Abstract of and comment on
‘Levels of Abstraction in Students’ Understanding of the Concept of Algorithm: the Qualitative Perspective’
by Perrenet and Kaasenbrood

This study is a follow-up of an investigation into the levels of thinking of computer science students. A summary and reflective essay of that investigation can be found at AREE (Perrenet, 2006). For better understanding of the following, the reader is advised to consult those texts first. In the first study we defined four levels of abstraction in students’ thinking about the concept of algorithm: the execution level, the program level, the object level and problem level. We constructed a questionnaire about algorithms to measure the answering level as an indication for the thinking level. Most of its items had the format of a proposition with the possibility to agree or disagree (or even both) followed by the obligation to give an argumentation for the choice made. It was mainly the supporting argumentation and not the choice of the answer that was used for further analysis. A scoring system was developed to determine the answering level as indication of the student’s thinking level. The questionnaire was presented several times to Bachelor students of three year groups. For details see Perrenet (2006). The main findings were as follows:

- The argumentations as measured with the constructed instrument were mainly on level 2 and 3, a few on 1 or 4.
- Within a students’ year group the answering level generally increased during the year.
- For successive year groups the answering level was generally higher.
- The reliability of the instrument was good (the scores given by several raters correlated well enough).

An important question remained after the first study, the question about validity: did we really measure abstraction level of thinking? The argumentations we analyzed consisted of a few lines of text. Could it not be that students sometimes only reproduced standard definitions, giving the false impression that they really did understand the terms used? It is this kind of question that typically asks for qualitative research, because it is the thinking process one is interested in. In the first study mainly quantitative research methods were used. This completion with a more qualitative study is in accordance with the pledge by Almstrum et al. (2005) for diversification of the research methods in CSE. Both kinds are necessary to provide a full picture of what is going on.

In the qualitative study an analysis was done on a large amount of data from only a few students: nine students from different bachelor years and different ability. These students were asked to complete the questionnaire, while thinking aloud, and subsequently were interviewed about their answers. The main goal was to investigate to what extent the students really understood the computer science terms they had used in their written answers.

In total the use of 34 computer science terms was analyzed, varying from terms like ‘algorithm’ and ‘implementation’ (used by all of them due to their appearance in the items of the questionnaire) to terms like ‘greedy’ and ‘upper bound’ only individually used. The most important result was that generally the students appeared to understand the terms they used. In only about 5% it was clear that their understanding was insufficient; in about 75% of the cases their understanding of the term is evaluated as (reasonably) well. In 20% it did not become clear within the given interview time.

From this result it can be concluded that not only the reliability but also the validity of our instrument to measure the abstraction level of thinking about the concept of algorithm has proved to be satisfactory.
The richness of qualitative data

We had quite a struggle in our process of data analysis. Analysis of qualitative data can be very time consuming and it is important to focus on a productive and efficient approach. In the end we had more data than we really needed and even the data we used gave us more information than expected.

In the reports about our study, including the AREE summary, a certain subtask with our subjects is not mentioned. The reason is that we did not use its data. It was a small set of algorithm-construction problems. The idea was that these tasks would give extra information about the level of understanding of the students. We had decided beforehand that the interviewer could choose to bring in these problems when the need was felt to get more clarity about the level of understanding. Afterwards it was clear that it would have been better to use these problems in a more standardized way, because it was difficult to draw conclusions when not all students had done the same tasks. So, after some time consuming initial analysis of the problem protocols and the accompanying written notes, the decision was made - also for efficiency reasons - to discard the data of this subtask.

In the process of analysis we even turned into a second blind alley. This time the idea was to determine our subjects’ level of thinking from their protocols and compare this level with the level determined from the written argumentations. Again, after some time consuming analysis, we had to halt this undertaking. Our scoring system had been developed for the students’ written argumentations and could not handle the more complex set of interview data.

In the end we confined ourselves to only one of the plans we started with: simply looking at the students’ understanding of the concepts used in the interview. This kind of analysis is comparable with the standard task of teachers when students take an oral or written exam; it could be done quite efficiently and gave us satisfactory results.

In hindsight we can conclude that an early investment of time in thinking about what kind of data we were looking for, what kind of analysis would be possible and – more fundamentally – what kind of validity questions we wanted to answer, would have saved us time later on.

In conclusion, what surprised us the most during the process of analysis was the richness of these qualitative data. Students of the same year group showed very different ways of thinking about the same concepts. They used different kinds of argumentations for the same answers or even used comparable argumentations to come to different answers. As a teacher one should not to easily turn to the multiple-choice test format where all these thought processes stay hidden. In the construction process the items should be tested on students in interview sessions or one should ask the students for short argumentations for their choices.
