A multifaceted students’ performance assessment framework for motion-based game-making projects with Scratch

Ioannis Altanis and Symeon Retalis
Department of Digital Systems, University of Piraeus, Piraeus, Greece

ABSTRACT
In the last few years, engaging students to create digital games has been a pole of attraction for many teachers and researchers, resulting in highly positive learning experiences and promoting their thinking skills, e.g., programming and computational thinking (CT) skills. Researchers have already stated about the need for further research not only around the evaluation techniques and tools of the quality of these complex educational interventions, but mainly about ways to ease the assessment of students’ performance from multiple perspectives with authenticity. This paper contributes to proposing a multifaceted assessment framework of the degree of students’ acquisition of multiple skills, when they get involved in digital motion-based touchless game-making course-projects with the MIT Scratch tool. The results of its implementation during a pilot study with computer science undergraduate students, which are presented, highlight the positive effects of combining and extending various assessment techniques and tools to draw holistic conclusions about students’ higher skills including computational and spatial thinking skills.

KEYWORDS
Game design; assessment tools; motion-based touchless interaction; Scratch; computational thinking (CT) skills

1. Introduction

It is common belief that students must be equipped with 21st century knowledge and skills. Thus, students are often called to actively participate in complex problem-solving projects such as creating digital games (Hava & Cakir, 2017; Kafai & Burke, 2015). Research studies have shown that such projects ultimately help students develop higher skills, such as computational thinking and problem-solving skills (Akcaoglu, 2016; Hoover et al., 2016; Moreno-León, Román-González, Harteved, & Robles, 2017) and/or comprehend concepts of subjects, such as computer science (Baytak, 2009; Denner, Werner, & Ortiz, 2012).

A highly interesting type of digital games is the motion-based touchless games which are based on natural user interaction benefitting from the affordances of 3D cameras like Microsoft Kinect. Researchers have documented
that students show high interest and motivation in interacting with such type of games that can be played with hand and body gestures (Bekesi & Sik-Lanyi, 2016; Hsu, 2011). The creation of such games gives students the opportunity to enhance computational thinking skills, multi-literacies and STEM (Science-Technology-Engineering-Mathematic) related skills, as well as their spatial thinking skills through the perception of objects and their body in space and the comprehension of geometrical notions during implementation of gesture-based interaction (Authors of this study, 2018; Tsai, Kuo, Chu, & Yen, 2015; Yu & Zacks, 2015). Spatial thinking “involves the location and movement of objects and ourselves, either mentally or physically, in space” and involves a considerable number of dimensions, e.g., locating objects, navigating and wayfinding, etc. (Ministry of Education in Ontario, 2014; National Research Council, 2006).

There are very few research studies related to students’ motion-based touchless game-making activities (e.g., Sinker, 2014; Sullivan, Byrne, Bresnihan, O’Sullivan, & Tangney, 2015). Their results state the positive influence of digital game-making activities. Nevertheless, researchers claim that the assessment techniques and tools that should be used in such studies need refinement thus making them an open research topic (Ilic, Haseski, & Tugtekin, 2018; Lockwood & Mooney, 2017; Romero, Lepage, & Lille, 2017). At the same time, it has become evident that in order to draw conclusions on the growth of skills through such educational initiatives, existing techniques and tools should be methodically combined and enriched (Grover, Pea, & Cooper, 2015; Román-González, Moreno-León, & Robles, 2017). Thus, this paper tackles these challenges as it proposes a multifaceted framework for assessing the degree of students’ acquisition skills in an authentic and ongoing manner when they are involved in digital motion-based touchless game-making projects. This framework makes use of a blend of well-chosen teacher-friendly authentic assessment techniques and tools already validated in various related studies. Authentic assessments include that the assessments “give students both feedback upon completion” as well as “guide their work along the way” (Schack, 1994).

The proposed multifaceted framework focuses on assessing:

1. skills about game elements analysis and game design,
2. computational thinking (CT) skills,
3. programming skills and
4. spatial thinking skills.

Up to our knowledge, it is the first time that such a framework is being proposed for measuring students’ skills including spatial thinking skills during the lifecycle of the motion-based touchless game-making process. This framework helps teachers score the students’ growth of the aforementioned skills per se and have a clear picture of the strengths and weaknesses of each student.
The structure of the paper is as follows: In section 2 an overview of existing teacher-friendly assessment techniques and tools per key skill which have already been validated via various related studies is given. Section 3 presents the specificities of the proposed framework, providing for each skill the specific measurement criteria/indicators and the techniques and tools that can be utilized. Section 4 and 5 contains a showcase of the implementation the assessment framework to an undergraduate course for computer science students as well as a discussion of the main findings. Finally, section 6 presents some conclusions and topics for future research.

2. Motion-based game-making activities and assessment of students’ skills

During game-making projects, students are called to follow a three-phase process:

- **Design Phase**: Conceive the idea, analyze game elements and visualize the game through a storyboard, a concept map or a flowchart
- **Implementation Phase**: Implement the design as an artifact, i.e., a prototype of a digital game using the MIT Scratch tool
- **Evaluation Phase**: Evaluate the quality of their artifact

Via systematic motion-based game-making project activities students can improve several skills as discussed below.

2.1. Analysis and design skills

During the design phase, students break their game idea down to the formal and dramatic elements, e.g., describing them in a game design document (Fullerton, Swain, & Hoffman, 2004). In order to measure students’ performance, regarding the designing and analysing skills, one common technique is to assess the deliverables of the designing phase (e.g., Game design document, storyboard or flowcharts), which help students to visualize these game elements and enrich the game design (Akcaoglu, 2016; Burke & Kafai, 2012; Rankin, Thomas, Irish, & Hawkins, 2014). Such analysis is exceptionally useful as it leads to the faster and more effective programming of a digital game (Claypool & Claypool, 2005).

2.2. Computational thinking (CT) skills

In the last few years, it has been stressed the lack of consensus on how CT skills can be effectively and efficiently assessed by using complementary techniques and tools (Román-González et al., 2017). Two of the most well-known techniques are
(a) to match commands in their game’s code, which are directly associated with CT concepts, through game code analysis (Moreno-León & Robles, 2015; Techapalokul & Tilevich, 2017) and (b) to match students’ answers (from an Interview or Journal) with specific CT practices such as Experimenting & iterating, Testing & debugging and Reusing & remixing (ScratchEd Harvard online community, 2017).

2.3. Programming skills

The assessment of students’ programming skills can be performed through a game code analysis. The main indicators measured are:

1. the students’ ability to implement their game by using different types of commands, such as control, motion commands and variables (Aivaloglou & Hermans, 2016; Fields, Kafai, Strommer, Wolf, & Seiner, 2014).
2. the students’ ability to follow the common best practices in programming, e.g., the avoidance of duplicated scripts and incorrect names (Gutierrez et al., 2018; Moreno-León & Robles, 2015).
3. the students’ ability to create a game code that covers quality criteria (code smells), such as the avoidance of unused variables, long scripts, etc. (Techapalokul & Tilevich, 2017).
4. the students’ ability to develop a digital game that covers common design heuristics, such as usability and playability (Barcelos, Costa, Muñoz-Soto, Noël, & Silveira, 2013; Wilson, Hainey, & Connolly, 2012).

2.4. Spatial thinking skills

During digital motion-based touchless game-making projects, students create hand and body gestures, in order to interact with game elements in a virtual environment. The creation of a successful interaction seems to promote spatial thinking skills, as students mentally and physically process the spatial relationships among their body joints in space. As a consequence, students transform this spatial thinking process to hand/body gesture algorithms, by utilizing a programming language and by following drill and practice activities. Up to our knowledge, not any other paper has been published as yet that offers ways to measure students’ performance, regarding their spatial thinking skills, during digital motion-based touchless game-making projects. However, there are few published studies that deal with measuring the improvement of school children’s spatial thinking skills when they perform digital game-making projects using authoring tools such as Minecraft, Lego Mindstorms EV3, Kodu or Koduble/Lightbot (Caci, Chiazzese, & D’Amico, 2013; Foerster, 2017; Francis, Khan, & Davis, 2016; Lux, LaMeres, Hughes, & Willoughby, 2018). In those studies, the assessment of children’s performance was done via direct observation, video-recording analysis, as well as pre/post-tests.
3. The proposed assessment framework

The proposed assessment framework has been created using a careful mix and match of known techniques and tools per skill of interest which have been aforementioned at the presentation of techniques and tools (Unit 2). Table 1 offers an overview of the dimensions of the proposed assessment framework. Specifically, it provides the proposed students’ assessment performance indicators (assessment criteria) for each one of the four (4) key skills that teachers could assess.

The proposed assessment framework helps teachers employ authentic and ongoing assessment techniques in motion-based game-making projects using the MIT Scratch tool (with the aid of the Kinect2Scratch plugin) which is one of the most popular authoring tool in K-12 education for such purposes. Thus, Table 2 shows the tools and techniques that could fit to the proposed assessment framework for such Scratch Kinect-based game-making projects. The selection of tools and techniques per dimension of the assessment framework has been done using two main criteria: (i) their relevance to the specific skill under measurement is and (ii) their usability and acceptability based on the comments made by educational practitioners and research findings. In the following subsections, the way the various skills can be assessed is presented.

3.1. Assessment of analysing and designing skills

The game design document and the storyboard that are requested in the respective game development phase are structured according to the guidelines that had been suggested by Rankin (2014). Their assessment can be done

| Table 1. Dimensions of the proposed assessment framework. |
|---------------------------------------------------------|
| **Skills** | **Performance Indicators** |
| 1. Skills of game elements analysis and design | 1.1 The ability to break the game into smaller parts (formal & dramatic elements) analysing them on a game design document (henceforward GDD) |
| | 1.2 The ability to visualize (e.g., in the form of animated scripts using Storyboards) the elements of the GDD |
| | 1.3 The ability to enrich game design process, e.g., through Storyboards |
| | 1.4 The ability to analyse and design a digital game, meeting specific quality criteria during the design phase |
| 2. Computational thinking skills | 2.1 The ability to use commands in their game’s code, which are directly associated with computational thinking skills concepts |
| | 2.2 The ability to apply computational thinking practices during creating their digital game |
| 3. Programming skills | 3.1 The ability to implement their game by using different types of commands |
| | 3.2 The ability to use best practices in the usage of code |
| | 3.3 The ability to develop a game code that meets specific coding quality criteria |
| | 3.4 The ability to develop a “good” game |
| 4. Spatial thinking skills | 4.1 The ability to design and develop hand/body gesture-based interaction in the game with high intricacy while keeping at the same time high level of implementation |
| Skills                              | Phases of drawing data | Students’ deliverables | Assessment Technique/Tool                                                      | Source (based on)                                      |
|-----------------------------------|------------------------|------------------------|-------------------------------------------------------------------------------|-------------------------------------------------------|
| 1. Skills of game elements analysis and design | Phase 1: Design        | - GDD & Storyboard     | - Assessment rubric                                                          | - Felder, 2011                                         |
| 2. Computational thinking skills  | Phase 2: Implementation | - Game                 | - Automated analysis tool (Dr. Scratch)                                      | - Moreno-León & Robles, 2015                         |
| 3. Programming skills             | Phase 2: Implementation | - Game                 | - CT Journal assessment rubric                                                 | - Burke & Kafai, 2012                                 |
| 4. Spatial thinking skills        | Phase 2: Implementation | - Game                 | - Automated analysis tool (Scrape, Dr. Scratch, Quality Hound)                | - Moreno-León & Robles, 2015                         |
|                                  |                        |                        |                                                                               | - Techapalokul & Tlevich, 2017                       |
|                                  |                        |                        |                                                                               | - Desurvire & Wiberg, 2009                           |
|                                  |                        |                        |                                                                               | - Maike, Neto, Baranauskas, & Goldenstein, 2014      |
|                                  |                        |                        | - 5-likert assessment rubric for the quality criteria of a digital game       | - Authors of this study, 2018                         |
|                                  |                        |                        |                                                                               |                                                       |
|                                  |                        |                        | - Observation during game Playtesting                                          |                                                       |
|                                  |                        |                        | - code analysis of the algorithms for hand/body gesture-based interaction      |                                                       |

Table 2. Basic Components of the proposed assessment framework for Scratch Kinect-based game-making projects.
using an assessment rubric (Felder, 2011) which measures various criteria such as the clarity of the elements and the scope, play flow, clear ending, mechanical defects, etc.

3.2. Assessment of CT skills

In order to reach a comprehensive assessment of CT skills, a combination of complementary assessments tools should be used (Román-González et al., 2017). Firstly tools such as Scrape (Happy Analyzing) and Dr. Scratch (Moreno-León & Robles, 2015) can automatically assign a CT score in terms of basic CT concepts such as abstraction and problem decomposition, parallelism, logical thinking, synchronization, flow control, user interactivity and data representation by analysing the game source code. Secondly, in order to assess the computational thinking practices followed by students while creating the game, the CT journal proposed by Scratch Ed internet group that students keep while developing their game is being analysed. This journal requires students to answer questions directly related to CT practices such as remixing, debugging, etc.

3.3. Assessment of programming skills

Tools for the automatic analysis of the game source code are being used for assessing the different dimensions of the programming skills. More specifically:

- The use of different types of commands by getting descriptive statistics via the Scrape tool. (Burke & Kafai, 2012).
- Checking the implementation of best practices with Dr. Scratch tool (Moreno-León & Robles, 2015)
- Checking the code’s quality based in Software quality criteria with Quality Hound tool (Techapalokul & Tilevich, 2017).

In addition, the overall quality of the submitted game using a 5-point scale assessment rubric that scores in detail the game design usability heuristics (Desurvire & Wiberg, 2009; Maike et al., 2014). This rubric is being used for self and peer-assessment purposes apart from the teacher grading needs.

3.4. Assessment of spatial thinking skills

The proposed assessment framework measures students’ performance, with regards to their spatial thinking skills, by analysing the intricacy of the hand/body gesture-based interaction imbedded in the game and the quality of its implementation. Specifically:
(1) Intricacy is related to the number of hand/body gestures, the number of body joints (e.g., hand right, head, spine) used for the needs of natural user interaction and the number of coordinate axes used to compare body joints (x, y, z axes). Observation of the game play and game source code analysis is being used as assessment techniques.

(2) Regarding the quality of implementation, game code analysis is being performed in order to collect data about: (a) the proper use of CT concepts in gesture algorithms (e.g., if, wait until), (b) the proper use of arithmetic (e.g., +, -) or logical operators (e.g., >, and, or) when calculating the spatial relations and (c) the number of logical errors within the algorithms that concern the hand/body gesture-based interaction.

Apart from the assessment of the particular skills, it is suggested to evaluate the degree of promoting positive learning experiences. Thus, the proposed assessment framework has been enriched with well-known practices (pre-post structured questionnaires) adopted from relative studies (Baytak, 2009; Cheng, 2009) and focus on assessing the degree of promotion of positive feelings, attitudes about programming and game design and social skills, and so on.

4. Evaluation case study

4.1. Context

A case study was performed with the participation of undergraduate computer science students at their last year of their studies, with the aim to validate the effectiveness of the proposed assessment framework. Students followed a systematic process (Authors of this study, 2018), in order to design, develop and self-assess the quality of a prototype of motion-based touchless games during 11 weeks lab-based course. Students had face to face collaborative sessions of 2 h per week at the lab.

4.2. Participants

The total number of participating undergraduate students was 42 (36 men, 6 women). All participants were given from the start the option to work either individually or in groups (of two members). Thus, six students preferred to work alone and eighteen (18) consisted of two (2) members were formed.

5. Results

This section presents indicative results about the assessment of the performance of two participated groups of students, with the aim to clarify aspects of the application of the proposed assessment framework.
5.1. Results about the promotion of skills of game elements analysis and design

Table 3 summarises the results from the assessment of the game analysis and design skills of two groups. Both groups very actively participated during in the design phase. The first group (group-4) completed all GDD fields, while the deliverable of the second group (group-9) had some gaps (index 1.1.1). Apart from this, the clarity of 14 game elements in GDD was quite high, reaching to 92.86% and 71.43, respectively (index 1.1.2).

Furthermore, both groups managed to visualize all the parts of their game in the storyboard (index 1.2). At the same time both groups enrich game design (index 1.3), by putting emphasis on storyboard analysis via the dramatic elements of their game such as the element of play or how the story develops over time. Finally, the assessment of these deliverables showed that both groups achieved to analyse and design complex systems like the one of the digital game, meeting the qualitative criteria of Felder’s rubric (index 1.4).

5.2. Results about the promotion of CT skills

Table 4 below gives, for each group separately, the analytical and average CT score, as it exported from Dr. Scratch tool and the analysis of students’ answers in their CT journal by using the related assessment CT rubric.

The results from Dr. Scratch showed that both groups incorporate the required conditions (blocks) in their games, which are directly associated with CT concepts and received a quite high score in the Dr. Scratch tool (94.74%). The analysis of the individual criteria of the CT journal’ rubric in total showed that the criterion with the lowest average score (65.63%) regarded the description of the degree of reusing and remixing others’ material. Therefore, it is ascertained that these groups did not utilize the code and elements from other games to create their own game or (for reasons not mentioned) they did not record it in their journals.

Table 3. Snapshot of the results about the promotion of skills of game elements analysis and design of two groups.

| Criteria                                                                 | Measurement indicators                                                                 | Students/Group 4 | Students/Group 9 |
|-------------------------------------------------------------------------|----------------------------------------------------------------------------------------|-----------------|-----------------|
| 1.1 The ability to break the game into smaller parts (14 formal &        | 1.1.1 GDD fields completion rate (%)                                                  | 100.00%         | 85.71%          |
| dramatic elements) analysing them on a GDD                              | 1.1.2 The clarity of 14 game elements in GDD (%)                                       | 92.86%          | 71.43%          |
| 1.2 The ability to visualize, in the form of animated scripts,          | 1.2 GDD’ elements visualized in the Storyboard (%)                                     | 100.00%         | 100.00%         |
| (henceforward Storyboard) the elements of the GDD                      |                                                                                        |                 |                 |
| 1.3 The ability to enrich game design process using a Storyboard        | 1.3 GDD’ elements were visualized in the Storyboard, providing more information versus | 30.77%          | 46.15%          |
|                                                                         | GDD’ elements description (%)                                                         |                 |                 |
| 1.4 The ability to analyse and design a digital game, meeting specific | 1.4 Meeting quality criteria for the deliverables of the Design Phase (GDD and       | 95.24%          | 90.48%          |
| quality criteria during the design phase                                | Storyboard)                                                                            |                 |                 |
Table 4. Results about the promotion of Computational Thinking skills.

| Criteria                                                                 | Measurement indicators                                                                 | Students/Group 4 | Students/Group 9 |
|--------------------------------------------------------------------------|----------------------------------------------------------------------------------------|------------------|------------------|
| 2.1 The ability to use command blocks (in the form of bricks) in their game’s code, which are directly associated with computational thinking skills concepts | Dr. Scratch Results (scale: 1–3 points for each CT concept)                             |                  |                  |
| 1.1 Flow control                                                         | 3                                                                                     | **94.74%**       | 3                |
| 1.2 Data representation                                                  | 2                                                                                     | 2                |
| 1.3 Abstraction                                                          | 1                                                                                     | 1                |
| 1.4 User Interactivity                                                   | 3                                                                                     | 3                |
| 1.5 Synchronization                                                      | 3                                                                                     | 3                |
| 1.6 Parallelism                                                          | 3                                                                                     | 3                |
| 1.7 Logic                                                               | 3                                                                                     | 3                |
| 2.2 The ability to apply computational thinking practices during creating their digital game | CT Journal Results (scale: 3–9 points for each CT practice)                            |                  |                  |
| 2.1 Experimenting & iterating: 4–12 points                              | 8                                                                                     | **66.67%**       | 10               |
| 2.2 Testing & debugging: 4–12 points                                     | 8                                                                                     | **66.67%**       | 10               |
| 2.3 Reusing & remixing: 4–12 points                                      | 7                                                                                     | **58.33%**       | 7                |
| 2.4 Abstracting & modularizing: 3–9 points                               | 6                                                                                     | **66.67%**       | 8                |
5.3. Results about the promotion of programming skills

The criteria for programming skills are evaluated according to the following indexes in Table 5. The analysis of the code of the digital games showed that all groups utilized code blocks from the seven (7) command categories of MIT Scratch (Index 3.1).

The results from the automatic source code analysis with the Scrape tool showed that both groups used blocks from each one of the seven (7) categories (Index 3.1). In addition, the students’ ability to apply very good programming practices (index 3.2) was calculated by combining data from Dr. Scratch & Scrape tool (see Table 6). According to Table 6, both groups applied the renaming and the dead code practice.

In addition, the results from the analysis of the quality of the source (12 Coding Quality Criteria) are presented in Table 7.

According to Table 7, for the first group (group 4) no problems have indicated by Quality Hound tool (10/12 criteria, 0.00%) thus scoring 10/12 = 83.33% (index 3.3) while the other group (group 9) score a bit lower for reasons that had been explained to them, i.e., they created unused variables.

Overall, the 24 games had been peer-evaluated using a proposed 5-Likert scale assessment rubric (index 3.4). These rubric’s criteria are:

1. Instructions
2. Unique/clear goal, simple & few rules
3. Challenge
4. Feedback

Table 5. Criteria and indexes for the evaluation of students’ performance in programming skills.

| Criteria | Measurement indicators |
|----------|-------------------------|
| 3.1 The ability to implement their game by using different types of commands | 3.1 The use of different types of code blocks (Control/Look/Sensing/Sound/Operators/Variables/Motion) (%) |
| 3.2 The ability to use best practices in the usage of code | 3.2 Coding best practices problems (dead code, sprite attributes, sprite naming, and variable naming) (%) |
| 3.3 The ability to develop a game code that meets specific quality criteria | 3.3 ”Coding Quality criteria” problems (12 code smells) |
| 3.4 The ability to develop a “good” game | 3.4 The existence of common design heuristics in the game |

Table 6. Coding best practices problems (Index 3.2).

| Coding best practices problems (Dr. Scratch results) | Students/Group 4 | Students/Group 9 |
|-----------------------------------------------------|------------------|------------------|
| Duplicated scripts | 2.98% | 0.55% |
| Sprite naming | 0.00% | 0.00% |
| Dead code | 0.00% | 0.20% |
| Sprite attributes | 34.38% | 57.14% |
From the qualitative analysis, it was found that the 24 games covered the total of the eight (8) proposed design principles at the average of 71.74%.

5.4. Results about the promotion of spatial thinking skills

The participants successfully conceptualized and implemented gesture-based interaction for the needs of their motion-based touchless games. Results about the performance of the aforementioned two groups are provided below:

- The gestures’ intricacy of the first group (group-4) is lower than the one of the second groups (group-9) i.e.,: (a) the number of gestures were only 2 versus 5 of the second group, (b) the number of body joints used in gesture’ algorithms were 3 versus 7 and (c) the number of coordinate axes used to compare body joints were 2 (y, z axes) versus 3 (x, y, z axes).
- With regards to the quality of implementation, the performance of the first group (group-4), is higher than the one of the second group (group-9) due to the better use of CT concepts (wait until vs if block) and the absence of logical errors within the algorithms gesture-based interaction. On the other hand, on the second’s group gesture algorithms we identified that 3 to 5 gestures had logical errors or no consistency between the algorithms and the gestures executed during playtesting. These results draw conclusions about the lack of programming skills for the second group (group-9).

| Quality Hound (% problems)                  | Students/Group 4 | Students/Group 9 |
|--------------------------------------------|------------------|------------------|
| 1. Broad Variable Scope                    | 42.86%           | 12.50%           |
| 2. Duplicated String                       | 0.00%            | 0.00%            |
| 3. Duplicated Code                         | 5.30%            | 0.00%            |
| 4. Feature Envy                            | 0.00%            | 0.00%            |
| 5. Inappropriate Intimacy                  | 0.00%            | 0.00%            |
| 6. Long Script                             | 0.00%            | 0.00%            |
| 7. Middle Man                              | 0.00%            | 0.00%            |
| 8. No-Op                                   | 0.00%            | 0.00%            |
| 9. Uncommunicative Name                    | 0.00%            | 0.00%            |
| 10. Undefined Block                        | 0.00%            | 0.00%            |
| 11. Unreachable Code                       | 0.00%            | 0.74%            |
| 12. Unused Variable                        | 0.00%            | 25.00%           |
| Total Score                                | 10/12 = 83.33%   | 9/12 = 75.00%    |
5.5. Results about the promotion of positive learning experiences

Students had very positive feelings about their digital game design and development projects. In the question “Successful completion of my own game made me feel satisfied” 87% of the students answered: “A lot” and “Very much”. In the question “What comes to your mind when I ask you about game design?”, the majority of students mentioned characteristics they had not mentioned at the beginning of the project e.g., “it is a process of resolving a complex problem”, “it is creative”. Likewise, they enhanced their perception of programming as, after the end of the project, they associated it with key elements like cooperation, CT skills, creativity, abstraction, troubleshooting and problem-solving. Finally, 70% of the students believes that the digital Kinect-based game creation project helped them very much to understand geometry and human-computer interaction issues by implementing natural interaction algorithmic structures, in contrast with 5% and 0% of the students who selected “A little” or “Not at all”.

Moreover, 70% of the students believes (a lot and very much) that they enhanced their ability to think of different solutions during the creation of their digital game. The analysis of the natural interaction code shows that 19 out of 24 groups created at least one (1) new NUI algorithmic structure, which did not come from the available learning material of the project.

Prior to the project, 27.50% of the students answered that they do not ask for any help from their fellow students or they do it very rarely. After the end of the project, the percentage of students that answered that they did not ask for help from their fellow students was reduced to 12.50%. On the contrary, 87.50% of the students asked for help during the project. In the question “Did the help you received solve your problem?”, 27 out of 35 students (77.14%) answered “a lot” and “very much”.

6. Conclusions

Several researchers have mentioned about the need for further research in finding ways to assess in authentic way students’ knowledge and skills when they got engaged in game-making projects. This paper presents a multifaceted framework of assessing the students’ performance during the whole lifecycle of a digital game-making process.

The enactment of this framework revealed some helpful findings:

- Regarding the assessment tools used, it is worth pointing out that they were adopted from the literature, in order to ensure that the produced scores will be reliable, valid and fair, with the exception of the tools for measuring spatial thinking skills. Specifically, we used automated assessment tools for the quality of the source code and rubrics for quantitative analysis. It was
easy and straightforward to apply these tools which were in line with the statements of their creators.

- The proposed assessment framework can help teachers in estimating whether students have cultivated spatial thinking skills in a systematic way. It is the first time teachers could systematically gather in-depth data regarding students’ ability to comprehend the coordinate axes, the position and distance of their body joints and the objects around them within arm’s reach and at the same time to indicate dimensions of spatial thinking, through motion-based touchless game-making activities. The assessment of spatial thinking skills through code analysis of their gestures’ algorithms though innovative gave a clear understanding of students strengths and weaknesses in understanding how to reason and calculate the spatial relations among body joints in space, in order to interact with game elements in a virtual environment.

- Due to a large number of assessment indexes, techniques and tools, the proposed framework requires about ninety (90) minutes to assess all the deliverables per group and to draw final results about the students’ performance, taking into consideration that the assessment is being made by only one (1) person (e.g., a teacher). Although this process can turn out to be time-consuming, it is clear that in order to perform an authentic and ongoing assessment the need of an in-depth data analysis is necessary.

- The fact that students involved in a peer-assessment process is already mentioned by other studies as an effective learning strategy (Hwang, Hung, & Chen, 2014). By using this framework, during the game creation process, in a self-assessment and peer-assessment process, is a testimony of user-friendliness, while at the same time helps them to comprehend in depth the assessment requirements and to create “good games”.

In this paper, we do not argue that students who involved in such game-making activities will certainly cultivate to large extent multiple higher order skills, but we support that in order to draw safe conclusions about the level of multiple skills improvement we should use a combination of assessment tools, as proposed by the framework. Future directions of this research include (i) the examination of the effectiveness of the proposed framework in other course settings, and (ii) the development of an e-portfolio tool to support the submission of the game-making process, the grading by embedding the assessment tools and automated calculations of the scores at the various performance indexes.

### Compliance with ethical standards

The authors declare that all procedures performed in this study that involved human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.
**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This study has not received any funding from any source.

**Open Access**

This paper is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**References**

Aivaloglou, E., & Hermans, F. (2016). How kids code and how we know: An exploratory study on the scratch repository. In Proceedings of the 2016 ACM Conference on International Computing Education Research, New York, NY, USA. (pp 53–61).

Akcaoglu, M. (2016). Design and implementation of the game-design and learning program. *TechTrends, 60*(2), 114–123.

Altanis I., Retalis S., & Petropoulou O. (2018). Systematic design and rapid development of motion-based touchless games for enhancing students’ thinking skills. *Education Sciences, 8*(1): 18.

Barcelos, T. S., Costa, G., Muñoz-Soto, R., Noël, R., & Silveira, I. F. (2013). Informal HCl: What may students learn from playability issues during a game design workshop?. *ACM international conference proceeding series* (pp. 116–119). doi:10.1145/2535597.2535613.

Baytak, A. (2009). An investigation of the artifacts, outcomes, and processes of constructing computer games about environmental science in a fifth grade science classroom (PhD Thesis). The Pennsylvania State University.

Bekesi, D., & Sik-Lanyi, C. (2016, September 20–22). Kinect controlled game to improve space and depth perception. In P. M. Sharkey & A. A. Rizzo (Eds.), *Proc. 11th international conference series on disability, virtual reality and associated technologies* (pp. 267–270). Los Angeles, CA.

Burke, Q., & Kafai, Y. B. (2012, February). The writers’ workshop for youth programmers: Digital storytelling with scratch in middle school classrooms. In Proceedings of the 43rd ACM technical symposium on Computer Science Education, SIGCSE ’12 (pp 433–438). ACM: New York.

Caci, B., Chiazzese, G., & D’Amico, A. (2013). Robotic and virtual world programming labs to stimulate reasoning and visual-spatial abilities. *Procedia - Social and Behavioral Sciences, 93*, 1493–1497.

Cheng, G. (2009). Using game making pedagogy to facilitate student learning of interactive multimedia. *Australasian Journal of Educational Technology, 25*(2), 204–220.

Claypool, K., & Claypool, M. (2005, September). Teaching software engineering through game design. *SIGCSE Bulletin, 37*(3), 123–127.

Denner, J., Werner, L., & Ortiz, E. (2012). Computer games created by middle school girls: Can they be used to measure understanding of computer science concepts? *Computers & Education, 58*(1), 240–249.
Desurvire, H., & Wiberg, C. (2009, July 19–24). Game usability heuristics (play) for evaluating and designing better games: The next iteration. In OCSC ’09 Proceedings of the 3rd international conference on online communities and social computing: Held as part of HCI international 2009 (pp. 557–566). San Diego, CA: Springer.

Felder, R. (2011). Writing for the Web: Creating compelling web content using words, pictures, and sound. Thousand Oaks, CA: 1st New Riders Publishing. ©2011. ISBN:0321794435 9780321794437. Retrieved from http://www.write4web.com/wp-content/uploads/2011/10/My.Rubric.storyboard_design.pdf

Fields, D. A., Kafai, Y. B., Strommer, A., Wolf, E., & Seiner, B. (2014, August 19–23). Interactive storytelling for promoting creative expression in media and coding in youth online collaboratives in Scratch. In Proceedings of Constructionism. Vienna, Austria.

Foerster, K. (2017, April 25–28). Teaching spatial geometry in a virtual world: Using minecraft in mathematics in grade 5/6. 2017 IEEE Global Engineering Education Conference (EDUCON) (pp. 1411–1418). Athens, Greece.

Francis, K., Khan, S., & Davis, B. (2016). Enactivism, spatial reasoning and coding. Digital Experiences in Mathematics Education, 2(1), 1–20.

Fullerton, F., Swain, C., & Hoffman, S. (2004). Game design workshop: Designing, prototyping, and playtesting games. San Francisco, CA: CMB Books.

Grover, S., Pea, R., & Cooper, S. (2015). Systems of assessments” for deeper learning of computational thinking in K-12. Presented at the annual meeting of the American Educational Research Association. Chicago, IL.

Gutierrez, F. J., Simmonds, J., Hitschfeld-Kahler, N., Casanova, C., Sotomayor, C., & Araya, V. P. (2018, May 27–June 03). Assessing software development skills among K-6 learners in a project-based workshop with scratch. 2018 IEEE/ACM 40th International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET) (pp. 98–107). Gothenburg, Sweden.

Hava, K., & Cakir, H. (2017, June 20). A systematic review of literature on students as educational computer game designers In J. Johnston (Ed.), Proceedings of the EdMedia: World conference on educational media and technology, Waynesville, NC (pp. 407–419). Washington, DC: Association for the Advancement of Computing in Education (AACE).

Hwang, G. J., Hung, C. M., & Chen, N. S. (2014). Improving learning achievements, motivations and problem-solving skills through a peer assessment-based game development approach. Educational Technology Research and Development, 62(2), 129–145.

Ilic, U., Haseski, H. I., & Tugtekin, U. (2018). Publication trends over 10 years of computational thinking research. Contemporary Educational Technology, 131–153. doi:10.30935/cet.414798

Kafai, Y. B., & Burke, Q. (2015). Constructionist gaming: Understanding the benefits of making games for learning. Educational Psychologist, 50(4), 313–334.

Lockwood, J., & Mooney, A. (2017). Computational thinking in education: Where does it Fit? A systematic literary review. International Journal of Computer Science Education in Schools, 2(1): 41–60. Retrieved from http://ijcises.org/index.php/ijcises/article/view/26.

Lux, N., LaMeres, B. J., Hughes, B. E., & Willoughby, S. D. (2018, June), Board 82: Increasing the spatial intelligence of 7th graders using the minecraft gaming platform. Paper presented at 2018 ASEE Annual Conference & Exposition. Salt Lake City, Utah. Retrieved from https://peer.asee.org/30116

Maike, V. R. M. L., de Sousa Britto Neto, L., Baranauskas M. C. C., & Goldenstein S. K. (2014) Seeing through the Kinect: A survey on heuristics for building natural user interfaces environments. In Stephanidis C., & Antona M. (Eds.), Universal Access in Human-Computer
interaction. Design and Development Methods for Universal Access. UAHCI 2014. Lecture Notes in Computer Science, (vol. 8513). Cham: Springer.

Ministry of Education in Ontario. (2014). Paying attention to spatial reasoning, K-12: Support document for paying attention to mathematical education. Queen’s Printer for Ontario. ISBN 1460638492, 9781460638491. Retrieved from http://www.edu.gov.on.ca/eng/literacy/numeracy/inspayingattention.pdf

Moreno-León, J., & Robles, G. (2015). Analyze your scratch projects with Dr. Scratch and assess your computational thinking skills. Amsterdam, The Netherlands: AMS.

Moreno-León, J., Román-González, M., Harteveld, C., & Robles, G. (2017). On the automatic assessment of computational thinking skills: A comparison with human experts. In Extended Abstracts of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI’17) (pp. 2788–2795). New York, NY: ACM. Doi:10.1145/3027063.3053216

National Research Council. (2006). Learning to think spatially: GIS as a support system in the K–12 curriculum. Washington, DC: National Academic Press.

Rankin, Y., Thomas, J., & Irish, I. (2014). From consumers to producers: African american middle school students as game designers. Spelman College Faculty Publications.

Román-González, M., Moreno-León, J., & Robles, G. (2017, June 26–30). Complementary tools for computational thinking assessment. In S. C. Kong, Sheldon, J. & Li K. Y. (Eds.), Proceedings of International Conference on Computational Thinking Education (CTE 2017) (pp. 154–159). Lviv, Ukraine: Lviv Polytechnical National University.

Romero, M., Lepage, A., & Lille, B. (2017). Computational thinking development through creative programming in higher education. International Journal of Educational Technology in Higher Education, 14(1), 42.

Schack, G. D. (1994). Authentic assessment procedures for secondary students’ original research. The Journal of Secondary Gifted Education, 6(1), 38–43.

ScratchEd Harvard online community. (2017). Scratch computational thinking journal & assessing development of computational practices. Retrieved from http://scratched.gse.harvard.edu/ct/assessing.html

Sinker, C. (2014). Creating an interactive learning resource for teaching scratch using the kinect for students throughout key stages 2 – 4. School of Computing at the University of Leeds. Retrieved from https://minerva.leeds.ac.uk/bbcswbdav/orgs/SCH_Computing/FYProj/reports/1314/Sinker.pdf

Sullivan, K., Byrne, J. R., Bresnihan, N., O’Sullivan, K., & Tangney, B. (2015). CodePlus — Designing an after school computing programme for girls. 2015 IEEE Frontiers in Education Conference (FIE) (pp. 1–5)El Paso, TX. doi: 10.1109/FIE.2015.7344113

Techapalokul, P., & Tilevich, E. (2017). Quality Hound - An online code smell analyzer for scratch programs. IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC) (pp. 337–338). Raleigh, NC. doi: 10.1109/VLHCC.2017.8103498

Tsai, C., Kuo, Y., Chu, K., & Yen, J. (2015). Development and evaluation of game-based learning system using the microsoft kinect sensor. International Journal of Distributed Sensor Networks, 11, 498560: 1–498560:10.

Wilson, A., Hainey, T., & Connolly, T. (2012). Evaluation of computer games developed by primary school children to gauge understanding of programming concepts. Retrieved from https://pdfs.semanticscholar.org/436c/852c4094a58ff9e629ace758612f58459342.pdf

Yu, A. B., & Zacks, J. M. (2015). How are bodies special? Effects of body features on spatial reasoning. Quarterly Journal of Experimental Psychology, 69(6), 1210–1226.