An instrument for evaluating problem solving, inquiry and programming skills in the context of robotics education

Heilo Altin[1] Margus Pedaste[2]

[1]University of Tartu, Faculty of Science and Technology
[2]University of Tartu, Faculty of Social Sciences and Education

Abstract

Educational robotics has been mostly practiced as an extracurricular activity. There are few studies evaluating the effect of robotics education on students' improvement. For such studies an instrument for evaluating study effect in the context of robotics is needed. We developed and tested such an instrument. The results showed that it is at an appropriate difficulty level and two versions of the instrument can be applied in pre- and post-test studies. The instrument evaluates students' skills to solve decision making and troubleshooting problems, inquiry skills and programming skills in the context of robotics education.

Keywords: educational robotics, inquiry skills, problem solving skills, programming skills

1. Introduction

Robotics has been applied in the context of ICT (Information and Communications Technology) education for more than a decade. The first widespread educational robotic kit was LEGO RCX; however, the ideas and methodology for teaching with robots are far older than the hardware itself [1]. Since the release of RCX, there have not been many substantial studies about the effect of using robots in education. Most results published are based on qualitative feedback from teachers and students [2, 3]. There are a few examples, like a Swedish project [4], where students' math skills have been tested quantitatively while using robots. The authors concluded that it is difficult to claim that robots help students better understand and improve their skills in math. They mentioned a need for more longitudinal and deeper studies. Each study in the field of robotics education is centered on specific skills and knowledge. However, our hypothesis is that robotics may empower students more in their cognitive development than current research suggests. Based on the relevant literature, there can be at least three types of skills that have to be evaluated when analyzing the effect of using robots in education. They are problem solving skills [5], inquiry skills [6], and programming skills [7]. Problem solving methodology is often used when learning with robots. Teachers present to students a problem to be solved with an ICT tool (e.g. a robotics kit). In this case, the outcomes of the project are not as important as the process itself. Testing ideas, debugging the program and troubleshooting the hardware are processes through which students learn most effectively [8]. There are also suggestions of learning models where robotics is used in integration with inquiry learning [6]. Here is one possibility to benefit from robotics in the context of compulsory curriculum while inquiry as a method is and robotics as a tool is not an obligatory part of many science curricula. However, inquiry learning can be also related with problem solving while it can be applied effectively for improving several analytical and problem solving skills [9, 10]. In that case, robotics in integration with inquiry learning could create a new powerful synergy. That is the reason to measure inquiry skills while using robots. And finally, programming is also one of the tasks when learning with most robotic kits. Only the simplest robots do not offer students opportunity to program, however most of the robotics kits do. Out of the 11 problem solving types de-scribed by Jonassen [5] decision making and troubleshooting problems are the most appropriate in the context of robotics education. Either working on a robotics project individually or in a group, students have to make decisions on how to solve problems with robots. Those decisions may be connected to specific hard-ware or software problems. The first step in building a robotic system is planning. As students are often in groups while working with robots, team members have to reason with other team members as to why they think some solutions are better than others. After planning comes the stage of developing hardware and software, which is then followed by a debugging process. During this step students have to find and correct errors in their robotic solution. This is similar to the process of solving troubleshooting problems. Therefore, skills related to these two types of problems should be evaluated when using robotics in education. Improvement of the third set of skills, programming skills, is measured because students program robots and through time use more complicated structures to reflect their thoughts. Through possibilities of programming the robots can be seen as tools for learning that can be used for building something [11]. Programming activities are also characteristic of learning with LEGO Mindstorms robots. Those robots have specific potential to give feedback to the student's programs. This means that students momentarily see their program working in real world with the help of robot. The programming activities can be performed while using many possible programming languages available for robots, but mostly students start with graphical environments while it is more connected to thinking algorithmically than in the case of textual programming languages. In this case, students also do not have to debug syntax mistakes. However, even the graphical programming language could have positive effects on improving learners' programming skills while using robots. To sum up, there is a need for an instrument that measures all of these skills in the context of robotics education. Currently there is missing such a widely applicable instrument, and it could be one of the reasons why the effect of robotics in education on students' skills has been measured in only a few studies. The effect of robotics education can be assessed according to an experimental design where pre- and post-tests are applied. Therefore, two versions of the instrument are needed for measuring students' skills in problem solving, inquiry, and programming. This will help lead to a broader evaluation of using robotics in schools. If the effect of robotics education is not carefully measured, we do not expect much development in learning methodologies and support for robotics in the classroom. Widespread implementation of robotics requires encouraging results from using robots in the classroom. To solve this problem the current study formulated two research questions:

- At what level is the theory- and practice-driven instrument designed in the current study applicable for evaluating students' problem solving, inquiry and programming skills?
- Is it possible to show that the two versions of the instrument are not significantly different from each other?

2. Methods

2.1. Research design and procedures

Development of an instrument for evaluating the effect of applying robots in education began with clarifying the skills that robotics can significantly affect (Fig 1). It was done by integrating the relevant literature and consulting teachers and educators who are in contact with educational robotics in the classroom. It was concluded that affected skills include problem solving skills, inquiry skills, and programming skills. Tasks for the instrument concerning each of the skills were developed together with evaluation guide with expected answers. Next, pilot tests for Validating the instrument were conducted. The test was improved according to the difficulties the students had in completing it. More explanations were added if the students did not understand particular questions and some questions were revised so that they contained more details. After that the first pilot test for assessing students' skills was conducted. The test was successful while students did not express more difficulties. Therefore, it was possible to start with another test where the second version of the instrument was compared with the first version. It was needed to validate comparability of the two versions according to the level of difficulty.

2.2. Instrument description

An instrument for evaluating students' problem solving skills, inquiry skills, and programming skills in the context of robotics education for students at age from 13 to 16 was developed. This instrument is web-based and can be accessed with specific guidance for using and assessment grid at www.robootika.ee/roboskill. In applying this instrument students cannot go back from one part to a previous

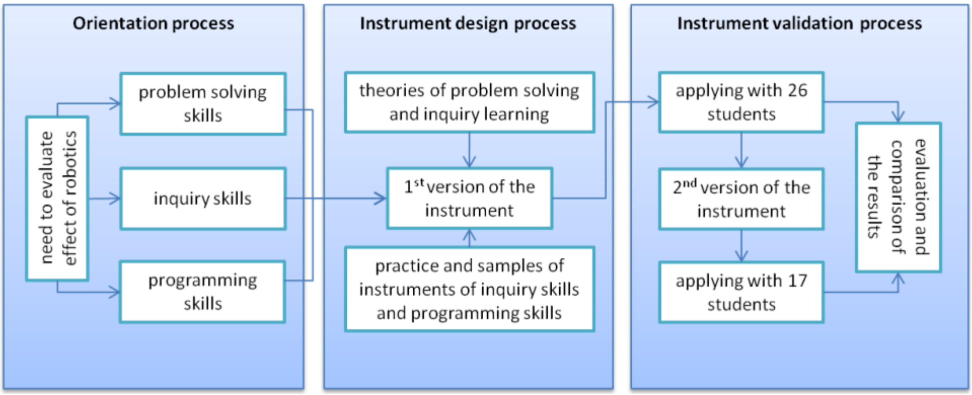

Fig. 1: Instrument design process of the study.

part to change their answers. This is because the questions have two dimensions –first students are asked in a theoretical context and then later instructed based on this theory to practically solve a problem. If they could return to previous pages, they could copy theoretical answers to previous questions. The structure of the instrument is shown on Fig 2.

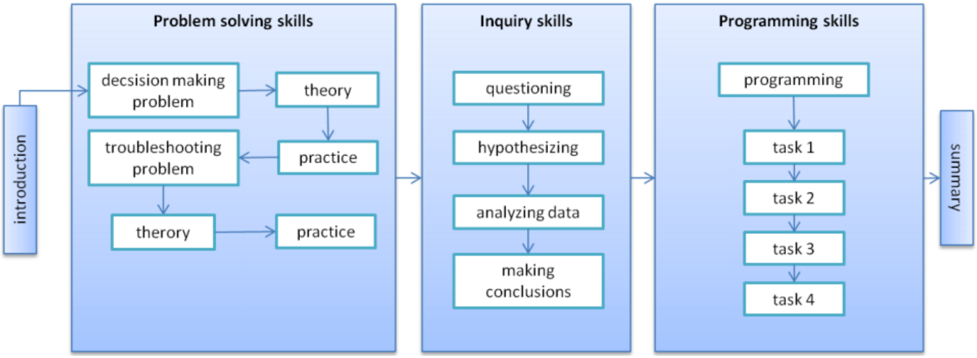

Fig. 2: Structure of the instrument designed for evaluating students' problem solving skills, inquiry skills, and programming skills in the context of robotics education.

According to the structure of the instrument the first step is introduction where students are introduced to the idea of the questionnaire, time consumption and work procedures. One of the rules students have to follow is that peer communication while answering questions is not allowed. This is something that the teacher or researcher should ensure since it cannot be guaranteed by the computer. There might be expected communication about sharing robots, because in most cases there are two or more students per robot. During the introduction students are asked about their age and name in order to link pre-and post-test results. In the second stage students are asked about a theoretical decision making problem. That question does not provide students with any additional background information about the expected answer. The reason to have a theoretical question is to have a better opportunity to evaluate answers because the second question does provide theory for the expected answer (in order to avoid the problem that skills are not applied due to lack of knowledge). It includes explanations about arguments and scales and asks students to follow that structure when solving a problem task. Students make up arguments that are for and against decision and give arguments scales. Scale is a number from one to ten to measure importance of arguments. In the instrument has been also applied the principle of using multiple items per task type in order to increase internal consistency [12]. Basically, students are presented with an actual decision making problem and asked to write down arguments with scales in favor and against the decision. They are also asked to write down decisions they would make. The given answers are evaluated by a specific grid (see www.robootika.ee/roboskill). For example, it is interesting to see if students made decisions based on scales or something else. The same method is used in the part for solving a troubleshooting problem. Students are first asked how they would theoretically solve a problem with the robot. Then, it is followed by a practical part where students have to download an example program to the robot and test it. This program is made with one intentional error and students are expected to troubleshoot the robot's behavior and write down all the changes they made to the program/robot and the results they got. It is expected that trials are based on previous results and students are using a trial-and-error method to find the error in the program. In the instrument's part for evaluating inquiry skills students are presented with a story problem connected to robotics education and asked to form a research question based on that problem [10]. Next, they have to form a hypothesis that is based on their research question. Then students analyze a graph that describes the problem. Students are also asked to analyze graph and find two ranges from the graph. The first range describes part of the graph where data change is not substantial. The second range describes part of the graph where data change is substantial. Answers are evaluated by the error of the range. A correct answer is if the error of a given range is below five, i.e. the range start and end point are with-in ±5. Values in graph range from 0-100. Students are then asked about the relationship between impact factor and characteristic of the changing value. They have to make conclusions and answer whether these conclusion support hypothesis or not. The third part of the instrument is for evaluating programming skills. This is the most practical part of the questionnaires and needs the most attention from students. Programming involves two types of assignments, both having two tasks. Previous stages of the questionnaire were based on theoretically proven examples, but measurement of programming skills lacks such example instruments. That is the reason why we developed new practical tasks. To increase effectiveness of the part of programming, there are more tasks than compared to previous stages. It is connected to internal consistency – programming skill is measured by four assignments in case of one assignment alone might not express programming knowledge. Evaluation of programming skill is very difficult [13]. Mostly quality and time can be evaluated. Quality being very narrow scaled, it is another reason for having more tasks to increase variation of points given in programming. There ought to be also a possibility to calculate time consumed for programming tasks. In this case, time was not measured, but it is important to use web based systems that enable users to do that.

3. Pilot study

There were piloted two versions of the instrument. The first test involved twenty six general school students (ages 13-16) and the second test seventeen students (ages 13-16). In the first test the aim was to evaluate if the items of the instrument could be applied within that age group. Therefore, we monitored if students indicate specific questions, misunderstandings or other problems concerning particular items. Some small changes were made to the instrument according to the feedback. After that we calculated the response rate and students' achievement in completing the tasks. There was expected that most of the students are able to solve the tasks and their average achievement will be between 20% and 80%. Otherwise, the questionnaire would have been either too difficult to assess students' initial level of skills or too easy to monitor their progress if the outcomes would be at a too high level already when using the instrument as a pre-test. When the first version of the test met these conditions, the second version was developed. In piloting the second version of the test the main aim was to evaluate if it is comparable with the first version. If these two are comparable then the versions can be applied in an experimental design where pre -and post-tests are applied. Data analysis will show if there is statistically important significance between answers of both tests. Answers were analyzed according to an evaluation tool (see www.robootika.ee/roboskill). The comparison of the answers was made using independent samples T-tests while the students’ scores distributed according to the normal distribution. Students’ answers were compared in three sections according to the structure of the instrument. There were calculated if there are comparable sections for assessing particularly problem solving skills, inquiry skills, and programming skills.

4. Results and discussion

In this study we had to first confirm if instrument developed based on theory and practice is applicable for evaluating students' problem solving, inquiry and programming skills. Therefore, the first questionnaire was applied to 26 students. Their completion rate and achievement was calculated separately for the sections of assessing problem solving skills, inquiry skills, and programming skills. The actual number of students piloting the first questionnaire was different between stages. All 26 students involved in the first pilot answered all problem solving and inquiry questions and 80% of them did programming tasks. Getting poor results was the biggest difficulty during piloting and might have been connected to lack of students’ programming skills. 20% of the students answered that they don’t know how to solve programming tasks. In the second pilot there were 17 students among whom only one did not do any programming tasks. Programming results were below 20% for the same reason mentioned before. But it is still acceptable as robotics is the first programming experience many students encounter. We cannot expect them to have programming experience before the experiment. It turned out that most of the answers were within limits, meaning that the test was not too hard or too easy. Data analysis showed that the students' results in using two versions of the instruments were not statistically significantly different. In conclusion it is possible to say that according to the completion rate and students’ achievement the questionnaire is applicable for measuring effectively problem solving and inquiry skills. However, measuring of programming skills is possible with some limitations. These can be measured if the students already have some experiences with programming. The applicability of the same instrument in the case of non-experienced students’ needs confirmation in additional studies. Next it was analyzed if the two versions of the same instrument are not significantly different from each other. This is important for studies where pre -and post-tests are needed. In this case the same version can be applied if the time between two measurements is long enough but often there are needed comparable versions of the same instrument; especially if the tasks of the instrument contain specific content and students could reach to higher score in the second measurement only because they have learned something from the first measurement. Pilot study results are described in Table 1. Mean results of problem solving and inquiry skills of questionnaires are above 20%. It fits the idea that questions were not too simple or difficult. It indicates that it is possible to measure development of skills on both problem solving and inquiry skills. Measuring programming skills needs some more attention and analysis. As programming skills are not that common among students, it is not taught on a curriculum basis and there are not many proven methodologies available to adapt [13]. At the moment, only quality of the programs was evaluated. Analysis of programming results indicated that time spent on programming should also be taken into account. Skill level on programming is low. Our hypothesis, like written before, is that it is because not all students in the sample had programming experience on robotics or in any other programming language. This is based on questions in programming where some student's answered that "I don't know", "I haven't done this before" or "I can't do it". Tasks given in programming assume possibility to use robots and programming environment. In this pilot test – NXT robots together with Edu NXT programming environment was used. As setting up robots and the physical environment to conduct the test requires a lot of time and effort from teachers – easier and simpler ways to measure programming skills should be considered. Students should have the possibility to fill out all the questions online without using any extra hardware. At the moment, programming skills can be measured on students who have some experience with programming LEGO Mindstorms robots. Another reason we considered to be a problem for measuring programming skills is that these tasks need more effort from students and they tend to skip these questions simply justifying their inability to solve problems. In one group who piloted the second questionnaire, all students had programming experience with LEGO Mindstorms robots. Two students out of ten did all the programming tasks. As opposite, more answers were given on questions that simply assumed writing text. There is possibility to use more supervision while conducting the test, but we find this to be more inconvenient for both researchers and teachers. As students are expected to work alone, any disturbance could affect results and answers. It was seen also in answers that some students cooperated and answered similarly.

As can be seen from Table 1, difference between problem solving, inquiry learning and programming of questionnaires is not that large.

Table 1. Comparison of students' achievement in problem solving, inquiry and programming skills in two versions of the instrument developed in the current study.

Skills	Mean scores of answers to the questionnaire: version 1	Mean scores of answers to the questionnaire: version 2	t	p
Problem solving skills	34.2%	41.2%	-1.7	>0.05
Inquiry skills	35.0%	34.4%	0.1	>0.05
Programming skills	18.1%	18.8%	-0.1	>0.05

T-test values show that difference between questionnaires is not statistically significant, so they can be used to measure these skills and results are possible to compare to see any development among skills of students.

5. Conclusion

In this study we designed two versions of an instrument for evaluating problem solving, inquiry and programming skills of students (aged 13 to 16) who use robotics. Item-completion rate and student' achievement showed that this instrument is not too complicated or too easy for the students. The tasks for measuring problem solving and inquiry skills were completed by all students and the average achievement was 34% and 35% respectively. Future studies are needed in the case of assessing students' programming skills because completion rate was low and average achievement only 18%. According to the qualitative analysis this part of the instrument is applicable in the case of students who have some experiences with programming; however, its applicability with other students needs additional confirmation. Next, there were compared two versions of the same instrument. The results showed that there is no statistically significant difference between the versions, so they can be used in any order as a pre-and post-test. Based on the findings of the current study, the instrument is applicable for measuring problem solving, inquiry, and programming skills while using robotics. However, it can be also be applicable in measuring students' skills when applying any other similar learning method. Measuring programming skills still needs some more attention.

6. References

[1] S. Papert, *Mindstorms - Children, Computers, and Powerful Ideas*. New York: Basic Books, 1980**.**

[2] A. Bredenfeld, A. Hofmann, and G. Steinbauer, "Robotics in Education Initiatives in Europe - Status, Shortcomings and Open Questions," in *Workshop Teaching Robotics, Teaching with Robotics*, Darmstadt, Germany, 2010, pp. 568-574.

[3] D. Alimisis, "Robotics in Education & Education in Robotics: Shifting Focus from Technology to Pedagogy," in *Robotics in Education*, Praha, 2012, pp. 7-14.

[4] S. Hussain, J. Lindh, and G. Shukur, "The effect of LEGO Training on Pupils' School Performance in Mathematics, Problem Solving Ability and Attitude: Swedish Data " *Educational Technology & Society*, vol. 9, pp. 182-194, 2006.

[5] D. H. Jonassen, "Toward a design theory of problem solving," *Educational Technology Research and Development*, vol. 48, pp. 63-85, 2000.

[6] H. Altin, M. Pedaste, and A. Aabloo, "Educational Robotics and Inquiry Learning: A Pilot Study in a Web-Based Learning Environment," in *11th IEEE International Conference on Advanced Learning Technologies*, Athens, Georgia, USA, 2011, pp. 224 - 226.

[7] D. T. Butterworth, "Teaching C/C++ Programming with Lego Mindstorms," in *Robotics in Education*, Praha, 2012, pp. 61-67**.**

[8] S. Papert, The Children's Machine: Rethinking School in the Age of the Computer. New York: Basic Books, 1993.

[9] M. Pedaste and T. Sarapuu, "Developing an effective support system for inquiry learning in a Web-based environment," *JOURNAL OF COMPUTER ASSISTED LEARNING*, vol. 22, pp. 47-62, 2006.

[10] M. Mäeots, M. Pedaste, and T. Sarapuu, "Transforming students' inquiry skills with computer-based simulations," in *8TH IEEE INTERNATIONAL CONFERENCE ON ADVANCED LEARNING TECHNOLOGIES*, Santander, SPAIN, 2008, pp. 938-942

[11] B. Brevik, S. Sjøberg, and A.-L. H. Tarrou, "Technology Education Through Open Ended Teaching Strategies Associated With Practical Learning Tools " in *XIII.IOSTE Symposium, The Use of Science and Technology Education for Peace and Sustainable Development.*, Kuşadası / Turkey 2008.

[12] Z. Dörnyei and T. Taguchi, Questionnaires in Second Language Research: Construction, Administration, and Processing: Routledge, 2009.

[13] G. R. Bergersen, J. E. Hannay, D. I. K. Sjøberg, T. Dybå, and A. Karahasanović, "Inferring Skill from Tests of Programming Performance: Combining Time and Quality " in *5th Int'l Symp.Empirical Software*

Engineering and Measurement (ESEM), 2011.