article:2411

A Multi-National Study of Reading and Tracing Skills in Novice Programmers

This paper was the report of a working group which convened as part of the ITiCSE 2004 conference (i.e. the conference on "*Innovation and Technology in Computer Science Education*). The ITiCSE 2004 conference was held in Leeds, and as there was only one working group that year, this working group is known as "The Leeds Group". Three years earlier, at ITiCSE 2001, the "McCracken Group" (so named after its leader, Mike McCracken) had assessed the programming ability of a large population of students from four universities in two countries, the United States and the United Kingdom. The students were asked to solve a given programming task. The majority of students performed much more poorly than expected. The McCracken group concluded that the students were weak at problem-solving.

The starting point for the Leeds Group was to question the conclusion drawn by the Leeds Group – that the students performed poorly because of a weakness in problem solving. There were other, simpler explanations that were not eliminated by the experimental design of the McCracken group. For example, perhaps the students did not have a strong grasp of the programming constructs they needed to use to produce a solution? The Leeds Group looked for simpler explanations of the McCracken Group's result by testing over 500 students (from seven countries) on their ability to read and comprehend code. The data for the Leeds group was collected by asking students to answer twelve Multiple Choice Questions (MCQs). These twelve MCQs were of two types. One type of question tested students on their ability to select, from the four options given, the outcome of executing a short piece of code. For the other type of question, students were given pieces of code with one or two lines missing. The students were told what the complete code should do, and their task was to select the correct missing line(s) from the four options given. Most students performed poorly on these two types of tasks, indicating that the conclusion of the McCracken Group (i.e. that students were weak at problem-solving) may not be the simplest explanation of the McCracken Group's results.

Apart from recording how many questions each student answered correctly, the Leeds Group also collected two other types of data. One of these other types of data consisted of transcripts. Thirty seven students were asked to "think out loud" as they answered the twelve MCQs. These students were recorded, and the tapes were transcribed. The other type of data was referred to as "doodles". This data was the annotations students made as they attempted to answer these questions. Both of these forms of data showed that students used simplistic approaches to answering the questions; approaches that would not equip them to write code of similar complexity.

References

[1] McCracken, M., V. Almstrum, D. Diaz, M. Guzdial, D. Hagen, Y. Kolikant, C. Laxer, L. Thomas, I. Utting, T. Wilusz, (2001) A Multi-National, Multi Institutional Study of Assessment of Programming Skills of First-year CS Students, SIGCSE Bulletin, 33(4). pp 125-140.

Author 1: Raymond Lister; raymond@it.uts.edu.au
Author 2: Elizabeth Adams; adamses@jmu.edu
Author 3: Sue Fitzgerald; Sue Fitzgerald

Article Link: http://portal.acm.org/citation.cfm?doid=1041624.1041673