of each node, represented by its size. In
order not to give too much weight to individual opinions and because one could name more than one category, only those edges are shown in the following figures that were above the third quartile \(Q_{0.75}\) according to the relative frequency of their codings. The latter was calculated by dividing the number of codings of each category in each of the groups G1, G2, G3 by the number of collected interviews of each group. For a detailed discussion of the application of concept maps in CSE, see [14]. This resulted in three different graphs for G1, G2 and G3 (see figure 4).

The resulting concept maps suggested some differences between the groups G1 to G3, regarding the codings of the 17 categories. Aiming to give statistical evidence to these results, we chose the following hypothesis: “The average number of codings per category is equal for the three groups G1, G2, G3”. More precisely, we set the following three hypotheses \(H_0\): “The average number of codings per category is equal within the three pairs (G1, G2), (G2, G3) and (G3, G1) respectively.”

To test these \(H_0\), we applied an approximative Gauss-Test according to [2] and found, that only the following two concepts fulfilled the preconditions of the test (\(5 \leq \sum x_i \leq n-5\)) and also differed in the averages significantly at the level \(p<0.05\).

Table 6. Results of the approximative Gauss test

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Different averages</th>
<th>Test value (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1, G2</td>
<td>Proliferation</td>
<td>-2.0011*</td>
</tr>
<tr>
<td>G2, G3</td>
<td>Program</td>
<td>2.8927*</td>
</tr>
</tbody>
</table>

*Rejection range: \(V < -1.9600\) or \(V > 1.9600\)

Therefore we could reject the respecting hypotheses \(H_0\) and conclude that the average coding frequencies of these concepts would differ actually between those groups. In both cases the persons with better education in CS mentioned these concepts more frequently.

In order to reveal more differences between the test persons according to the background in CS, we joined the groups G2 (formal school education in CS) and G3 (Higher educational background) to G2-3, aiming to increase the number of test persons per group and thus the chances for significance of the differences.

Figure 4. Concept map of G1, G2, G3.
Nevertheless, this reproduced exactly the same results: At the level of \(p \leq 0.05\) the only concepts that were applied significantly different between G1 and G2-3 were Proliferation and Program again. Apparently CS education raises the awareness that viruses are programs essentially and also that they bear the danger of proliferation. In the light of the fact that the term virus does not appear in the curricula of the subjects CS respectively IT of the two relevant school types in our state (Gymnasium and Realschule) at all, this result is not surprising.

In summary, the results for our second research question suggest that the conception of viruses is not strongly depending from the educational background in CS. Therefore, it might be interesting for a teacher to take research results into account, even if those were obtained in a different educational context.

Nevertheless, there are certain differences between the groups. For example, the test persons of G2 (having attended CS courses at school) were less aware of the camouflage aspect of the viruses and did not mention protection software as of the as the other groups. This might indicate that their CS education in school had some deficits regarding these aspects.

### 4.3 Teachers Perspectives

In 2009 we have investigated the views of active teachers on the recently installed compulsory subject of informatics in the state of Bavaria [24], [13]. We had asked more that 1200 teachers via E-Mail to answer an extensive questionnaire that comprised 48 questions. We had received 448 responses, as though these were not complete in some cases. One of the questions was: “How would you rate the following items of the curriculum regarding their importance for the future (professional) life of the students?”.

The items of the curriculum were presented in a table that can be found [24]. The response items represented a 4-step ordinal scale: very important, important, rather unimportant, unimportant. We found that the teachers could be assigned to one of three different clusters, based on their answers on this question [24]:

1. “office users”, concentrating on the application of software systems,
2. “fans of the curriculum” and “anti-programmers”, preferring object-oriented modeling (OOM) instead of programming and algorithmic concepts,
3. “traditional computer scientists” that focus on traditional algorithmic views instead of OOM.

This might illustrate how different the conceptions of teachers regarding the same phenomenon, e.g. on a computer virus, might be. While the members of cluster 1 might focus on the infection by a Virus via office software, the members of cluster 2 would try to design an object oriented model of virus and the infected system, and cluster 3 could be interested in the algorithms of the virus predominantly.

In [23] we presented first results about the subjective theories of computer science teachers at secondary schools, collecting and analyzing semistructured interviews. The pretest showed that computer science teachers really had very diverse perceptions on how to design lessons on the topic networks and the Internet. As first results we can sum up: Computer science teachers have different ways in designing lessons on a given topic. They set different focuses and handle the topic differently, do not have a clear idea of the students’ perspectives, have diverse ways of becoming a computer science teacher, have various points of views on computer science compared to their other subject, often feel a lack of knowledge on computer science, think they are not a proper computer science teacher because of their (missing) qualifications.

Nevertheless, these findings have to be regarded with care because the main investigation will follow.

### 4.4 Validity of our Results

As already mentioned in the preliminary remarks of section 4, the research that was presented in this paper should serve as an illustration of suitable research questions or methodologies for the empirical evaluation of our model, but does not claim to produce results that are valid in any way, as there were many serious threats to validity. Both surveys were conducted in the context of a teacher education course, which encompasses the risk of results that were faked by the interviewers. Additionally, the circumstances of the interviews were not prescribed or recorded and might have been very different. Finally, as the categorizations of the test show, those were very different in many respects.

### 5. CONCLUSION AND FUTURE WORK

In this paper we have presented the derivation of a specific model for Educational Reconstruction for CSE that is based on the models of [17] and [9]. We have applied this model already for the design of several courses as well as a guiding principle for students’ theses. So far it has turned out to be very suitable for these purposes.

Additionally we have presented three “model surveys” that illustrate, how future research regarding three very important parts of our model (Selection of Phenomena, Teachers’ and Students’ Perspectives) could be designed, conducted and evaluated. Of course we are well-aware that these surveys can be regarded as starting point of future research only. More precisely, we intend to conduct the following research activities about these parts of our model:

Interviews with school students from different age groups and school types concerning the phenomena they regard as particularly interesting and attractive, according the Grounded theory method [4]. The focus will be on phenomena of the categories 2 and 3 according [16] that don’t have apparent origin in CS. This could result in a questionnaire for a quantitative survey. Further, we plan to conduct a second series of interviews with school students concerning their views and explanations of certain phenomena that will be selected according to the results of the first series of interviews. Thirdly, we are elaborating several contexts (“sports”, “social networks” and “intelligent house”) exemplarily in order to get more information about the requirements that have to be met by “good” contexts.

Concerning the teachers’ conceptions, we are going to focus more on their knowledge about students’ perspectives on the topic of networks and the Internet. This might give us more detailed answers about when and where teachers think that students have difficulties in understanding this topic. We are going to confront the teachers with a specific perspective of students such as that students think the Internet consists of only one big computer and ask how he would react if this view occurs in his lessons. We hope that his paper will stimulate other researchers to join these activities by similar or (totally different) surveys.

### 6. ACKNOWLEDGMENTS

We have to thank the students of our teacher course for gathering the data of the two described surveys.

### 7. REFERENCES