# Envisioning a Computational Thinking Assessment Tool

Ilenia Fronza
Free University of Bolzano
Piazza Domenicani 3, 39100 Bolzano, Italy
ilenia.fronza@unibz.it

Claus Pahl
Free University of Bolzano
Piazza Domenicani 3, 39100 Bolzano, Italy
claus.pahl@unibz.it



## Abstract

Recent work on Computational Thinking (CT) has focused on proposing new curricula but in many cases the assessment phase has been overlooked. The issue is critical because appropriate assessment is needed to facilitate the incorporation of CT in the curriculum. What is now clear from the existing literature is that there is a need to build on top of the existing multiple forms of assessments, in order to integrate multiple approaches and reach a comprehensive assessment of CT learning. In this paper, we envision a system that integrates different types of assessments while providing an intuitive interface in order to allow teachers to see and supervise the overview of the learning process, with the possibility to assess individually the student's learning. To assess the suitability of our idea, we describe the Proof of Concept of a mobile application to assist CT assessment, and we discuss the challenges that need to be solved to create such an application.

## 1 Introduction

Computational Thinking (CT) is considered as a key set of skills that must be learned by today's generation of learners, both in the context of STEM and other subjects [Gro17]. Therefore, CT has caught the attention of a broad academic community. Many studies have first tried to capture the essence of CT and to create an agreed definition, as CT was rather a broad term [BR12, Win14]. Then, researchers and educators have focused on designing new activities to foster CT across the curriculum. Unfortunately, in many cases more weight has been given to the development of methods to teach CT than on proposing how learning will be assessed and evaluated. This is a relevant issue, because appropriate assessment instruments are needed to incorporate CT in the curriculum [HL15].

Most of the existing assessment procedures focus on code analysis. This approach might be convenient in a professional environment [CF15], but it neglects CT broader aims [Cro14, Net13]. Other procedures, while valuable for research and for providing a view of student's learning, are subjective, time-consuming, and not easily usable regularly in classrooms [Gro17].

Nowadays, many researchers agree that using only one type of assessment can lead to misunderstand the development of CT skills [RGMLR17]. Therefore, in order to reach a comprehensive assessment of CT learning, a "system of assessments" needs to be adopted, i.e., a combination of omplementary assessments tools. To this end, S. Grover [Gro15] suggests to consider multiple measures that are complementary, encourage and reflect deeper learning, and contribute to a comprehensive picture of students' learning.

At this point it should be considered that the use of a system of assessments could make the teacher's work even more complicated, which could result in limiting considerably the adoption of this type of systems in the classroom. In fact, the teacher would need to regularly apply different types of assessment (e.g., code analysis, interview) and then integrate the results to obtain an overall assessment.

To address this issue, we envision a system that integrates different types of assessments while providing an intuitive interface in order to allow teachers to see and supervise the overview of the students' learning process, with the possibility to assess individual learning. To assess the suitability of our idea, we describe the Proof of Concept (PoC) of a mobile application,



In: A. Piotrkowicz, R. Dent-Spargo, S. Dennerlein, I. Koren, P. Antoniou, P. Bailey, T. Treasure-Jones, I. Fronza, C. Pahl (eds.): Joint Proceedings of the CC-TEL 2018 and TACKLE 2018 Workshops, co-located with 13th European Conference on Technology Enhanced Learning (EC-TEL 2018), 03-09-2018, published at http://ceur-ws.org

and then we discuss possible challenges that need to be solved in order to create such application.

Section 2 reports the state of the art of existing CT systems of assessments. Section 3 describes the proposed Proof of Concept of a mobile application to assist CT assessment, and Section 4 discusses the challenges that need to be faced in order to create such application. Section 5 draws conclusions from this work, also proposing possible directions for future work.

## 2 State of The Art

Recent work on Computational Thinking (CT) has focused on environments and tools that foster CT [GP13, RWI10], new curricula for CT (e.g., [SFH+12, FEIJ+14]), also using CT as a medium for teaching other subjects (e.g., [FEIC15, FZ15, Edw11]). In this scenario, a major gap has emerged in research on CT assessment. Indeed, in many cases more weight is given to the development of methods to teach and foster CT without thinking about how those methods will be assessed and evaluated. This issue is critical, because assessment not only determines whether or not educational goals are being met, it also drives the design of the curriculum [HL15]. Grover and Pea make the gravity of the lack of CT assessment clear: "Without attention to assessment, CT can have little hope for making its way successfully into any K-12 curriculum" [GP13].

Given the absolute need for an assessment methodology, many efforts in the last years aimed specifically at tackling the issue of CT assessment. An overview of the proposed approaches, divided according to their perspective (e.g., summative assessment, perceptions-attitudes scales, etc.), can be found in the recent work of Román-González et al. [RGMLR17].

What is clear from the existing literature, is the need to build on top of the multiple forms of assessments that have been proposed so far, in order to reach a comprehensive assessment of CT learning. Research is now moving in this direction, and some examples of "systems of assessments" have been proposed [Gro15].

For example, Brennan and Resnick [BR12] described three approaches for assessing the development of CT concepts, practices, and perspectives. Fronza et al. [FIC17] developed a framework to assess the development of computational concepts and practices. Román-González et al. [RGMLR17] aimed at studying the convergent validity of their CT summative-aptitudinal assessment test with respect to other assessment tools. S. Grover [Gro17] described the multiple forms of assessments designed and empirically studied in a middle school introductory computing curriculum.

If we want these systems of assessments to make it to the classroom, in addition to building these systems, another goal must be to reflect on how to facilitate their adoption in the classroom, without leaving on the teacher's shoulders the task of manually integrating different types of assessments to achieve an overall assessment.

Despite this need, few if any tools exist that enable real-time, overall, formative assessment of CT. As we describe in the next section, we aim to fill this research gap envisioning an assessment tool that assists CT assessment, by integrating different types of assessments and providing an intuitive interface.

## 3 CT Assessment Tool: A Proof Of Concept

We envision a mobile application that supports CT assessment as follows. It integrates different types of assessments, also providing an intuitive interface in order to allow teachers to see and supervise the overview of the students' learning process, with the possibility to assess individually the students' learning. In this section, we describe a Proof of Concept of this application. First, we describe a possible approach to define the assessment framework that the application should implement. Then, we detail possible design choices that would ease the assessment process. Finally, we report on architectural considerations.

### 3.1 Assessment Framework

Following the guidelines of the existing literature, our assessment framework, being a systems of assessments, should assess different skills, such as: CT concepts, practices, and perspectives [BR12], cognitive skills, and also social and relational skills [Wor94]. The "Goal-Question-Metrics" (GQM) measurement model [BCR94] could be adopted in order to provide an effective view of student's learning. In fact, the GQM approach helps to clearly specify the object of study, the aspect of study, the purpose of the assessment and the environment in which the data is collected.

The GQM approach foresees the definition of measurement goals, questions, and metrics. In our case, *goals* specify the assessment needs (i.e., CT skills) in a formal way; *questions* define information gaps that need to be filled to understand whether a measurement goal can be achieved or not; and *measurements* help to answer the measurement questions.

An example of such a model is depicted in Figure 1: the problem of assessing the learning of computational concepts is modeled as a GQM goal. The goal is then assessed using measurement questions: in the example of Figure 1, these questions are how well the student understands conditionals, sequences, and loops. Each question is then answered using one or more measurements, for example, one measurement to understand whether the student understands the correct usage of a loop is the presence of a "for" construct in the source code provided by the student, when this is required by the given problem.

### 3.2 Design of the Visual Support

A dashboard would be an effective means to visually display the outcome of the assessment, which is modelled in form of GQM models. In dashboards, in fact, pre-attentive properties (e.g., color, shape, location) are used to maximize the understanding of the displayed information and to guide attention [Few13].

Each tile in the dashboard shall represent a specific measurement. Suppose that, for example, the tile in Figure 2 shows the number of "for" statements in the code. In this case, there are 12 loops in the code and this number is higher respect to the previous measurement (as shown by the upward arrow). The tile is colored in green to show that a sufficient level has been reached, therefore the teacher does not have to focus on this tile anymore. If, instead, no loops in the code are found, then the tile shall be red. The teacher, in this case, needs to find out if there was a need of a loop in the project and the students did not use the appropriate block but, for example, just repeated the same command many times. In this case, the computational concept "loops" would not be assessed positively [FIC17].

Figure 3 shows an example in which three skills (i.e., goals) are depicted with associated questions and metrics. For example, metric B is colored in red and therefore requires the attention of the teacher. Metric E is colored in yellow, therefore depicts a warning level which should be kept under observation. Tiles that are colored in green depict a satisfactory fulfillment of the metric.

The same rule applies to the each *skill tile*. When all the corresponding metrics are colored in green, also the skill tile is colored with the same color in order to show immediately to the teacher that there is no need of attention for that skill (for example, "CT skill 2" in Figure 3). The skill tile is red, instead, when immediate intervention is needed because all the metrics are colored in red (for example, "CT skill 3"). A yellow skill tile (for example, "CT skill 1") indicates that the corresponding skill should be kept under observation.

The existing literature recommends to use an assessment framework in several points of the learning process [BR12]. The proposed dashboard would help the teacher to focus on critical aspects: she/he will ignore the green tiles and focus on the orange and red ones, which means she/he would assess and provide feedback on those aspects that have not been learned yet or require a more detailed explanation. In the example shown in Figure 3, it would be immediately clear to the teacher that "CT skill 2" has been achieved.

The rules defining whether a tile becomes green, yellow, or red embed the knowledge extracted during the GQM definition, i.e., the conditions under which teachers can assess whether a skill has been acquired or not.

### 3.3 Architectural Considerations for the Assessment Tool

The architectural considerations for the CT measurement tool we consider are the following:

- Functionality. The assessment tool shall have the necessary features to receive the input of key concept or areas from each student, as well as the comments and preliminary assessment of the teacher. It shall allow the track, follow up, comparison, evaluation and projection of each considered skill. Due to the scope of the tool, the modules and features pertaining data reports will be of particular importance through the development and acceptance of this tool.
- Maintainability. The CT model, formulated as a GQM measurement model, is of crucial importance for the interpretation of the data. Therefore, it has to be implemented in a modular, configurable way for two reasons: it has to possible to configure an existing model to accomplish small to medium changes but we want also that it is possible to replace the entire model with a new one, if new considerations arise. Moreover, changes could be necessary to allow using the assessment tool at different education levels (such as K-12 or university). In consequence, the assessment framework has to be implemented as a pluggable, configurable component of the final tool, which interacts with the designed dashboard through clear interfaces.
- Usability. A desirable characteristic for the assessment tool is to be designed to be used on a tablet because it would allow teachers to move easily in the room (from one student's workstation to another one).

The next section discusses possible challenges that need to be faced in order to create our envisioned CT assessment tool.

## 4 Discussion

Before starting the development of a real application from the PoC that we have envisioned in Section 3,

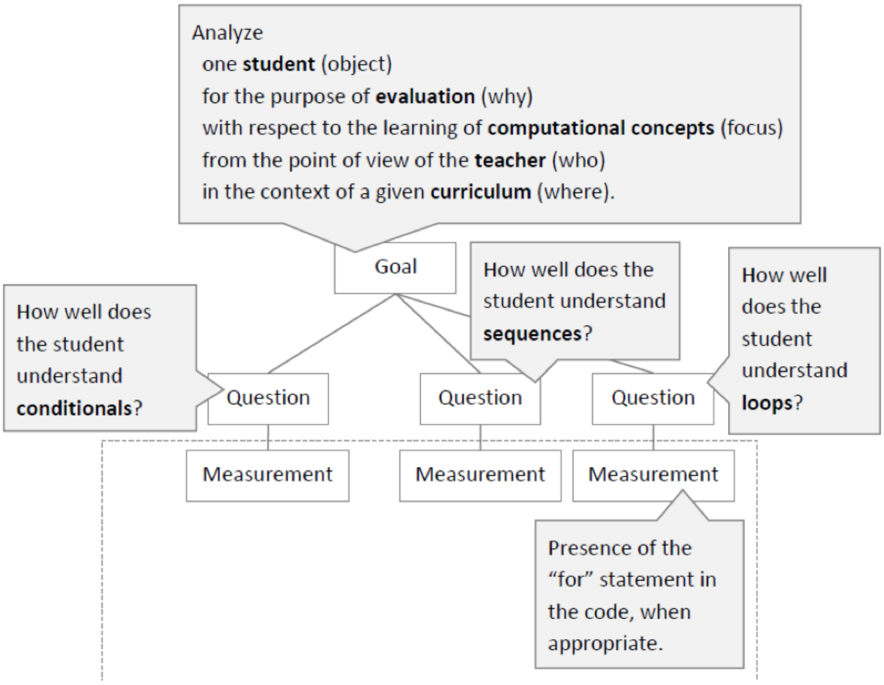


Figure 1: Illustration of CT assessment, modeled in form of a GQM.


Figure 2: Structure of a tile in the assessment dashboard.

a number of challenges need to be solved. In the following, we discuss the main challenges that we have already identified.

**Definition of measurements and thresholds.** The definition of the CT assessment model (Section 3.1) is a crucial step towards the implementation of our PoC, and requires the definiton of measurements and thresholds. This requires extensive effort and empirical research, by iterations, also involving teachers to collect their feedback.

**Customization.** To facilitate the adoption of this tool in the classroom, attention needs to be taken to make it customizable. For example, a teacher may decide to carry out activities to improve only a specific subset of skills; in this case, he/she should be able to select (and integrate) only a part of the available measurements in the tool.

**Code analysis.** The tool shall be adaptable to each level of education. From the code analysis perspective, this means for example foreseeing the possibility of analyzing code written in different programming languages (e.g., block- or text-based).

**Social skills.** One of the most critical aspects will probably be assessing social and relational skills. Indeed, a comprehensive method has not been proposed so far, and the current approach is to evaluate life-skills trainings through observations, questionnaires [AVM15], or self-report [AM15]. The tool shall provide an opportunity to collect and analyse data from these sources, at different points in the learning process.

**Iterative development and empirical research.** The tool shall be developed in an iterative fashion, performing empirical research with students and teachers, and adapting it to the feedback obtained. A research method that supports such approach is "design research". By its nature, design research is relevant for educational practice as it aims to develop research-based solutions for complex problems in educational practice or to develop or validate theories about processes of learning and teaching [VdAGMN06]. Design research incorporates systematic educational design processes and, like all systematic educational and instructional design processes, it is cyclical in character: analysis, design, evaluation and revision activities are iterated until an appropriate balance between ideals ("the intended") and realization has been achieved [VdAGMN06].

Clearly, this type of method requires numerous classroom experiments to be carried out, and replicated, to draw relevant information from them.

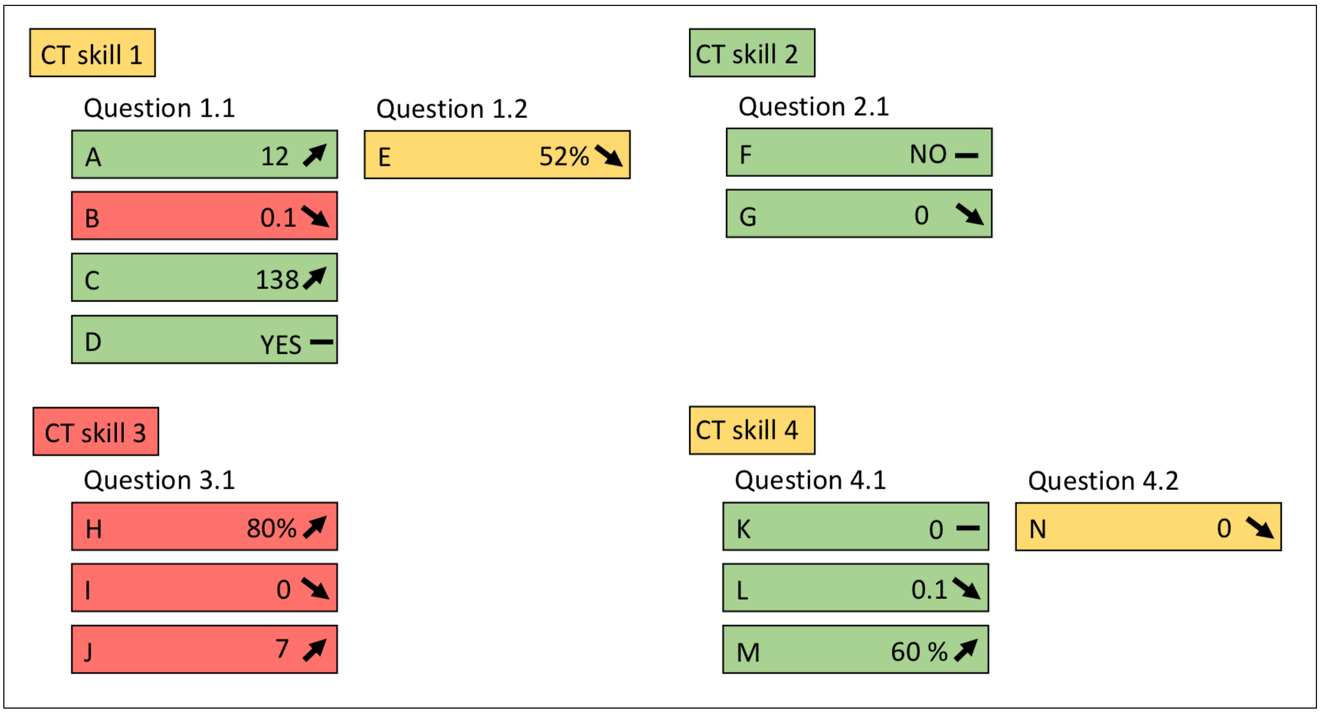


Figure 3: Dashboard structured using pre-attentive visualization techniques (adapted from [JSS13]).

## 5 Conclusion and Future Work

In this paper, we described our vision of a mobile application to assist CT assessment, and then we discussed possible challenges that need to be solved in order to create such application. The results of this analysis allow us to outline a research agenda that we believe will help to implement the idea we have described. An interdisciplinary approach is needed, which draws from software engineering, didactics, and cognitive psychology. Research efforts should focus primarily on: *i)* definition of the assessment framework (using a three-step approach: what? so what? now what?); *ii)* implementation of a software product that will sustain operationally the assessment framework; *iii)* data mining and effective visualization; *iv)* empirical research in classroom.

## References

[AM15] Hayley Anthony and Louise A McLean. Promoting mental health at school: short-term effectiveness of a popular school-based resiliency programme. *Advances in School Mental Health Promotion*, 8(4):199–215, 2015.

[AVM15] Narges Adibsereshki, Abas Mahvashe Vernosfaderani, and Guita Movallali. The effectiveness of life skills training on enhancing the social skills of children with hearing impairments in inclusive schools. *Childhood Education*, 91(6):469–476, 2015.

[BCR94] Victor R. Basili, Gianluigi Caldiera, and H. Dieter Rombach. The goal question metric approach. In *Encyclopedia of Software Engineering*. Wiley, 1994.

[BR12] Karen Brennan and Mitchel Resnick. New frameworks for studying and assessing the development of computational thinking. In *2012 Annual Meeting of the American Educational Research Association (AERA'12), Vancouver, Canada*, pages 1–25, Vancouver, Canada, 2012. AERA.

[CF15] L. Corral and I. Fronza. Better code for better apps: A study on source code quality and market success of android applications. pages 22–32, 2015.

[Cro14] Dan Crow. Why every child should learn to code, feb 2014.

[Edw11] Michael Edwards. Algorithmic composition: computational thinking in music. *Communications of the ACM*, 54(7):58–67, 2011.

[FEIC15] Ilenia Fronza, Nabil El Ioini, and Luis Corral. Students want to create apps:

Leveraging computational thinking to teach mobile software development. In *Proceedings of the 16th Annual Conference on Information Technology Education*, SIGITE '15, pages 21–26, New York, NY, USA, 2015. ACM.

[FEIJ+14] Ilenia Fronza, Nabil El Ioini, Andrea Janes, Alberto Sillitti, Giancarlo Succi, and Luis Corral. If i had to vote on this laboratory, i would give nine: Introduction on computational thinking in the lower secondary school: Results of the experience. *Mondo Digitale*, 13(51):757–765, 2014.

[Few13] Stephen Few. *Information Dashboard Design: Displaying data for at-a-glance monitoring*, volume 5. Analytics Press Burlingame, CA, 2013.

[FIC17] Ilenia Fronza, Nabil El Ioini, and Luis Corral. Teaching computational thinking using agile software engineering methods: A framework for middle schools. *ACM Transactions on Computing Education (TOCE)*, 17(4):19, 2017.

[FZ15] I. Fronza and P. Zanon. Introduction of computational thinking in a hotel management school [introduzione del computational thinking in un istituto alberghiero]. *Mondo Digitale*, 14(58):28–34, 2015.

[GP13] Shuchi Grover and Roy Pea. Computational thinking in k–12. a review of the state of the field. *Educational researcher*, 42(1):38–43, Jan/Feb 2013.

[Gro15] Shuchi Grover. "systems of assessments" for deeper learning of computational thinking in k-12. In *Annual Meeting of the American Educational Research Association*, pages 1–9, Washington, DC, 2015. AERA.

[Gro17] Shuchi Grover. Assessing algorithmic and computational thinking in k-12: Lessons from a middle school classroom. In *Emerging research, practice, and policy on computational thinking*, pages 269–288. Springer, 2017.

[HL15] Roxana Hadad and Kimberly A Lawless. Assessing computational thinking. In *Encyclopedia of Information Science and Technology, Third Edition*, pages 1568–1578. IGI Global, 2015.

[JSS13] Andrea Janes, Alberto Sillitti, and Giancarlo Succi. Effective dashboard design. *Cutter IT Journal*, 26(1), 2013.

[Net13] Warren Nettleford. Primary school children learn to write computer code, jul 2013.

[RGMLR17] Marcos Román-González, Jesús Moreno-León, and Gregorio Robles. Complementary tools for computational thinking assessment. In *Proceedings of International Conference on Computational Thinking Education (CTE 2017), S. C Kong, J Sheldon, and K. Y Li (Eds.). The Education University of Hong Kong*, pages 154–159, 2017.

[RWI10] Alexander Repenning, David Webb, and Andri Ioannidou. Scalable game design and the development of a checklist for getting computational thinking into public schools. In *Proceedings of the 41st ACM Technical Symposium on Computer Science Education*, SIGCSE '10, pages 265–269, New York, NY, USA, 2010. ACM.

[SFH+12] Amber Settle, Baker Franke, Ruth Hansen, Frances Spaltro, Cynthia Jurisson, Colin Rennert-May, and Brian Wildeman. Infusing computational thinking into the middle- and high-school curriculum. In *Proceedings of the 17th ACM Annual Conference on Innovation and Technology in Computer Science Education*, ITiCSE '12, pages 22–27, New York, USA, 2012. ACM.

[VdAGMN06] Jan Van den Akker, Koeno Gravemeijer, Susan McKenney, and Nienke Nieveen. *Educational design research.* Routledge, 2006.

[Win14] Jeannette M. Wing. Computational thinking benefits society, January 2014.

[Wor94] World Health Organization and others. Life skills education for children and adolescents in schools. In *Life skills education for children and adolescents in schools*, pages 2pts–in. 1994.