Università Politecnica delle Marche
Scuola di Dottorato di Ricerca in Scienze dell'Ingegneria
Curriculum in Ingegneria informatica, gestionale e dell'automazione
----

Study, assessment and identification of Educational Robotics experiences at school

Ph.D. Dissertation of:

Laura Screpanti

Advisor:

Prof. David Scaradozzi

Curriculum supervisor:

Prof. Francesco Piazza

XVIII edition - new series

Università Politecnica delle Marche
Scuola di Dottorato di Ricerca in Scienze dell'Ingegneria
Curriculum in Ingegneria Informatica, Gestionale e dell'Automazione

---

Study, assessment and identification of Educational Robotics experiences at school

Ph.D. Dissertation of:

Laura Screpanti

Advisor:

Prof. David Scaradozzi

Curriculum supervisor:

Prof. Francesco Piazza

XVIII edition - new series

Università Politecnica delle Marche
Dipartimento di (nome del dipartimento dove la tesi e' stata sviluppata)
Via Brecce Bianche — 60131 - Ancona, Italy

Acknowledgements

No man is an island entire of itself; every man
is a piece of the continent, a part of the main
(John Donne, Meditation XVII
Devotions upon Emergent Occasions)

For what I achieved in this thesis is part of long journey begun in the past years and, hopefully, doomed to continue in the future. I would not have been able to achieve so much without the support, the open debates and the critical reviews of the scientific community which I belong to.

Most of all, I would like to express my special appreciation and thanks to my advisor, Professor David Scaradozzi, who has unceasingly provided guidance, advice and example.

I wish to thank my colleague, Ing. Lorenzo Cesaretti, for the witty conversations spanning from Educational Robotics to new sports and food, for sharing passion and curiosity.

A special thanks to my family, friends and boyfriend. Even though you could not imagine what was going on in my head and in my professional life, you chose to stand by my side, no matter how low my feelings were or how unnecessarily happy I was. Thank you.

Abstract

The aim of this thesis is to study and model the learning process of K12 classrooms during Educational Robotics (ER) activities. Approaching the study of learning in K12 classrooms during ER activities for modelling purpose, the first question to answer is what kind of variables are identified as representative of learning in the given context; subsequently, the causality principle between input and output should be demonstrated; finally, variables should be observed to provide a measure which could be related to the internal state of the system. In this work answers to these questions are explored by choosing the definition of learning as the process of acquiring knowledge, competence and abilities, by relating ER activities carried out in K12 classroom (procedures) with changes in those variables and by formulating and validating instruments to capture quantitative information about those variables. Some studies reported in literature present a model of learning for different context, none in the ER field. Moreover, in assessing ER activities there are fewer quantitative studies than qualitative studies, and ER itself lacks a formal definition of what it exactly entails. In the present work four case studies are presented. Each case study reports an ER experience in a K12 set of classrooms which employed the developed sensors to measure the outcome of the experience. After testing "traditional sensors" of learning, namely the state-of-the-art procedure to validate those instruments, some final considerations are provided on the metrology and causality issues for modelling.

Contents

List of Figures

List of Tables

Introduction

The aim of this thesis is to study and model the learning process of K12 classrooms during Educational Robotics (ER) activities. Several researchers tried to shape the learning process as a dynamical system (Abbona, Del Re and Monaco, 2008; Gattie, Kellam, Schramski and Walther, 2011; Jacobson, Levin and Kapur, 2019; Nicolescu and Petrescu, 2013; Piech, Sahami, Koller, Cooper and Blikstein, 2012; Steenbeek and van Geert, 2013; White, Smigiel and Levin, 2017); most of their studies does not have grounding in the experimental field, and their focus is on learning in general; models are formulated starting from a theoretical approach and end stating the need to be implemented. Acknowledging this gap several institutions are investing on the research field of education as a complex system. In Italy, for example, the national institute INVALSI created the research group on Education as a Complex System, called “E-CAS”, with the aim of collecting and analysing more information, scarce at the moment, to understand the educational system on the basis of the socio-cultural context of society and of the single student in a classroom (INVALSI, 2006).

The study of the dynamic evolution of a student’s learning is challenging because, even if the concept of “learning” is part of common knowledge, its scientific definition is still not well acknowledged. In fact, it is possible to define “learning” from several points of view - neurobiological, psychological and pedagogical for instance - but in order to measure and to quantify “learning” a more formal and unifying concept must be used. Research in education proposed some models of learning and identified some variables pertaining to learning; for example, the output of the learning process is often defined in terms of learning objectives of a certain activity. According to national and international laws, K12 education should develop knowledge, skills, behaviours and attitudes (European Union, 2006; Mazzer, 2018; MIUR, 2012). Like the educational activity, also the modelling process is required to focus on a set of aspects of reality; it is required to define the learning processes at play and which of them are taken into consideration in the model of the reality.

Another aspect to keep into consideration while building a model is the causality principle. Classical systems theory is an example of axiomatic theory of dynamic phenomena that verify the causality principle (i. e. given an input, the system produces an output). To examine whether this principle holds for the variables involved in the learning process, the causal relation needs to be proven by means of sound scientific findings. In the educational, psychological and sociological fields, tools and methods have been examined through the lens of scientific method only recently, if compared with other sciences like Physics, Astronomy, Biology, etc. . Many theories in those fields define abstract elements, like self-efficacy (Bandura, 2010) or computational thinking (Brennan and Resnick,2012; Wing, 2006, 2008); this abstract concepts are the building elements (the variables) of the experiment whose aim is to verify the causal relationship between input and output; moreover, these building elements can vary across time, increasing or decreasing.

Another element to keep into consideration while building a model of learning is the measurement itself of the accessible variables. The "observed variables" are those variables that can be accessed by means of "sensors", while the "latent variables" are those variables that cannot be directly measured. By means of repeated observations (samples) of an accessible variable, which is a repeatable phenomenon, it is possible to have a numerical evaluation of the associated "latent variable", which is a phenomenon comparable to the observed variable. In order to establish causal relationships among variables and to provide valid and reliable measurement of the intended constructs, research in the education field found its scientific soundness basing its claims on philosophical stances. Stating the theoretical stance a researcher can be able to derive a sound methodology of enquiry, establishing a point of view through which one can see the reality and thus defining how to measure the variables of interest (Figure 1).

The last phase of modelling learning is the one that refers to "cybernetics", the art of controlling the system (Wiener 1948) and deals with providing the output of the "sensors" measuring learning as an input to the "actuator" of a correction strategy. Actuators are those elements in the system in charge of transforming the measures received as an input into commands to the system under control in order to make the observed variables closer to the reference objectives established. This feedback mechanism processes information on the past/present measurements to determine the future behaviour of the systems.

	Level		Contrasting stances	
Theoretical stance	**Ontology** Beliefs about the nature of being or reality		There is one objective reality	There are multiple realities
	Epistemology Belief about the nature and scope of knowledge (how we come to know the world)		You uncover the reality – there is one true explanation	Meaning is culturally defined
Approach	**Methodology** Based on paradigmatically different ontological and epistemological assumptions		**Quantitative** Positivist, Objectivist, Empiricist, Nomothetic	**Qualitative** Hermeneutic Interpretivist
	Design		Overarching strategy for collecting data, such as: Experimental, Quasi-experimental, Random Controlled Trials	Case study, Action research, Ethnography
	Emphasises		deductive reasoning	inductive reasoning
	Data (numerical or non-numerical)	**Methods**	Techniques for collecting data, such as: Survey/questionnaire; Interview/Focus group; Document analysis; Observation	
		Instruments	Specific data collection tools, such as: a specific questionnaire or interview schedule	
		Analysis	How the data are processed in order to make sense of them (to answer your research questions)	

Figure 1: quantitative and qualitative approaches deriving from contrasting theoretical stances (Jacobson, Levin and Kapur, 2019).

The first phase of this study reflects on the modelling of a frontal lesson: a teacher explaining concepts to a learner. Ideally, information flows from the teacher to the learner, the teacher provides the learner with a reference (Figure 2). Subsequently, the teacher evaluates how this information has been acquired by the student through the evaluation of knowledge, skills, attitudes and competences.

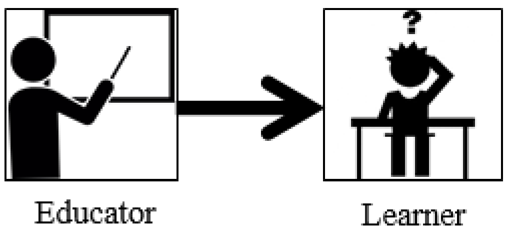


Figure 2: scheme of a frontal lesson.

The concept of providing the reference input to the system can also be explained through a block diagram as expressed in Figure 3. This simple way of expressing the model pose some simple questions: what is r(t)? what is y(t)? Can they be related and transformed into quantitative variables? If so, which kind of instruments are capable of sensing the variable of interest and transduce it into something measurable and understandable and usable into a cybernetic system?

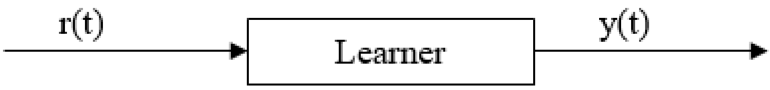


Figure 3: block diagram of the system (learner) in the open loop.

To answer these questions a typical measuring chain can be used to represent the sensors needed (Figure 4). Firstly, a primary sensitive element captures the variable of interest, then, the variable of interest is transduced and amplified adequately to be processed and put into a comprehensible format for a teacher. This information can be used by teachers to adjust their “control strategy”, to make his or her teaching strategy more effective. The metrology issue poses serious questions about the nature of the variables involved as well as how to represent them.

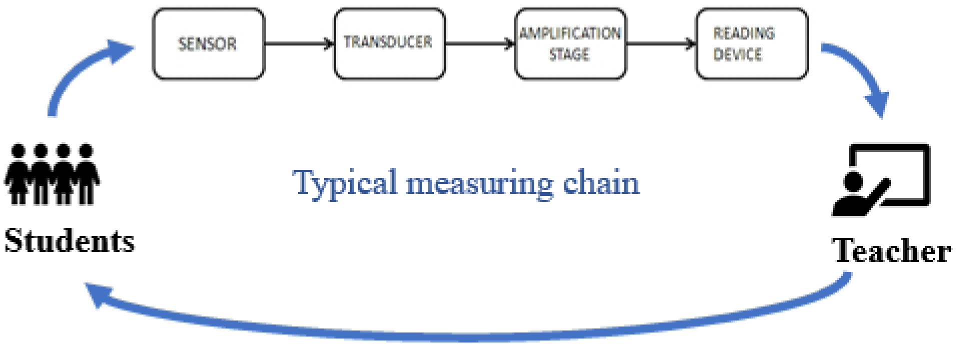


Figure 4: schematic representation of a sensor of learning.

Focusing on the ontological and epistemological basis, the approach carried out in the present thesis is the positivist/logical empiricist. The empirical scientific method is empirical, namely related to experience, and logic, namely it is possible to consistently apply logic to make sense of the world, not trust, not authority, neither belief (Figure 5).

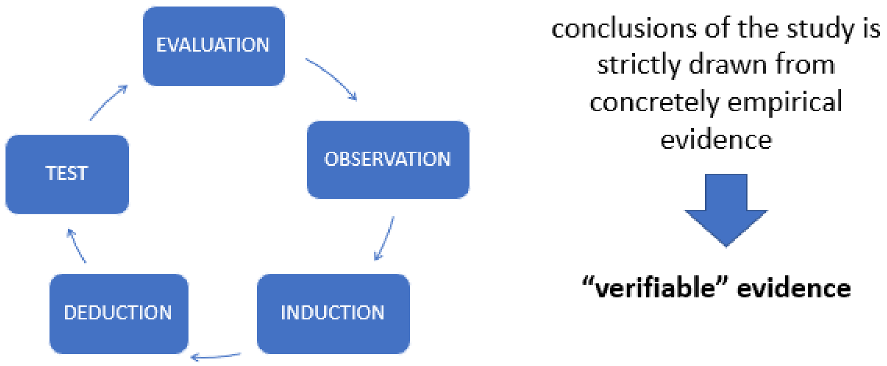


Figure 5: the scientific cycle.

From these stances it derives an approach to measurement in education science: tools and methods must be tested for their validity and for their reliability. Validity implies an evaluation not only of the tools but also of methodologies applied during the experiment (research design). In measuring variables, it must be clear that the results are derived only from the object of the study and that the results depend entirely on the variables of interest and not from other unwanted variables (noise). To evaluate the validity of a research it should be taken into consideration if the results are a useful model to account for reality adding new elements to understand it and represent it. Four aspects to examine the validity of a study:

- internal validity
- construct validity
- external validity and ecological validity
- statistical validity

Each one of these types of validity can be threatened by situation that may occur usually both in a laboratory experiment and in the real-life scenario. For this reason, to formulate a measure experiment of the variable y(t) it should be considered:

- research question
- research design

- tools to measure the variable of interest, tested for validity and reliability
- statistical validation of the research hypothesis

Thus, in the present work the focus is on the variable of interest (those regarding learning) and the development of "sensors" to capture those data. Which sampling interval to choose to better represent the trajectory of learning is a fundamental question. Two approaches can be developed: a discrete and "classic" approach, based on tools like questionnaires and test, and a discrete and real-time approach (Berland, Martin, Benton, Petrick Smith and Davis, 2013; Cesaretti, Screpanti, Scaradozzi and Mangina, *in press*; Jormainnen and Sutinen, 2012), based on the analysis of log files from students' programming (Scaradozzi, Cesaretti, Screpanti and Mangina, *in press*; Scaradozzi et al., 2018). Both these two approaches are based on an intervention in the classroom, both these approaches are discrete measurements of students' learning trajectory. "Classic" sensors can be developed to test any variable being it knowledge, competence, skills or attitude. At the moment "real-time sensors" provide feedback on students' learning trajectory in terms of styles of problem-solving and they can provide feedback on students' way of programming robots or apps. "Real-time sensors" acquire data from students' processes with a shorter sampling period than "classic sensor". Moreover, "real-time" sensors are transparent to students while programming, while "classic" sensors can be used only with an active participation of the student. In the present thesis the focus is on classic sensors.

Capturing information from the system as described by Figure 3 can lead to two interesting outcomes: the characterisation of the system "learner" and the closed loop chain of control (Figure 6). The first will be useful because it will help understand the profile of a student while learning through Educational Robotics; the second will be useful because the information acquired with sensors can be provided back to teachers to support them in the evaluation process or in adjusting the way they teach to each student.

The former aspect is the focus of the present thesis, the latter lead us to a final series of thoughts. To face the second aspect in fact it is important to assign a quantity to r(t), which is a reference signal to the system and as such it is an arbitrary choice of the modeler. It should be noted that, even if it is fundamental from the point of view of the control theory to arrive to a standard r(t), rules for the validity of tools still hold. It is right then to trust experts, namely people trained and competent to "deliver" the content.

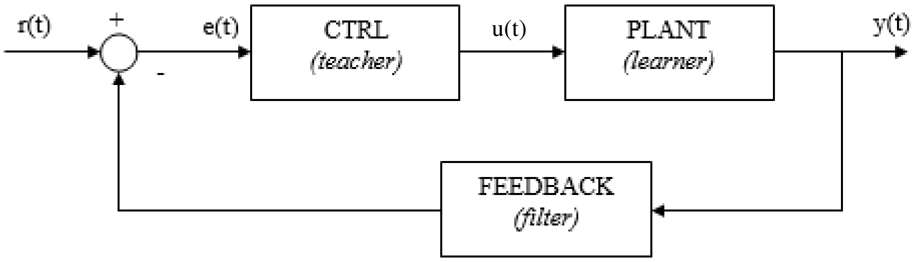


Figure 6: block diagram of the system (learner) in the closed loop.

To better understand what it means to quantify r(t) it is possible to make an example. A single activity can be defined in terms of information passed down from teacher to student (i.e. activity on "sensor in a robotic system" has the aim to bring the classroom to a mean level of 7/10 on a scale which has been a priori validated). The teacher can start the activity and pass down the concepts, develop skills and so on (they all constitute u(t)). At the instant t* the teacher can administer the validated sensors to measure the level of the information retained by the student, which is the output of the system "learner", namely y(t). A proper filter from y(t*+dT) should extrapolate information comparable with the dimension of r(t) and with the chosen reference level. In this design, the teacher (the controller of Figure 6) can choose the action to perform on the system: keep on adjusting its own output with corrections (u(t*+dT)) depending on e(t*+dT) or stop. The model of the process will keep on representing reality as long as the teacher will not close the temporal axis (t_f) and all its sub-Phases. The choice of adopting the scheme reported Figure 6 only for t_0 and t_f in the activity or to give pre-eminence to single sub-phases within that period represent the sampling time of the system.
Limits to the linear dynamical model can be found in the way reality is modelled and in the emergent and nonlinear properties of the single components of the systems. It remains the problem of how to represent and quantify the variables of interest at play.
In the following sections these considerations are explored starting from Section 1, which will show the context of the present work, namely Educational Robotics and measurement of learning for modelling purposes. Section 2 will present some sensors to measure the variables of interest and their validation. Section 3 will demonstrate how these sensors were used in

the K12 classrooms during ER experiences. Finally, Conclusions will provide final remarks.

Section 1 – Context and state of the art

The context of the present work of thesis is Educational Robotics and systems theory applied to learning. The next subsection will provide details on the main features of Educational Robotics. It cannot be formulated a model of reality without observing it, thus information on relevant studies describing Educational Robotics is provided. It should be pointed out that it does not exist at the moment a unifying vision on Educational Robotics. Literature reports several different stances and advocates for a definition of Educational Robotics (Alimisis, 2013; Alimisis, Alimisi, Loukatos and Zoulias, 2019; Angel-Fernandez and Vincze, 2018; Benitti, 2012; Eguchi, 2014; Mikropoulos and Bellou, 2013; Toh, Causo, Tzuo, Chen and Yeo, 2016). The subsequent subsection will report literature on how to represent learning. Notably, no research paper was found on how to represent learning in a constructionist environment for Educational Robotics, but all of these papers mention the vital need for adequate measurements in terms of philosophical stance, methodology of research and tools for measuring intended construct (Benitti, 2012; Castro et al., 2018; López, Valenzuela, Nussbaum and Tsai, 2015).

1.1 Educational Robotics

Robots in education are increasingly used as tools to promote the acquisition of knowledge, disciplinary and cross-disciplinary competencies, attitudes and skills, including but not limited to the STEM (Science, Technology, Engineering and Mathematics) subjects. This trend can be easily linked with the recent policies of Italian and European governments aiming to reshape curricula (Eurydice, 2012). Figure 7 sums up the key competences for lifelong learning identified by the European Commission as guidance for all member states in developing national curricula. Furthermore, researchers like Dede (2010), Voogt and Roblin (2012), van Laar, van Deursen, van Dijk and de Haan (2017) reported a set of particular abilities that an individual should develop to cope with the requirements of the society of the XXI century.

European Union

2006	2018 Proposal for revision
1. Communication in mother tongue	1. Literacy competence
2. Communication in foreign language	2. Languages competence
3. Mathematical competence and basic competences in science and technology	3. Science, technological engineering and mathematical competence
4. Digital competence	4. Digital competence
5. Learning to learn	5. Personal, social and learning competence
6. Social and civic competence	6. Civic competence
7. Sense of initiative and entrepreneurship	7. Entrepreneurship competence
8. Cultural awareness and expression.	8. Cultural awareness and expression

Source: https://www.consilium.europa.eu/en/meetings/eycs/2018/05/22-23/

Figure 7: key competences for lifelong learning.

Amongst these skills we can find the traditional literacies like reading and computing, but more complex skills seems to be on demand. Capability to work in team, ability to express by means of digital media, ability to master technology, also to face the consequences of a digital citizenship: these are completely different skills from the traditional skills developed in the education of the XX century. The school is changing accordingly to try to provide students the best possible option, but teachers and administrative staff are having a hard time figuring out which methodology is effective in fostering these new skills and which is not.

Within innovative methods proposed to start a revolution in school, Educational Robotics (ER) has become increasingly popular because it seems to help in developing transversal skills, soft skills, attitudes toward STEM subjects and career, technical and digital skills and all the so called 21st century skills (Atmatzidou and Demetriadis, 2016; Atmatzidou, Demetriadis and Nika, 2018; Caballero-Gonzalez, Daniela and Lytras, 2019; Muñoz-Repiso and García-Holgado, 2019; Eguchi, 2015, 2016; Kandlhofer and Steinbauer, 2016).

In ER the underlying pedagogical approach is that of Constructionism, a learning theory suggested by Seymour Papert on the basis of the work of Jean Piaget (Ackermann, 2001; Papert, 1980, 1991, 1999). The learning approach is a construction and re-construction of mental representation more than a transmission of knowledge. An effective learning takes place with the usage of manipulative materials (cognitive artefacts), by means of which students enter into a construction activity of a meaningful product. In this activity building knowledge is the natural consequence of several elements: an experience of creation, experimentation, direct observation of the effects of the actions performed and the sharing of ideas in a highly motivating context. From this point of view, technology and innovative learning environments increase the chances for students to learn.

A reference for building and carrying out activities with ER is also made to the psychological theory on multiple intelligences, especially to the US psychologist Howard Gardner (1992) and to his conceptualization, as the framework for the technical laboratory's activities. In fact, ER activities are intended to give students the chance to discover by himself/herself his/her own main abilities. Each person is characterized by a distinctive profile given by the combination of distinct areas of our mind (logical-mathematical, linguistic, visuo-spatial, bodily kinaesthetic, interpersonal, intrapersonal). Although all these abilities are something we are endowed with, they can be developed through experience; moreover, being more conscious of his or her own areas of excellence and deficiency could really help the student to acknowledge his or her personal value, thus envisaging a path of personal growth, study or career.

The use of project-based learning in ER activities indicates an overall approach to the design of learning environments, characterized by a peculiar emphasis on the cooperative research of feasible and effective solutions to a starting problem, involving systematically new technology and trying to produce real and tangible products as an outcome of the activity. Project-based learning (PBL) is a model that organizes learning around projects. Projects are complex tasks, based on challenging questions or problems, that involve students in design, problem-solving, decision making, or investigative activities; give students the opportunity to work relatively autonomously over extended periods of time; and culminate in realistic products or presentations. Project-based learning is an educational strategy for designing learning environments, characterized by a peculiar emphasis on the cooperative research of feasible and effective solutions to a starting problem, involving systematically new technology and trying to produce real and tangible products as an outcome of the activity. This approach is based on "learning by doing" philosophy (Jones, Rasmussen and Moffitt, 1997; Papert, 1972; Resnick, 1987; Schank, 1982).

Peer tutoring is another useful technique which use technologies to teach in classroom: some students will be facilitators in the learning process to help other students of the same age or younger. Positive attitude towards social relation can be thus developed into the school, acting as a protective factor for the risk of absenteeism and preventing school dropouts (Daniela and Strods, 2019). Students learning and developing their skills are in charge of the responsibility of working with their companion in order to fill the team gaps.

In a typical, ER activity students are divided in teams composed of 3-4 people, and within the team some roles are defined: the **designer** (responsible for the project and coordinator of the team, the person who has the task to communicate to the others building instructions of the robot), the **warehouse worker** (responsible for the robotic kit, the student who has the task to look for the Lego/Arduino pieces inside the box), the **technical-assembler** (responsible for the robot assembling, the student who has the task to build the robot receiving instructions from the designer and Lego/Arduino pieces from the technical-assembler), and the **validator** (responsible for the check of the robot assembly, observing the instructions on the computer). This roles assignment is very important for student to learn to cooperate and simulate a real-life scenario where workers are building a solution out of the available resources.

All these characteristics of ER make it stand out from the crowd of possible applications of robots in the field of education. Many authors use the term Robotics in Education (RiE) as a synonym for Educational Robotics (ER), but this cannot be the case. Even if some literature uses "Robotics in Education" and "Educational Robotics" as synonyms (Benitti, 2012; Eguchi, 2017), a distinction should be made between the two labels to bring clarity in reporting such experiences and thus enhancing analytic approaches in reviewing them, which is preliminary to the synthesis of a definition for ER that could be used in further research in this field and establishing a common ground to develop sound research methods.

Robotics in Education (RiE) is a broader term referring to what Robotics can do for people in Education. For example, it can help impaired students to overcome limitations or it can help teachers to gain attention or to deliver content to pupils. Educational Robotics (ER) refers to a specific field which is the intersection of different kind of expertise like Robotics, Pedagogy and Psychology. ER builds on the work of Seymour Papert, Lev Vygotsky, Jean Piaget (Ackermann, 2001; Mevarech and Kramarski, 1993; Papert, 1980;

Vygotsky, 1968) to bring not just robotics in education, but to create meaningful experiences on Robotics since an early age (Scaradozzi, Cesaretti, Costa, Screpanti and Zingaretti, 2019; Scaradozzi, Sorbi, Pedale, Valzano, & Vergine, 2015). ER is made of robots allowing a construction/deconstruction and programming activity, teachers/experts facilitating the activity, methodologies enabling students to explore the subject, the environment, the content of the activity and their personal skills and knowledge. These key elements of ER make it an integrated approach to STEM (Brophy, Klein, Portsmore and Rogers, 2008) and an interdisciplinary and transdisciplinary subject (Eguchi, 2014).
Scaradozzi, Screpanti and Cesaretti (2019) identified four different features to describe a RiE experience or project: the learning environment, the impact on students' school curriculum, the integration of the robotic tool in the activity, the way evaluation is carried out. Regarding how the robotic tool is integrated into the activity we can distinguish ER as a subset of RiE (Figure 8).

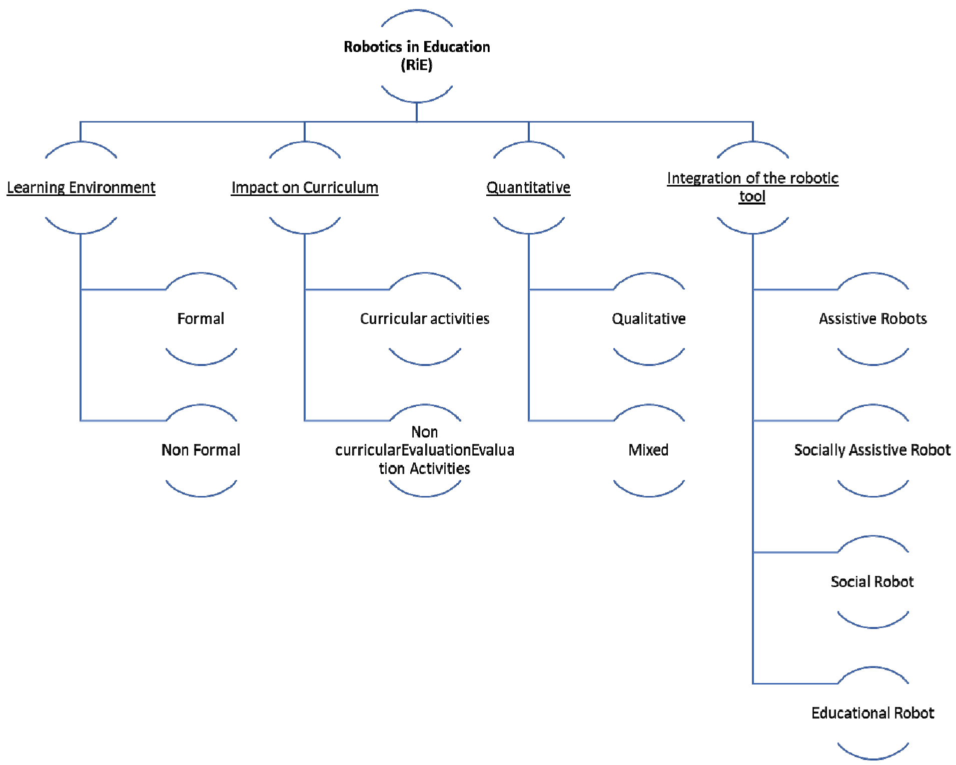


Figure 8: classification of RiE.

Students can learn in a variety of settings (e. g. at school, at home, in an outdoor environment). Each setting is characterised by the physical location, learning context and cultures. Usually, each setting holds specific rules and ethos to defines relationships, behaviours and learning activities. It's authors opinion that it is important to specify in a RiE activity whether the **learning environment** is formal or non-formal. Formal education is usually delivered by trained teachers in publicly recognised organisations providing structured activities and evaluation. Non-formal education can be a complement to formal education, but it may be apart from the pathway of the national education system, consisting in a shorter activity. Usually, non-formal activities lead to no qualification, but they can have recognition when they complete competences otherwise neglected. Formal environment is where formal education usually takes place (e.g schools) and non-formal environment is where non-formal education usually happens (e.g. private houses, company's headquarters, museums). Teaching methodologies, spaces, furniture and many other variables influence the outcome of a RiE or an ER activity, but they are out of scope in this part of the classification, which intends to make a distinction at a broader level.

The **way activities are integrated in education** strongly impacts their design and their expected outcomes. Activities carefully designed to fit the curriculum needs, carried out regularly in the classroom to support students' learning of a concept and whose evaluation is recognized in the final evaluation of the school on students is a curricular activity. Seldom activities organized to better support the teaching of particular concepts, both inside and outside the classroom, and that lead to no final formal recognition are non-curricular activities. There may be activities performed at school (formal learning environment) that do not account for the final evaluation of the student (non-curricular activity). On the other hand, there may be an activity performed outside the classroom environment (non-formal learning environment) that is recognised into the final evaluation of the student provided by school (curricular activity).

Robotic tools that are used into the activities should be distinguished according to the purpose they serve in the educational context. First, they can reduce the impairments for students with physical disabilities. These tools are usually medical devices that help people in their activities of daily living, and they compensate for the lost function. This kind of robots are

Assistive Robots and they are not intentionally produced to meet the need of education, but to meet the needs of impaired people. Second, some robots can help people with a social impairment (e. g. Autistic Spectrum Disorder). This kind of robots can be defined as Socially Assistive Robots, because they are capable of assisting users through social rather than physical interaction (Matarić and Scassellati, 2016). Socially assistive robots “attempt to provide the appropriate emotional, cognitive, and social cues to encourage development, learning, or therapy for an individual” (Matarić and Scassellati, 2016, p. 1974). Third, some robots can be companions to students’ learning or to teachers’ while teaching (Belpaeme, Kennedy, Ramachandran, Scassellati and Tanaka, 2018). These robots are called Social Robots, because they are designed to interact with people in a natural, interpersonal manner to accomplish a variety of tasks, including learning (Breazeal, Dautenhahn and Kanda, 2016).

Fourth, robots can be a tool to study Robotics, STEAM subjects and to develop transversal skills. ER projects use this kind of robots. Generally, they are presented to students as disassembled kits to give the possibility to create meaningful interdisciplinary pathways, letting students be free to build original artefacts. To build an artefact with fully functioning actuators and sensors, students need to master the fundamental concepts about Robotics. Only when these concepts are reworked and absorbed by students, they can feel confident in reusing that kind of knowledge in another context. So, one of the main features of ER it’s the basic understanding of Robotics fundamentals.

Evaluation in RiE activities could be carried out by using a qualitative method, a quantitative method or a mixed-methods approach. Qualitative methods in education pertains to research and to everyday practice. Teachers and researchers can analyse essays, focus groups, scenarios, projects, case studies, artefacts, personal experiences, portfolios, role play or simulation and many other outputs of the activities. This is a deep and rich source of information on students’ learning, but sometimes impractical in a crowded classroom and always vulnerable to personal biases or external influence. On the opposite, quantitative methods are easier to replicate and administer. They try to summarize with numbers the outcome of an activity. Common tools in quantitative methods are based on questionnaires, tests and rubrics. Anyway, experiments and empirical method should be applied to prove these methods valid, reliable and generalizable. Moreover, a quantitative evaluation in education is often deemed as poor and reductive.

Lately, researchers in education have been overcoming the historical distinction between qualitative and quantitative methods to exploit the beneficial aspects that both methods provide. Researchers have been proposing the mixed methods approach as an appropriate research method to address problems in complex environments, like education. The choice of mixed-method design is usually well motivated because it could imply a lot of work as it requires that both quantitative and qualitative data are collected. In the last years some novel real-time techniques have been introduced to monitor students during their activities. Technology and artificial intelligence seem to be promising in providing feedback on students' learning and in integrating both qualitative and quantitative methods of assessment. Moreover, it could be deployed into classroom seamlessly and give response on the activity to support the assessment.

1.2 Measurement and modelling in ER research

As already stated in the previous subsection, evaluation of activities pertains to teachers and observer. An observer can evaluate the phenomena in front of him/her mainly by means of two methodological approaches: quantitative and qualitative (Firestone, 1987).
Quantitative methods, including experimental and quasi-experimental, are rooted in a positivist philosophical tradition and they are used in education research (Kapur, Hung, Jacobson, & Voiklis, 2007; Suthers & Hundhausen, 2003) to establish causal or quasi-causal explanations of design or intervention effects versus control or comparison conditions. Qualitative approaches have a phenomenological philosophical basis and are used to describe and to understand educational contexts and environments (Twinning et. al, 2017).
Since the late 1980s it has become increasingly common for educational researchers to use both quantitative and qualitative methods in a complementary manner (Firestone, 1987). Nonetheless Benitti (2012), Alimisis (2013) and Toh et al. (2016) report that there is a lack of quantitative studies in the literature about Educational Robotics.
Considering the target of the evaluation, evaluation can focus on performance, attitude and behaviour. Performance measurement can be a test whose aim is to evaluate the knowledge acquired on the subject and/or the ability to use it to perform a task (Blikstein, Kabayadondo, Martin and

Fields., 2017; Di Lieto et al., 2017; Screpanti et al., 2018b) or it can be based on neuropsychological measures (Di Lieto et al., 2017).

Complex task evaluation can be related also to the development of skills, not only knowledge. Moreover, written tests more often reflect theoretical knowledge, while practical exercises or tests demonstrate applied skills. Attitudes and skills are more often measured through surveys and questionnaires (Atmatzidou and Demetriadis, 2016; Cesaretti et al., 2017; Cross, Hamner, Zito and Nourbakhsh, 2017; Di Lieto et al., 2017; Goldman, Eguchi and Sklar, 2004; Lindh and Holgersson, 2007; Screpanti et al., 2018a; Weinberg, Pettibone, Thomas, Stephen and Stein, 2007), which are easy to administer and useful for triangulation. Measures of student's behaviours in ER activities can help the design of the learning environment as well as deepen understanding of how students learn (Kucuk and Sisman, 2017).

Another distinctive feature of evaluation regards when to measure. Measurements (or evaluation of a student's state) can be performed before the activity, iteratively during the activity and after the activity. In addition to this, stating the purpose of evaluation can help researchers and teachers to clarify how and when to perform such assessment. Summative assessment (or assessment *of* learning) is often related to the outcome of the activity and it is often regarded as the post-activity evaluation which relates to benchmarks. Formative assessment (or assessment *for* learning) is often a kind of evaluation taking place before the activity, but it can also be iterative, occurring periodically throughout the ER activity. The purpose of formative assessment is to adjust teaching and learning activities to improve student's attainment. More recently, the field of assessment *as* learning brought the idea that formative assessment, feedback and metacognition should go together (Dann, 2014; Hattie and Timperley, 2007).

At the end of an ER activity it would be interesting to investigate the process that led to the resolution of a specific problem, or to the design of a software sequence. During an ER activity, students experiment and modify their sequence of instructions or robot's hardware structure, to obtain a specific behaviour. They usually work in team in a continuous process of software and/or hardware improvement, as specified by the TMI model (Martinez and Stager, 2013). It would be very interesting for an educator to have the chance to observe and analyse this process, but it is not realistic to have one teacher per group that keeps track of the students' development inside the classroom. New experimentations in constructionist research laid the way into new possibilities of insights into the students' learning processes. Evaluation can be performed using the "offline" or "online" method. The offline methods are those assessments gathering information one or more times during the activity and then usually processed later by a human evaluator. The online methods are those assessments "continuously" gathering information on students' activity (e.g. camera recording students' behaviour, sensors collecting physiological parameters, log system recording students' interactions) aiming at providing an analysis of the student's learning while the student is still exploring the activity. Online methods are usually automated and rely on Educational Data Mining (EDM) and Learning Analytics (LA).

The first applications of these technologies tried to extrapolate information from data gathered from structured online learning environments (Baker, Corbett and Koedinger, 2004; Beck and Woolf, 2000; Berland, Baker and Blikstein, 2014; Merceron and Yacef, 2004): in this type of condition it was easier to deduce relations and recognize patterns in the data. Recent studies (Asif, Merceron, Ali and Haider, 2017; Ornelas and Ordonez, 2017) tried to predict students' success using machine learning algorithms on data gathered from closed environments.

Blikstein et al. (2014) collected the code snapshots of computer programs to investigate and identify possible states that model students' learning process and trajectories in open-ended constructionist activities. Berland, Martin, Benton, Petrick and Davies (2013), extending the previous work by Turkle and Papert (1992), registered students' programming actions and used clustering to study different pathways of novice programmers. This led to the identification of three general patterns: tinkering, exploring, and refining. To evaluate different aspects of constructionist activities, other works relied on external sensors (cameras, microphones, physiologic sensors), and automated techniques, like text analysis, speech analysis, handwriting analysis, and others (Blikstein and Worsley, 2016). A key for future developments and experimentations will probably be connected to the availability and cost of implementation of such technological solutions for classroom assessment. External sensors may be more expensive, whereas embedded software solutions and machine learning algorithms could be effective and reliable in extracting evidence of students' learning process and helping teachers to provide personalised feedback to students. Anyway, as stated by Berland et al. (2014), EDM and LA in constructionist environment aim at generating complementary data to assist teachers' deep qualitative analysis with quantitative methods.

A first experimentation that used data mining in the field of ER was conducted by Jormanainen and Sutinen (2012). They adopted the Lego Mindstorms RCX and collected data from students' activities with the main functions of a new graphical programming environment that they designed. They created an Open Monitoring Environment (OME) for the teachers involved, obtaining promising results with decision trees algorithm (J48 implementation) for classifying students' progress in the ER setting. But probably there were some weaknesses in this experimentation: the kit chosen for the study was anachronistic, indeed in 2012 the new model of Lego Mindstorms (the NXT version) had been on the market since 2006; only 12 students and 4 teachers from primary school were involved, a very low number of participants to validate the method; a new graphic programming environment was developed, but it was without a block-based approach, maybe not so friendly for primary school pupils.

However, the existing quantitative and qualitative methodologies, whether separate or in combination, seems to be not sufficient for providing appropriate information and understandings of the dynamics of educational systems, because the commonly used in quantitative research (e.g., differential equations and statistical modeling) are linear tools that work by breaking a system into its components or parts, studying the parts individually, and then adding the parts together to form the whole. For example, studying the single relation between exposure to ER training and STEM related attitudes may fail to describe the whole picture. Attitudes, knowledge and successful learning trajectories apparently emerge from many stimuli a learner may receive during the treatment. Emergent phenomena in an educational complex system generally have nonlinear properties, which cannot be analysed by adding up the parts because the patterns at the macro-level of a complex system generally have different properties to the constituent parts at the micro-level of the system.

Furthermore, quantitative and qualitative approaches are best suited to explain and understand what has already emerged, providing explanations by analysing measure of central tendency and descriptions by structured or unstructured observation. However, the same trajectory of interactions may not have occurred, even if there had been similar initial conditions. For this reason, Jacobson, Levin and Kapur (2019) suggest that it is necessary to study and explain the patterns that actually unfolded, as well as the space of possible trajectories that could have unfolded. Moreover, the space of possible trajectories for an educational system is very important for policy purposes.

Traditionally, science has tended to study phenomena in isolation. Today, there is a greater need to develop systemic approaches for designing and understanding the world. Thinking in levels can have considerable impact on interpreting phenomena in all fields of research and Complex systems (CS) theory is having considerable impact on the pure sciences and within many professions such as engineering, medicine, finance, law, and management. For example, the biological perspective of complex systems that highlights interdependence and co-evolution, with emergent patterns formed by self-organization, is now seen as equally important as the traditional perspective of competitive selection in understanding biological evolution (Jacobson and Wilensky, 2006; Kauffman, 1995;).

The foundation of a CS approach may be leveraged to close the gap between theory and method considering the system of a student learning as a whole and calling upon existing frameworks and philosophy in the field, the ontological framework defines the complex macro-behavior of systems, the dynamic micro-interactions of system components, and the emergent mechanisms that produce system outcomes (Mitchell, 2009).

A CS is a collection of interacting components (i.e., those that interfere, cooperate, or collaborate) that gives rise to complex behavior (Mitchell, 2009; Strogatz, 1994). System components may be a material, conceptual, or semiotic forms such as individual students, teachers, and technological objects; motivation, behavioral, affective, epistemological, and cognitive variables; or words, text, symbols, and discourses (Bunge, 2000). Components within CS interact over time to produce emergent outcomes at higher levels of analysis that are characterized by nonlinear behavior such as sudden transitions from one state to another or bifurcations in topological structure. Emergent outcomes are more than the sum of their parts, meaning the complex behavior cannot be reduced to the components that make up the system (Holland, 2006).

To preserve the complex, dynamic, and emergent ontology of interaction of CS models, research should be formulated in a way that represents the changing and interdependent relationships among individuals and observations over time, as well as interactions amongst multiple levels of analyses (Hilpert and Marchand, 2018).

Research questions leading to CS methods and analyses are focused on time-intensive, relation-intensive, or time-relation intensive processes. Questions should correspond to data collection techniques that can provide evidence for interaction dominant processes.

Research questions that reflect time-intensive processes may focus on any unit of analyses undergoing a process that is expected to unfold over time. Methods for collecting time-intensive data range from the use of experience sampling techniques such as the use of diaries and surveys to the use of sensors and video recordings. What defines the method is the microgenetic observation of variables over time.

Relation-intensive processes describe the relationships or interactions among individuals or variables in a system. Research questions about relation-intensive processes focus on identifying the structure of the relationships in a system and the purpose and weight or value of exchanges. Relation-intensive questions may focus on interpersonal relationships, or the focal unit of analysis may be relationships between organizations or psychological constructs. Methods for collecting relationship-intensive data range from observations of behaviour to analysis of existing documents and the use of surveys. Both interpersonal and intervariable relations can be targeted.

Time-relation intensive processes describe both within-and between-element changes over time. Research questions about time-relation intensive processes focus on microgenetic correspondence among social partners, individual and contextual elements, psychological constructs, or intergroup/organizational change over time. Time-relation intensive questions may focus on how individuals and their social partners reciprocally influence each other over time, how multiple psychological phenomena vary together, how group membership change over time, or how individuals influence change in group behavior, to name a few areas of inquiry. Methods for collecting time-relation intensive data range from surveys to observations, video recordings, and sensors (Hilpert and Marchand, 2018).

A range of analytic techniques can be used to investigate CS research questions and study complex, dynamic, and emergent processes. In the quantitative field, Parunak and others (1998) distinguished between equation-based modelling and agent-based modelling for CS research. In agent-based modelling, the goal is to emulate the system by programming components (or agents) that follow behavioural rules, thereby producing emergent outcomes. In equation-based modelling, the goal is to evaluate a system using observations that are entered into equations. Both are based on the notion that two kinds of entities can be examined, individuals and observables, both with temporal resolution.

Three nonlinear equation-based techniques that can offer novel insight into interaction-dominant educational systems: nonlinear time series analysis, dynamic modelling, and network analysis. These three techniques may be attractive to educational psychologists because they can be used with data sources and data collection methods that are already widely utilized among researchers in the field (Hilpert and Marchand, 2018).

An example of students' learning that is close enough to ER environment is to be found in Piech et al (2012) where authors focus on the modelling of how a student learns to program by means of a Hidden Markov Model (HMM) where a student is supposed to be in a "state", or "high level milestone", that cannot be directly observed (it is a latent variable), but can be evaluated by means of a code snapshot, which is a "noisy sensor" of the latent variable. To identify the finite set of High-level States, the transition probability and the emission probability, authors sampled 2000 snapshot from a programming exercise and clustered the sample by means of K-mediods. From the analysis of the high-level states dead end of exercises and other recognisable students' developmental progress were found. Moreover, the model was used to predict performances of students in midterm exam with successful results from implementation. This accounts for the belief of the authors on this approach being more promising than the data-driven methodology for improving programming instruction and pedagogy, and for gaining deeper insights into learning in general.

Another example in modelling learning is provided by Stenbeek and van Geert (2013). These authors focus on learning-teaching trajectories as emergent and dynamic phenomena resulting from the interactions in the entire educational context, in particular the interaction between students and teachers. Two dynamic models are proposed in relation to the two levels of the phenomenon under investigation:

- one focusing on the short-term dynamics of learning-teaching interactions as they take place in classrooms, an agent model whose parameters depend on the preceding history of the student-teacher dyad; the short-term patterns of learning-teaching interaction will cause these parameters to change on the long-term, or to stabilize in the form of a particular attractor states. The dynamic agent model suggested that student's learning concerns were not fixed, but that their strengths changed during and as a consequence of the ongoing interaction. It highlighted the dynamic interplay of both student's and teacher's concerns and emotional appraisals in forming short-term positive or negative learning experiences. It was shown how not only the student's, but also the teacher's level of balance in autonomy-relatedness concern, and a sufficient level of competence concern dynamically constituted teaching-learning processes.

- one focusing on the long-term dynamics of interactions in a network of variables encompassing concerns, evaluations, actions and action effects; technically, every node in the network corresponds with a particular variable, e.g. the student's current level of his competence concern, but mathematically, each variable is modelled by means of a logistic support equation. The support parameters link every variable to a subset of other variables in the network. The long-term model showed that there exist dynamic relationships between the student's and teacher's concerns, emotional appraisals, and the student's performance level. These dynamic relationships can be supportive or competitive, competing, and symmetrical or asymmetrical, demonstrating that particular patterns of relationships act as control parameters producing emergent learning-teaching trajectories, such as "unsuccessful" or underachievement trajectories.

Another kind of representation of learning is provided by Canuto (2008), whose model assimilates a student to an accumulator and a teacher to a reference provider. This model, even if well-grounded in control systems theory, has a mere comparative purpose. The author himself states that this work is not supported by experimental data.

These few examples of models of students' learning show that there can be several approaches to modelling learning, but all have in common the need for real data in order to be validated and therefore used to describe and predict reality. Quantitative methods are the best instruments for observing and capturing information on a quantitative level, thus providing the basis to build, evaluate and validate such models.

Section 2 – Evaluation tools

In order to study Educational Robotics learning systems with the Systems theory approach, the sub elements of the system that are of interest should be identified. To simplify this concept, four phases in modelling were identified to represent the fundamental blocks of the system:

1. How can it be studied the learning process? What is learning? Which variables should be considered to be included in the model?
2. Is it possible to establish the causality principle (i.e. given an input at a time t, we have a certain output at the time t)?
3. Metrology: how can we measure the intended variables?
4. Cybernetics: how can we transform information on the output to determine which input to provide?

To answer the first question, the dimensions of the construct learning should be identified and defined accordingly. The quantitative methodology can help to establish the causality principle and to measure the intended variables. In fact, introducing the empirical cycle (Figure 9) into pedagogical research has helped in realising studies whose conclusions are strictly drawn from empirical evidence.

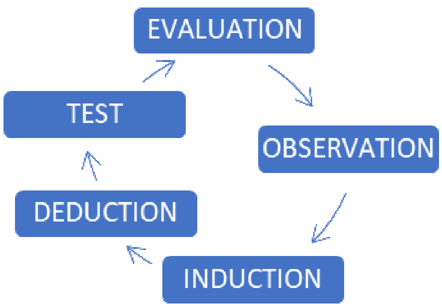


Figure 9: empirical cycle.

Testing a hypothesis derived from a theory or from observation means that we are respecting six basic rules. First, a hypothesis is empirically testable. This means that it should be possible to collect empirical or physical evidence, or observations, that will either support or contradict the hypothesis. Second, a study and its findings should be replicable, meaning that it should be possible to consistently repeat the original study. A hypothesis is more plausible if it is repeatedly confirmed; this requires that it is possible to repeat or replicate a study. If the expected result occurs only once or in very few cases, then the result could just have been found by chance. Third, anybody should be able to get the same results based on the description of the assumptions and procedures. A researcher should therefore be as objective as possible about assumptions, concepts and procedures. This means that all these elements should be clearly and explicitly defined, leaving no room for subjective interpretation. Fourth, transparency is needed because in science anyone should be able to replicate your results for themselves. Fifth, a hypothesis should be falsifiable, it should be possible to imagine finding observations that will contradict the formulated hypothesis. Sixth, a hypothesis should be logically consistent or coherent, so there shouldn't be any internal contradiction and the conclusions based on our observations should also be logically consistent. Furthermore, in social sciences it is very important to comply with ethical requirements of respect, beneficence and justice.

Introducing the empirical cycle and the positivist/empirical philosophical stance in the present work is very important because it allows the definition of the proper sensors to capture information to solve the first three points of the modelling phase, in particular the metrology and the causality principle. Introducing the quantitative methods as a way to solve the metrological issue requires to follow the rules of the gold standard for this kind of research. Specifically, it will be necessary to demonstrate the internal and external validity of the study and of the instrument. Internal validity means that the experiment, or quasi experiment, avoids confounding (more than one possible independent variable [cause] acting at the same time). The less chance for confounding in a study, the higher its internal validity is. External validity analyses how well data and theories from one setting apply to another. Moreover, reliability should be demonstrated, because a measure is considered reliable if it would give us the same result over and over again. The objects under measurement in the present thesis are constructs: explanatory variables which are not directly observable. For example, an object's center of mass is certainly a real thing, but it is a construct (not another object). For example, the concepts of intelligence and motivation are used to explain phenomena in psychology, but neither is directly observable. How is it possible to know whether the measure is accurate and reflects the intended construct? The validity of an instrument or manipulation method is commonly referred to as measurement or construct validity. How is it possible to assess construct validity? How is it possible to determine if this score actually reflects the property? Some criteria are available. Firstly, a test can be said to have "face validity" if it appears to measure what it is supposed to measure; typically, an expert in the field provide this kind of validity by evaluating the test. Secondly, the extent to which a measure is related to an outcome is called Criterion validity, which is often divided into concurrent and predictive validity. Concurrent validity refers to a comparison between the measure in question and an outcome assessed at the same time. Concurrent validity reflects only the status quo at a particular time. Thirdly, convergent and discriminant validity can be a proof for construct validity. They can be thought of as a combination of measures of constructs that theoretically should be related to each other are, in fact, observed to be related to each other (that is, you should be able to show a correspondence or convergence between similar constructs) and measures of constructs that theoretically should not be related to each other are, in fact, observed to not be related to each other (that is, you should be able to discriminate between dissimilar constructs).

Not only validity of measurement, but also reliability of measurement should be demonstrated to account for an instrument's consistency or stability or precision. A reliable instrument will result in highly similar scores if we repeatedly measure a stable property in the same person. There are three types of consistency: over time (test-retest reliability), across items (internal consistency), across different researchers (inter-rater reliability) and over time for the same observer (intra-observer consistency). To visualize the relatedness of validity and reliability of a measure a shooting target can be used. Figure 10 shows four different shooting results, where the center of the target (red dot) represents the true score and the repeated shootings scattered around the center (blue dots) represent the measures taken by our sensor. A reliable measure will result in a dense cloud of blue dots around the same center. A valid measure will result in a cloud of blue dots somehow centered around the true score. A reliable and valid measure will result in a dense cloud of points around the true score.

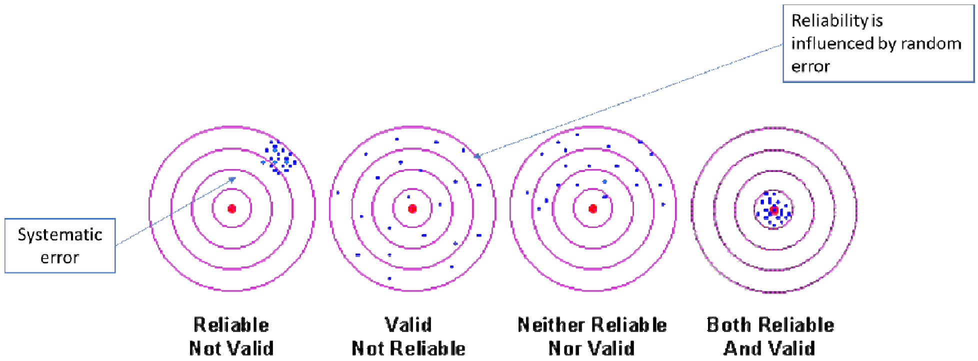


Figure 10: validity and reliability of a measure.

Some examples of traditional sensors from the psycho-socio-pedagogical field to capture information on the variables of interest are:

- 'Survey' is a general term that can refer to a list of questions asking about biographical information, opinions, attitudes, traits, behaviour, basically anything. Surveys generally cover a variety of topics.
- 'Questionnaire' is used when the focus is on one construct, or a related set of constructs, usually psychological traits, emotional states or attitudes.
- 'Test' is used when the aim is to measure an ability, such as general intelligence or math proficiency.

Surveys, test and questionnaires all consist of a series of questions, referred to as "items", usually accompanied by a set of discrete response options or a continuous range to choose from.

When a quantitative approach is used in a study, it should be reported detailed information to allow a knowledgeable reader to replicate the study, assess the rigorousness of the research design and evaluate the robustness of the results and generalisability of the conclusions. As already stated, to assess the robustness of the study it should be provided evidence to show that the instrument used to measure the variables is valid and reliable, namely that it measures what it is supposed to measure and does so with a suitable degree of accuracy. Not only content validity (whether or not the test covers a representative sample of the variable to be measured) and predictive validity (the degree to which the variable of interest can be effectively predicted), but also construct validity is a key element to assess the robustness of the instrument when measuring variables that are not directly observed. Whenever such variables are measured it is essential to reference the validation study and, in case there isn't such study reported in literature, the factorial structure may be reported. This can be done using Exploratory Factor Analysis (EFA), which is employed to understand the shared variance of measured variables that is believed to be attributable to a factor or latent construct. It is not required to have any specific hypotheses about how many factors will emerge, and what items or variables these factors will comprise. If these hypotheses exist, they are not incorporated into and do not affect the results of the statistical analyses. The goal of EFA is to identify factors based on data and to maximize the amount of variance explained. The main parameters of EFA must be reported: number of factors extracted, total variance explained by extracted factors, Kaiser Meyer Olkin index (KMO) (Kaiser, 1970; 1974), sample size, and ratio of number of participants to number of variables factored. The requirements for carrying out such analysis are: Correlation Matrix with significant correlation coefficients and non-zero determinant, measures of sample adequacy (KMO index) and Bartlett's sphericity test (Bartlett, 1937). The Kaiser-Meyer-Olkin Measure of Sampling Adequacy is a statistic that indicates the proportion of variance in the variables analysed that might be caused by underlying factors. High values (close to 1.0) generally indicate that a factor analysis may be useful. If the value is less than 0.50, the results of the factor analysis probably won't be very useful. Bartlett's test of sphericity tests the hypothesis that the correlation matrix is an identity matrix, which would indicate that variables are unrelated and therefore unsuitable for structure detection. Small values (less than 0.05) of the significance level indicate that a factor analysis may be useful. After the assumption check a dimensionality reduction is performed to identify which factors underlie the observed variables; methods to obtain this information can vary according to the issue at stake. Once the eigenvalues of the matrix are extracted, they can be represented by a Scree plot and different methods for choosing the number of factors extracted can be used. Kaiser criterion (1959) or Guttman criterion (1954) establish that we may choose the absolute value of eigenvalues which is greater than one. Cattell (1966) proposed that the scree plot can be used to graphically determine the optimal number of factors to retain. The Scree-test involves finding the place where the smooth decrease of eigenvalues appears to level off to the right of the plot. To the right of this point, presumably, you find only "factorial scree" - "scree" is the geological term referring to the debris that collects on the lower part of a rocky slope. Thus, no more than the number of factors to the left of this point should be retained. Horn's parallel analysis (1965) is a statistical method used to determine the number of components to keep; the method compares the eigenvalues generated from the data matrix to the eigenvalues generated from a Monte-Carlo simulated matrix created from random data of the same size. Gorsuch (1983) suggests that it is necessary to do more than one analysis and retain the factors that appears in different analysis.

It should be kept in mind the fundamental difference between Confirmatory Factor Analysis and EFA: CFA is a confirmatory technique, it is theory driven, so the planning of the analysis is driven by the theoretical relationships among the observed and unobserved variables. When a CFA is performed, the researcher uses a hypothesized model to estimate a population covariance matrix that is compared with the observed covariance matrix. Technically, the researcher wants to minimize the difference between the estimated and observed matrices (Schreiber, Nora, Stage, Barlow, & King, 2006). EFA is a statistical method used to uncover the underlying structure of a relatively large set of variables, the variance of large number of variables can be described by few summary variables, i.e., factors. (Thomson, 2004). EFA transform the current set of variables into other variables such that each new variable is a weighted combination of the current ones. Hence, information is not added nor removed but only transformed. A typical way to make this transformation is to use eigenvalues and eigenvectors: each eigenvector provides the direction of the transformation and the eigenvalues represent all the new variables or "factors". An eigenvalue more than 1 will mean that the new factor explains more variance than one original variable, so they can be sorted in decreasing order of the variances they explain. Thus, the first factor will be the most influential factor followed by the second factor, the third and so on. Factors are retained based on the percentage of variance explained collectively by factors themselves and the interpretation of each factor should be linked to the combination of the original features (items) composing them.

However, it is not enough for the instrument to be valid; it must also be reliable, i.e. it must accurately measure what it aims to measure. While there are several alternatives for demonstrating reliability, the index that is most widely used by researchers continues to be Cronbach's alpha (Cronbach, 1951).

Traditionally, an acceptable level for Cronbach's alpha is considered to be >0.7, although this criterion is adjusted depending on the type of research.

For example, if the results of the research have an impact on important decisions that are to be made, a much higher level of reliability is required (Cortina, 1993; Streiner, 2003). Cronbach's alpha supposes the one-dimensional nature of the construct that is being evaluated, i.e. that the set of items or indicators only measure a single dimension, so if the construct that is evaluated has more than one dimension (e.g. subscales), the reliability index for each of these dimensions must be indicated. Lastly, this index is especially sensitive to the number of items included in a scale, meaning that an increase in the number of items may lead to an increase in Cronbach's alpha.
In the following subsections some of the instruments developed and validated to capture the information useful to build the model of learning are presented. The instruments themselves are fully reported in the Appendix section.

2.1 Knowledge and competence

Knowledge and competences are strongly subject related. Teachers and policy makers are often interested in understanding if school achieved the task of educating young minds. To this purpose governments provide indications to school about the curriculum they can implement. According to Recommendation of the European Parliament and of the Council of 23rd April 2008 on the establishment of the European Qualifications Framework for Lifelong Learning (European Union, 2008): *"Learning outcomes" means statements of what a learner knows, understands and is able to do on completion of a learning process, which are defined in terms of knowledge, skills and competence."* These terms have been defined as follows in the same Recommendation:

- "***knowledge***" means the outcome of the assimilation of information through learning. Knowledge is the body of facts, principles, theories and practices that is related to a field of work or study;
- "***skills***" means the ability to apply knowledge and use know-how to complete tasks and solve problems. In the context of the European Qualifications Framework, skills are described as cognitive or practical;

- "***competence***" means the proven ability to use knowledge, skills and personal, social and/or methodological abilities, in work or study situations and in professional and personal development.

Knowledge is sometimes viewed as if it was a concrete manifestation of abstract intelligence, but actually it is an outcome from the interaction between intelligence (capacity to learn) and situation (opportunity to learn), so is more socially constructed than intelligence. Knowledge includes theory and concepts and tacit knowledge gained as a result of the experience of performing certain tasks. Understanding refers to more holistic knowledge of processes and contexts and may be distinguished as know-why, as opposed know-that (Winterton, Delamare, Le Deist and Stringfellow, 2006). From this perspective, it is often argued that acquiring explicit factual knowledge (declarative knowledge) must precede developing procedural knowledge (know how), which relates to using knowledge in context.

Skill is usually used to refer to a level of performance, in the sense of accuracy and speed in performing particular tasks (skilled performance), which has been considered both in its physical psychomotor aspects and mental cognitive aspects as a subject of psychological studies.

Competence is a term subject to such diverse use and interpretation that it is impossible to identify or impute a coherent theory or to arrive at a definition capable of accommodating and reconciling all the different ways the term is used. It can be identified the root out of which stemmed this concept, namely the distinction between competence and performance used by the linguist Noam Chomsky (1980).

Chomsky described linguistic competence as a universal, inherited, modularised ability to acquire the mother tongue, as distinct from performance (ability to understand and use the language). Chomsky's model of linguistic competence and performance has influenced similar models of numerical competence, spatial competence and other areas of domain-specific knowledge (Winterton et al., 2006).

In Italy the "Regolamento recante Indicazioni nazionali per il curricolo della scuola dell'infanzia e del primo ciclo d'istruzione" (MIUR, 2012) put into action the European recommendation establishing which kind of knowledge, skills, competences and values are to be pursued by Italian schools.

Within this framework, in order to evaluate knowledge from ER activities at school, three instruments were developed by researchers responsible for the study, evaluated by experts of robotics and teachers for face validity, and finally used in different experimentations. The instruments developed are:

- a crossword puzzle for primary school's students (Scaradozzi et al., 2016; Screpanti et al., 2018)
- a knowledge test for students from 11 to 13 years old
- a knowledge test for teachers (Scaradozzi et al., 2019)

They have been created for the purpose of assessing some of the concepts introduced during the activities designed for specific learning objectives. The common element of these activities was that they were partially or fully carried out as introductory lessons to robotics, focusing on sensors, actuators, programming and building a coherent mechanical structure.

2.1.1 Crossword puzzle for primary and secondary school's students

In Scaradozzi et al. (2016) and Screpanti et al. (2018), a crossword puzzle was formulated (see Appendix A). Two Crossword puzzles were prepared, one for primary (18 definitions) and one for secondary school (22 definitions). Students had 30 minutes to finalize the work, that was performed like a summative test. Theoretical concepts addressed during the project were the solutions to fill the empty boxes. Definitions can be grouped into 4 categories. For primary school:

- Concepts related to software functions (Scratch): 8,11,12;
- Concepts related to Robotics and hardware (kit Lego Wedo): 1,2,3,7,13,16,17;
- Concepts related to Energy issues: 4,5,9,14,15;
- Aspects regarding teamwork: 6,10,18.

For lower secondary school:

- Concepts related to software functions (Scratch): 1,6 across, 6 down, 14, 16;
- Concepts related to Robotics and hardware (kit Lego Wedo): 2,4,8,9,12,17,18;
- Concepts related to Energy issues: 5,13,15,20,21;
- Aspects regarding teamwork: 3,7,10,11,19

Both questionnaires and crosswords puzzles were delivered to students at the completion of the project. All data from these two kinds of evaluation were pro-cessed using MS Excel. No data was discarded. Mean values were computed per each question in each of the two level of education, then normalized to the highest value of the range (3 for the Elementary school and 5 for the Lower Secondary School) and multiplied by 100 to obtain the percentage value. Figure 11 and Figure 12 show the results from the crossword puzzles scored by primary and lower secondary school, respectively, and grouped by the four main areas of knowledge (Screpanti et al., 2018a).

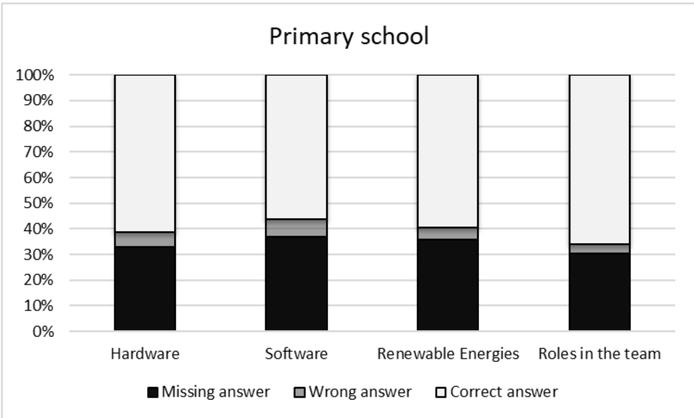


Figure 11: results of the crossword puzzle tests administered to primary school students.

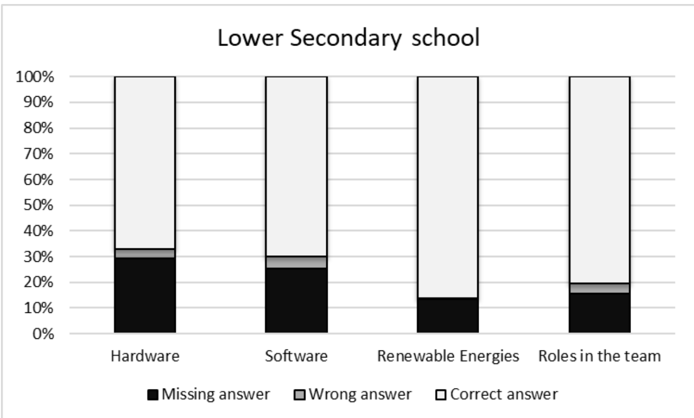


Figure 12: results of the crossword puzzle tests administered to lower secondary school students.

Results from the crossword puzzles were processed assigning each answer to one of the following categories: missing answer, wrong answer, correct answer. Then, percentage of each categories were computed and divided into four categories relating to four areas of knowledge addressed by the activities: hardware, software, renewable energies and roles in the robotic activity. To examine to which extent the didactic objectives concerning these four areas of knowledge have been achieved, a threshold of 80% was set to see how many students answered correctly to this level. This threshold is suggested by (Calvani and Menichetti, 2015) as a reference point to measure the success of the activities. Results from this kind of analysis are shown in Table 1, where rows present the percentage of students above the threshold (>80%), below the threshold (<80%) and the sum of these shares (Tot.). Columns present percentages related to the overall result in the test (Tot), to Hardware (Hw), to Software (Sw), to Renewable Energies (En) and to Roles in the Group (RG).

Table 1: percentage of students who scored above or below the threshold expressed as the percentage of correct answers.

	Primary school					Lower Secondary school				
	Tot	Hw	Sw	En	RG	Tot	Hw	Sw	En	RG
>80%	33	33	39	45	35	40	29	62	86	79
<80%	67	67	61	55	65	60	71	38	14	21
Tot.	100	100	100	100	100	100	100	100	100	100

Looking at Figure 11 and Figure 12 it can be noted that there are more missing answers than wrong answers. This can relate to the peculiarity of the test, because it seems likely that a student not knowing the answer, would seek to match the possible words coming to his/her mind with the boxes in the crossword and if the word didn't fit, he/she would discard the answer leaving the boxes blank. It should be considered that crossword puzzles could be an easier way to answer a question, because they give hints on the length of answers and sometimes even on some letters. Maybe, if students were presented with open questions, they probably would have ventured more to give tentative answer. In general, it seems that lower secondary school presents more positive results in Crosswords puzzle test. It can be noted that the percentage of missing/wrong answers in primary school was on average 40%, whereas in the lower secondary school the average missing/wrong answer seem to be considerably inferior (about 25%). Looking at Table 1, we can see that in the field of Renewable energies at the lower secondary school 86% of the students are above threshold, whereas at the primary school only the 45%. Another curious result is that with a share of 79% at the lower secondary school, students clearly recalled the roles they were assigned during the project, whereas at the primary school only the 35%. This gives credit to the possible side effect of technology. Robots are very attractive to children, but they can draw attention on themselves, to the detriment of contents they are supposed to convey. In fact, even if Robotics is a useful aid to raise interest toward STEM subjects in schools and it has the potential to change the way students learn STEM subjects, technology by itself do not necessarily translate into better learning.

2.1.2 Knowledge and competence test

To eliminate the hardship of a linguistic barrier (recalling a word that exactly matches the number of cells of the crosswords) a test made by 12 question was thought and implemented (See Appendix B). First 10 items were designed to test knowledge and understanding of some fundamental concept of robotics. Questions 11 and 12 are intended to evaluate competence; students have to identify the core elements of a robotic tool that can be found in daily life. Basic concepts chosen are: input and output, block schematization of the functioning of a robot, human-robot comparison with respect to the function of our parts, navigation in a structured environment, sensors, motors, conditions in the programming task, and flowcharts about generic everyday task.

This assessment is strongly related to what has been the focus during the activities. It is expected that shorter interventions (a limited number of hours spent in ER activities) may lead to worst results (students may still enjoy the time spent in such playful activities and are willing to know more, but knowledge in these areas may not be fully achieved). Some concepts represent basic knowledge for each ER course (what is a sensor, what is a motor, what is a central unit to control the robot, what is a visual language), but some items in the test ask students to make some level of abstraction

and generalisation (flowcharts, navigation of an arrow on a grid, systems thinking).
In Appendix B the full tests are reported. Four version of the same test were created: one version for Primary school and one for the Lower Secondary school. For each version a pre-test and post-test set of questions were formulated.
The evaluation of each test could reach a maximum of 12 points: 1 point is assigned to each question. If an item of the test had several answers, the final score for that item would be the sum of all correct answers divided by the total number of answers to provide. For example, a question asked to connect seven parts of the robot with the parts of the human body that are comparatively similar in their functionality within the whole system (i.e. sensors are similar to the human senses, mechanical structure of the robot is similar to the human skeleton, motors are similar to muscles and so on); if a student connected correctly five parts out of seven, the final score for that question would be 5/7=0.71.
Questions 11 and 12 are connected to competence, because it is required to have knowledge of a robotic system, of its parts, why these parts are needed and how to arrange them; this knowledge should be used to think about an object that students did not see in the classroom but can be somehow familiar both because they could have seen it in their daily lives, and because the functioning is similar to the mobile robot they all assembled during the activities, whether in the kit used, or in the shape and functioning, or in the parts to assemble. Question 11 asks about the usage of sensors and actuators in a different context to accomplish a task. So, without a robotic kit in their hands students should mention at least the three elements they could use to realise that robotics system (for example a robotic wheelchair or a domotic system for opening the window). The answer to this question is measured on 3 points: mentioning 3 correct items (the basic solution could be: one sensor, one actuator and the central unit of control) and coherently and correctly explaining their position and functioning within the system they created. One point per correct item identified and explained up to three. Students mentioning more than three, even if correctly, were considered having 3 points out of three. To the purpose of this evaluation creativity or extra-expertise was not taken into account. This kind of question can obviously involve the linguistic competence of the student, but the creators and evaluators deemed as fair enough the endeavour of writing three lines of text. Question 12 ask to

identify the main parts of a robot's functioning. Examples of robots displayed are a mobile robot to cut the grass or to vacuum clean the room, or a Lego NXT system with motors, mechanical structure and intelligent brick. Given the basic structure and the similarity with what students approached during the activities, the final result from this question was evaluated by the number of correct options identified divided by the number of maximum correct choices identified by the creators of the test (this number is set to five).

These measures assessing performance on technology by asking students to identify major parts of commonly encountered electronic devices and asking them to design a solution by means of electronic parts, are not uncommon in the fabrication and exploration activities. Blikstein et al. (2017) proposed a study using a similar method with questions on a key-fob and a blender.

A study was performed in three schools with a total number of 76 students during different projects. Two schools organised a summer school (10 days, 4 hours/day) aimed at fostering STEM attitudes, especially amongst young girls, in a project call "In Estate si imparano le STEM" and one school organised an outreach program (10 days, 3 hours/day) to foster computational thinking, creativity and digital citizenship amongst students. A description of the sample can be found in Table 2. Figure 13 shows the percentage of students who had previous experience with digital technology and social media.

Table 2: descriptive statistics of the sample.

Genre	
M	44,74 %
F	51,32 %
Age	
media	11,52 years
dev std	1,32
Favourite subject at school	
SU	34,21 %
TU	59,21 %
both	2,63 %

We can predict that students having past experience with robots could perform better in the pre-test condition. To test this hypothesis and demonstrate predictive validity, it was performed a t test for the pre-test condition (H0: equal mean scores for students with past robotic experiences and students without past robotics' experiences, Ha: students with past robotics' experiences achieve higher mean scores than the students without past robotics' experiences). Results from the t test on past experiences with robots suggest that students with past robotics' experiences achieve higher mean scores than the students without past robotics' experiences in pre-test condition (t = 3.7472, df = 68.748, p-value = 0.0001842). The mean score of students having past experiences with robots (score= 7.138966) is significantly higher than the mean score of students without any past experience with robots (score= 5.574754).

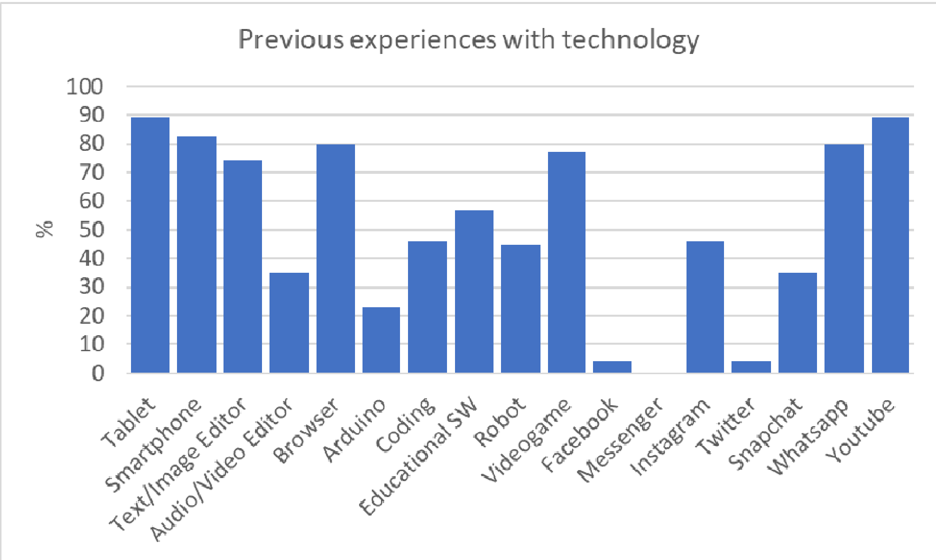


Figure 13: shares of students who had previous experiences with technology and social media.

2.2 Attitudes towards STEM studies and careers

In a robotics' class K12 students need to mobilize several skills and knowledge from different areas, but to really pursue ad engage in robotics they need to improve their attitudes and interest toward STEM subjects. This area lacks comprehensive, validated, STEM-related measurement

instruments but for the S-STEM (Student Attitudes Toward Science, Technology, Engineering, and Math) questionnaire.
S-STEM is the result of a joint effort of the Friday Institute for Educational Innovation's Evaluation Group and the MISO project (2012) at North Carolina State University to design and validate an instrument that could be used to assess the impact of K-12 STEM education in both formal and informal settings.
This instrument was validated in the original English form (Faber, Walton, Booth and Parker, 2013) and subsequently spread across the world (Navarro, Förster, González and González-Pose, 2016; Unlu, Dokme, and Veli, 2016; Jabarullah and Hussain, 2019). This instrument, validated in the US, is originally written in English and developed for US students and culture, therefore it is not straightforward that it can be useful to represent the information derived from students of other countries. Each time the tool has been translated it was also adapted to the culture that language represents, thus it was needed to provide new evidence of its validity and reliability. S-STEM was validated in the US in its two versions, Upper Elementary S-STEM (4^{th} and 5^{th} grade) and Middle/High S-STEM (12^{th} grade), with over 10,000 4-12th grade students from across North Carolina that were in special STEM initiative promoted by public schools in rural areas in 2011-2012.
The areas under investigation in the S-STEM survey are Science, Math, Engineering and Technology and 21^{st} century skills, which include critical thinking, complex communication skills, problem-solving, and self-management skills. The dimensions guiding the formulation of items in each area are expectancy-value theory (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000), which refers to the regular assessment that individuals do on their expectancy of success in a specific task and the value perceived of that task, and self-efficacy (Bandura, 1986), which is the belief in one's ability to complete tasks or influence events that have an impact on one's life. A 5-point Likert-type response scale (strongly disagree to strongly agree) was used for all four subscales. In addition to such attitudes toward STEM disciplines, the S-STEM survey has a section on students' interest in STEM careers because it is important in investigating future participation in the STEM workforce. Sample items for constructs listed above can be found in Table 3.

Results from validation are reported in Unfried, Faber, Stanhope and Wiebe (2015). Some of the studies that use this kind of assessment can be found in Wiebe et al. (2013) and Wiebe, Unfried and Faber (2018).

Table 3: example of items from S-STEM questionnaire.

Construct	Sample Items
Science	I am sure of myself when I do science.
	I will need science for my future work.
Mathematics	I would consider choosing a career that uses math.
	I am the type of student to do well in math.
Engineering and Technology	I like to imagine creating new products.
	I believe I can be successful in a career in engineering
21st Century Skills	I am confident I can lead others to accomplish a goal.
	I am confident I can set my own learning goals.

The questionnaire reported in Faber, Walton, Booth and Parker (2013) was translated in Italian by a well-informed researcher with a certified understanding of English at level C1 of CEFR (Council of Europe, 2011). Subsequently, the questionnaire was adapted to the need of the research.
Even if STEM disciplines are part of Robotics and Robotics can be used to teach those disciplines, the evidence for a connection between Robotics' activities and attitudes to STEM may not be as straightforward as it is lead to believe. For this reason, a fifth construct was added to the original four: Robotics. The attitude toward Robotics was investigated by means of 12 items. In building items, the expectancy value and self-efficacy structure of the whole S-STEM questionnaire was applied. Example items from the fifth construct are: "*I am good at building/repairing robots*"; "*I like to program robots*".
For an overview of the constructs investigated see Table 4. The full version of the Italian questionnaire is reported in Appendix C.

Table 4: description of the S-STEM questionnaire adapted for the Italian sample.

Survey section	**Psychometric profile and answers' scale**	**Measurement Application**
Math Attitudes	Construct 5-point Likert scale (from very disagree to very agree)	items measuring **self-efficacy** related to math and **expectations for future value** gained from success in math
Science Attitudes	Construct 5-point Likert scale (from very disagree to very agree)	items measuring **self-efficacy** related to science and **expectations for future value** gained from success in science
Engineering and Technology (E&T) Attitudes	Construct 5-point Likert scale (from very disagree to very agree)	items measuring **self-efficacy** related to E&T and **expectations for future value** gained from success in E&T
21° century learning	Construct 5-point Likert scale (from very disagree to very agree)	items measuring student's **confidence** in communication, collaboration, and self-directed learning
Your Future	Items 4-point Likert type (from "I do not care at all" to "I care a lot")	**interest** in 12 broad categories of STEM career fields
More about you	Items 3-point scale (Not very well/well/very well or yes, no, I do not know)	predication of future academic performance, plans to take advanced classes, postsecondary plans
Robotics Attitudes	Construct 5-point Likert scale (from very disagree to very agree)	consists of items measuring **self-efficacy** related to robotics and **expectations for future value** gained from success in robotics.

2.2.1 Validation of the Italian modified version of the S-STEM questionnaire

Data from eight schools were collected during the year 2018. The study was carried out in agreement with the school that provided consent to realise activities of ER in the classroom and to pilot measurements. Students and parents could refuse to participate in this study. Incomplete data were discarded. 109 remaining data were used to perform an Exploratory Factor Analysis (EFA) and a Confirmatory Factor Analysis (CFA). Not all the questions were used to perform the Factor Analysis because only 50 items were considered to observe 5 constructs (Attitudes toward Math, Science, Engineering and Technology, Robotics and 21st century skills).
The correlation matrix showed a determinant different from zero. The KMO index was greater than 0.60 and in particular, KMO index = 0.82. The Bartlett's sphericity test should be statistically significant to perform the EFA and in this case the test statistic is chi squared = 3649.536(1225) and p-value<0.01. To extract the factors underlying the dataset both Principal Component Analysis (PCA) and Principal Axis Factorization (PAF) were performed. Both PCA and PAF can be seen as ways of dimension reduction. In PCA the goal is to create variables (components) that maximize interindividual variance, trying to create an index where people differ most, components are always orthogonal, each component explains non-redundant information and components are linear combination of indicators. In PAF, factor scores at not real scores, instead they are estimates of the underlying latent constructs, that exist but cannot be measured directly. The variables are linear combinations of latent factors and residuals (Fabrigar and Wegener, 2012). Moreover, the purpose of factor analysis is to assist researchers in identifying and/or understanding the nature of the latent constructs underlying the variables of interest. Technically, these descriptions exclude component analysis, which is a method for reducing the dimensionality of a set of observed variables through the creation of an optimum number of weighted composites. A major difference between factor and component analysis is that in the latter all of the variance is analyzed, whereas in factor analysis, only the shared (common) variance is analysed (Bandalos and Finney, 2018).
The Scree test (Figure 14, left) reports the eigenvalues from the correlation matrix. Figure 14 (right) shows the results from the parallel analysis. From

the Adjusted eigenvalues greater than zero it could be chosen the factors to retain. In this case the results show that 7 factors should be retained. Horn’s parallel analysis suggested that 7 are the number of factors to be extracted. Accordingly, the pattern matrix after the oblimin rotation with Kaiser normalization is reported in Table 5.

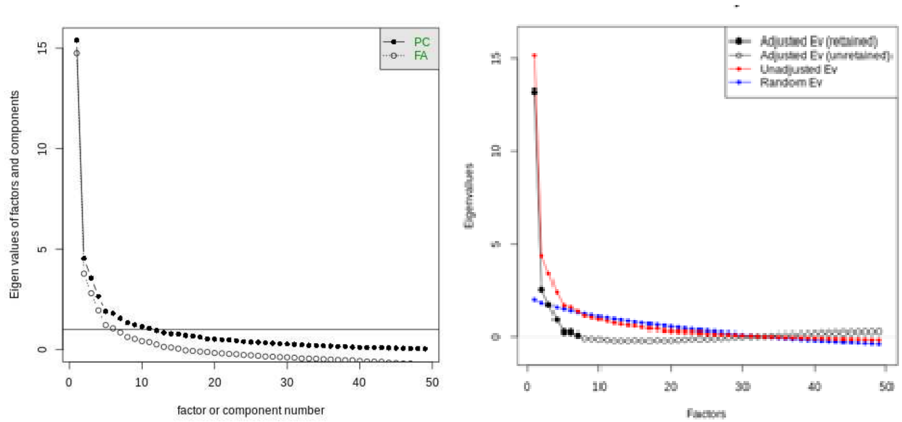


Figure 14: Scree test comparing PCA and PAF results (left) and results from Horn’s parallel analysis (right).

We can see from Table 5 that the items seem to correlate with factors in a coherent way. The cell highlighted in red correspond mostly to a unique factor. Anyway, items loading on more than one factor could be noticed. Gorsuch (1983) notes that second order factors are of interest when they explain 40% to 50% of the extracted variance.
Even if KMO index and Bartlett’s sphericity test let us perform the Factor Analysis, several thumb rules state differently. Tabachnick’s rule of thumb (Tabachnick and Fidell, 2007) suggests at least 300 cases are needed for factor analysis. Hair et al. (1995) suggested that sample sizes should be 100 or greater. A number of textbooks (Pett, Lackey and Sullivan, 2003; Gorsuch, 1983; Tabachnick and Fidell; 2007) deem 100 as poor, 200 as fair, 300 as good, 500 as very good, and 1000 or more as excellent (Williams, Onsman and Brown, 2010; Hogarty, Hines, Kromrey, Ferron and Mumford, 2005).

Table 5: pattern matrix after the oblimin rotation with Kaiser normalization.

	1	2	3	4	5	6	7
Math							
1	0,033	0,815	-0,046	0,148	-0,100	0,021	0,106
2	0,022	-0,601	0,217	-0,183	-0,117	0,063	0,247
3	0,131	0,735	0,209	-0,053	-0,064	-0,019	-0,059
4	0,176	-0,754	0,015	0,159	-0,016	-0,171	0,005
5	-0,110	0,753	-0,066	-0,098	0,069	0,236	0,155
6	0,051	-0,456	0,226	-0,303	0,058	0,023	0,516
7	-0,194	-0,478	0,072	0,134	0,108	0,197	0,033
8	-0,087	-0,794	0,015	0,150	-0,075	0,083	0,073
Science							
9	0,045	0,102	0,313	0,290	0,071	-0,063	0,433
10	-0,072	0,009	0,800	0,038	0,033	0,187	0,073
11	0,024	-0,045	0,846	0,080	0,045	-0,049	0,019
12	0,243	-0,063	0,367	0,055	-0,112	0,099	0,381
13	0,049	-0,065	0,730	0,143	0,070	-0,108	-0,082
14	0,026	0,156	0,235	0,460	0,184	-0,017	0,398
15	0,023	-0,116	0,824	-0,030	0,005	-0,040	-0,006
16	-0,053	-0,050	-0,412	-0,255	-0,002	0,159	-0,210
17	0,004	0,054	0,868	-0,033	-0,020	0,063	0,083
Engineering and technology							
18	-0,097	0,162	0,214	0,031	0,075	0,540	0,063
19	-0,096	0,005	0,318	-0,040	0,461	0,146	0,120
20	0,138	-0,097	0,006	-0,006	0,495	-0,050	0,171
21	0,129	0,012	0,029	0,113	0,675	0,001	-0,138
22	0,151	-0,108	0,370	-0,206	0,461	0,114	-0,148
23	0,062	0,117	-0,015	-0,004	0,601	0,252	0,100
21st century skills							
24	0,336	-0,030	0,133	0,028	0,312	0,067	-0,065
25	0,329	-0,092	0,167	-0,058	-0,061	0,279	0,218
26	0,253	-0,167	0,056	-0,100	0,651	0,076	-0,092
27	0,002	-0,104	0,010	0,102	0,169	0,304	0,508
28	0,010	-0,177	-0,059	0,482	0,084	0,279	0,076
29	0,095	-0,200	0,207	0,097	0,131	0,166	0,206
30	-0,102	-0,098	0,189	0,726	0,028	0,131	-0,097
31	0,206	-0,152	0,165	0,549	-0,256	0,121	0,012
32	0,074	-0,089	0,084	0,492	0,260	-0,154	0,184
33	0,074	0,083	-0,019	0,395	0,523	0,008	0,105
34	0,061	-0,071	0,019	0,462	0,118	0,044	0,140
35	-0,072	0,106	0,174	0,521	0,313	-0,069	-0,032
36	0,287	0,149	0,134	0,358	-0,212	0,052	0,269
37	0,194	-0,123	-0,064	0,561	-0,193	0,091	-0,056
Robotics							
38	0,898	0,184	0,005	0,039	0,036	-0,127	0,025
39	0,755	0,016	-0,005	0,069	-0,018	0,065	0,133
40	0,320	-0,100	0,054	0,058	0,069	0,291	-0,040
41	0,536	-0,126	-0,094	0,021	0,220	-0,076	0,148
42	0,298	-0,077	-0,266	0,104	0,172	0,492	0,042
43	0,450	0,021	0,161	0,039	0,172	0,296	-0,289
44	0,101	0,038	-0,112	0,117	-0,021	0,734	0,026
45	0,447	0,051	0,171	0,017	0,116	0,411	-0,237
46	0,329	-0,010	0,365	0,087	0,187	0,354	-0,360
47	0,340	0,026	0,233	0,066	0,361	0,317	-0,240
48	0,628	0,068	0,142	0,024	0,103	0,200	-0,142
49	0,595	-0,104	0,100	-0,047	0,218	0,161	-0,058
50	0,530	0,071	0,003	0,334	0,141	0,045	0,085

2.3 Questionnaire on attitudes

S-STEM questionnaire is suitable for students of Secondary school, but for primary school students it could be too long and complex to answer. A shorter version of a questionnaire on attitudes was formulated focusing on three main areas. It took the shape of a self-reported expectations and satisfaction questionnaires. Before starting the activities, the expectations questionnaire was administered:

- self-efficacy on building robots (Q1) and programming robots (Q2)
- interest in robotics activities (Q3, Q7, Q8)
- teamwork attitude (Q4, Q5, Q6)

After the completion of activities, the satisfaction questionnaire was administered:

- relationship with the educator (Q1, Q2, Q3)
- self-efficacy on building robots (Q4) and programming robots (Q5)
- interest in robotics activities (Q6, Q10, Q11)
- teamwork attitude (Q7, Q8, Q9)

Answers to this questionnaire were on a 3-point Likert type scale, ranging from “not so much” to “very much”. Several activities into classroom were carried out during 2018 and data from these experiences were collected. To validate these questionnaires, two separate Factor Analysis were performed, one for the expectation questionnaire and another on for the satisfaction questionnaire.

2.3.1 Expectation questionnaire

The correlation matrix was computed (Figure 15) with a determinant very small but different from zero (d=0.096). The KMO measure was 0.76 and Bartlett’s sphericity test resulted in a chi squared = 413.8488 (df= 28) and p-value<0.01. Cronbach alpha as a measure of reliability is 0.75. The chosen method for factors’ extraction was the Principal Axis Factoring and results are shown in Figure 16 along with results from Horn’s parallel analysis. The number of retained factors at the end of the analysis is three. Factors’ loadings are reported in Figure 17.

0	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0
1.0	1.000000	0.420450	0.270490	-0.019704	0.097295	0.094484	0.154574	0.269347
2.0	0.420450	1.000000	0.359412	0.063092	-0.020427	0.035371	0.257309	0.353375
3.0	0.270490	0.359412	1.000000	0.290070	0.183520	0.330473	0.627941	0.556583
4.0	-0.019704	0.063092	0.290070	1.000000	0.318794	0.250083	0.282358	0.228232
5.0	0.097295	-0.020427	0.183520	0.318794	1.000000	0.476793	0.293349	0.161652
6.0	0.094484	0.035371	0.330473	0.250083	0.476793	1.000000	0.388392	0.308535
7.0	0.154574	0.257309	0.627941	0.282358	0.293349	0.388392	1.000000	0.617507
8.0	0.269347	0.353375	0.556583	0.228232	0.161652	0.308535	0.617507	1.000000

Figure 15: correlation matrix for the expectation questionnaire.

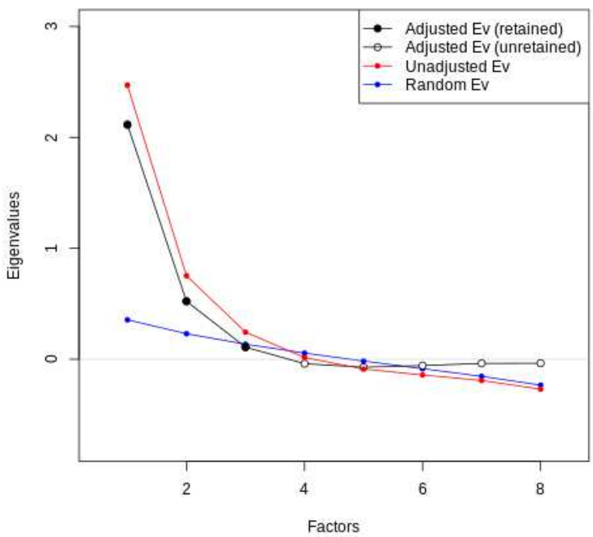


Figure 16: results from Horn's parallel analysis of the expectations questionnaire.

Loadings:

	PA1	PA2	PA3
X1.0			0.780
X2.0	0.284	-0.150	0.466
X3.0	0.749		
X4.0	0.316	0.247	-0.130
X5.0		0.903	
X6.0	0.318	0.437	
X7.0	0.834		
X8.0	0.722		0.103

Figure 17: factors' loadings for expectations questionnaire.

Items 1 and 2 (“It will be easy to build robots and “It will be easy to use the software on computers”) are observed variables from factor PA3; items 3 (“Can’t wait to do the activities of this laboratory), 7 (“This activity seems to me fun and engaging”) and 8 (“I’d like to do more activities on Robotics in the future”) are from factor PA1 and items 5 (“In my group there will be cooperation and support”) and 6 (“Working in group will improve my relationship with one or more of my classmate”) loads on PA2. Item 4 (“Classroom climate during the activity will be positive”) seem to match no one of the factors. Ideally, we would expect it to be loading on PA2 which we would associate with teamwork attitude. From this analysis we can expect PA1 to be related with interest in robotics activities, PA2 with teamwork and PA3 with self-efficacy in Robotics.

2.3.2 Satisfaction questionnaire

The correlation matrix was computed (Figure 18) with a determinant very small but different from zero (d=0.051). The KMO measure was 0.77 and Bartlett’s sphericity test resulted in a chi squared = 446.8081 (df= 55) and p-value<0.01. Cronbach alpha as a measure of reliability is 0.75. The chosen method for factors’ extraction was the Principal Axis Factoring and results are shown in Figure 19 along with results from Horn’s parallel analysis.
The number of retained factors at the end of the analysis is five. Factors’ loadings are reported in Figure 20.

0	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0
1.0	1.000000	0.362543	0.420236	0.235511	0.255086	0.352153	0.204185	0.156223	0.180467	0.368748	0.307989
2.0	0.362543	1.000000	0.371120	0.084449	-0.115378	0.301037	0.271548	0.310268	0.210317	0.277947	0.202024
3.0	0.420236	0.371120	1.000000	0.063807	0.178832	0.433815	0.146981	0.126663	0.127120	0.463742	0.450474
4.0	0.235511	0.084449	0.063807	1.000000	0.295085	0.137992	0.130776	0.068023	0.035734	0.020618	0.083688
5.0	0.255086	-0.115378	0.178832	0.295085	1.000000	0.066690	-0.018009	-0.180473	0.047966	0.121441	0.169562
6.0	0.352153	0.301037	0.433815	0.137992	0.066690	1.000000	0.376590	0.268501	0.222032	0.659894	0.578895
7.0	0.204185	0.271548	0.146981	0.130776	-0.018009	0.376590	1.000000	0.363212	0.202307	0.156058	0.182101
8.0	0.156223	0.310268	0.126663	0.068023	-0.180473	0.268501	0.363212	1.000000	0.390654	0.175397	0.151257
9.0	0.180467	0.210317	0.127120	0.035734	0.047966	0.222032	0.202307	0.390654	1.000000	0.252495	0.127779
10.0	0.368748	0.277947	0.463742	0.020618	0.121441	0.659894	0.156058	0.175397	0.252495	1.000000	0.617168
11.0	0.307989	0.202024	0.450474	0.083688	0.169562	0.578895	0.182101	0.151257	0.127779	0.617168	1.000000

Figure 18: correlation matrix for the satisfaction questionnaire**.**

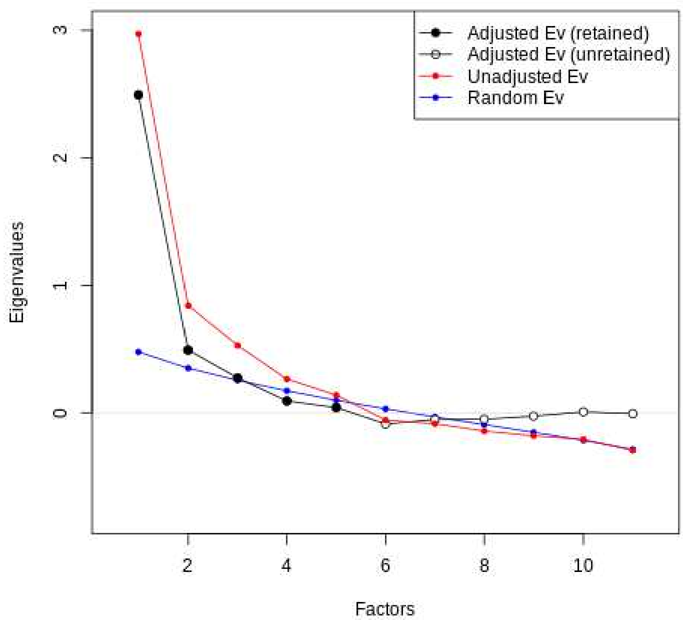


Figure 19: results from Horn's parallel analysis for satisfaction questionnaire.

Loadings:

	PA1	PA4	PA2	PA3	PA5
X1.0		0.514	0.295		
X2.0		0.700	-0.145		
X3.0	0.336	0.465			-0.143
X4.0		0.129	0.404		0.292
X5.0			0.826		
X6.0	0.805				0.253
X7.0	0.148	0.110		0.125	0.487
X8.0		0.130	-0.173	0.428	0.300
X9.0				0.758	
X10.0	0.811				-0.183
X11.0	0.724				

Figure 20: factors' loadings for the satisfaction questionnaire.

Figure 20 shows that items 1 (“I understood all the instructions that the educator said”), 2 (“The educator was open to questions and needs”) and 3 (“The way educator carried out the activity was engaging”) are observed variables from factor PA4; items 4 (“It was easy to build robots”) and 5 (“It was easy to use the software on computers”) are from PA2; items 6 (“I enjoyed the activities of this laboratory”), 10 (“This activity was fun and engaging”) and 11 (“I’d like to do more activities on Robotics in the future”) are from factor PA1; items 8 (““(“In my group there was cooperation and support”)”) and 9 (“Working in group improved my relationship with one or more of my classmate”) loads on PA3. Item 7 (“Classroom climate during the activity was positive”) seem to match no one of these factors. Ideally, we would expect it to be loading on PA3 which we would associate with teamwork attitude, instead it loads on a completely different factor PA5.

Item 4 in the pre-activity questionnaire (Expectations) and Item 7 of the post-activity questionnaire (Satisfaction), namely “*The classroom climate will be/was positive*”), is related to the classroom climate. It seems to do not match with others existing factors. That is probably because classroom climate can be associated not only with a teamwork attitude, but also with the prevailing mood, attitudes, standards, and tone that teachers and students feel when they are in the classroom. A negative classroom climate can feel hostile, chaotic, and out of control. A positive classroom climate feels safe, respectful, welcoming, and supportive of student learning. So, not just peer relationship, but also relationship with the teacher. Not only relationship, but also shared values and rules. Also, the physical learning environment, with its structure can contribute to classroom climate. For these reasons, the classroom climate could be considered another factor, which was not expected in the first place, but not unwelcome.

Section 3 – Case studies

In this section some examples of how to apply the instruments for the measurement of learning during ER activities will be presented. The instruments shown in the previous section will be applied in to K12 case studies where ER is the main focus of the classroom activity. Each of the case studies focus on a different research question and therefore uses a different measurement tool. Procedure in the classroom is not reported fully since the focus is on the instruments and their results.

3.1 Case study: questionnaire on attitudes

In the previous section the expectation and satisfaction questionnaires were described and validated. In this section results from a measurement campaign by means of these sensors are shown.
In the years 2017-2018 eight Italian schools (4 Istituti Comprensivi, which are institutions grouping both Primary and lower Secondary Schools, and 4 upper secondary schools) located in two regions, Marche and Emilia Romagna were involved in a series of projects whose aim was to realise some activities of Educational Robotics in the classroom. 198 students were involved. Their age ranged from 10 years old to 20 years old. Background variables included in the survey were: 'group' in which the student worked in, 'genre' (male, female), 'age' in years, 'favourite subject at school' (socio-humanities or techno-scientific subject), use of technology in class or at home: 'tablet', 'smartphone', 'audio-video editor', 'text-image editor', 'browser', 'Arduino', 'coding', 'software for educational purposes', 'robot', 'other technological devices or software application'; also social media were investigated by listing some of the most popular until 2018: 'Facebook', 'Twitter', 'Instagram', 'Messenger', 'Whatsapp', 'Snapchat', Youtube' and 'other social media'. Descriptive statistics of the sample are reported in Table 6 and Figure 21.

Table 6: Descriptive statistics of the sample.

Gender		
	M (%)	67,32673
	F (%)	32,67327
Age		
	Mean	14,72277
	Standard Deviation	2,409
Favourite discipline		
	Socio humanities	22,11055
	Techno-Scientific	76,88442

To check whether there is an association between the four factors in the expectation questionnaire (See Section 2) and the background variables a chi squared test was performed (Figure 22).

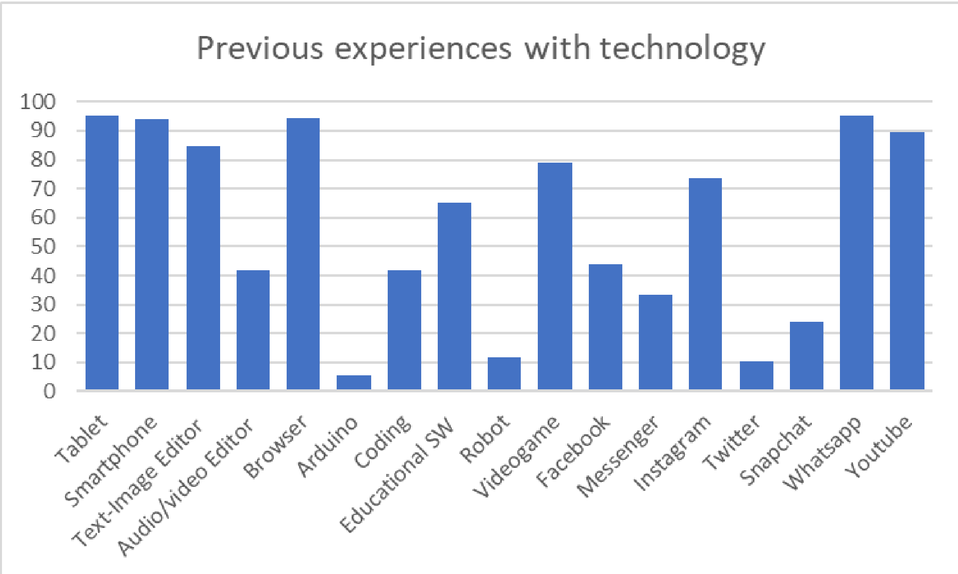


Figure 21: previous use of technologies and social media among the students of the sample.

In Figure 22 the four factors analysed were self-efficacy on building robots (Q1) and programming robots (Q2), interest in robotics activities (Q3, Q7, Q8), teamwork attitude (Q5, Q6) and classroom climate (or environment, Q4).

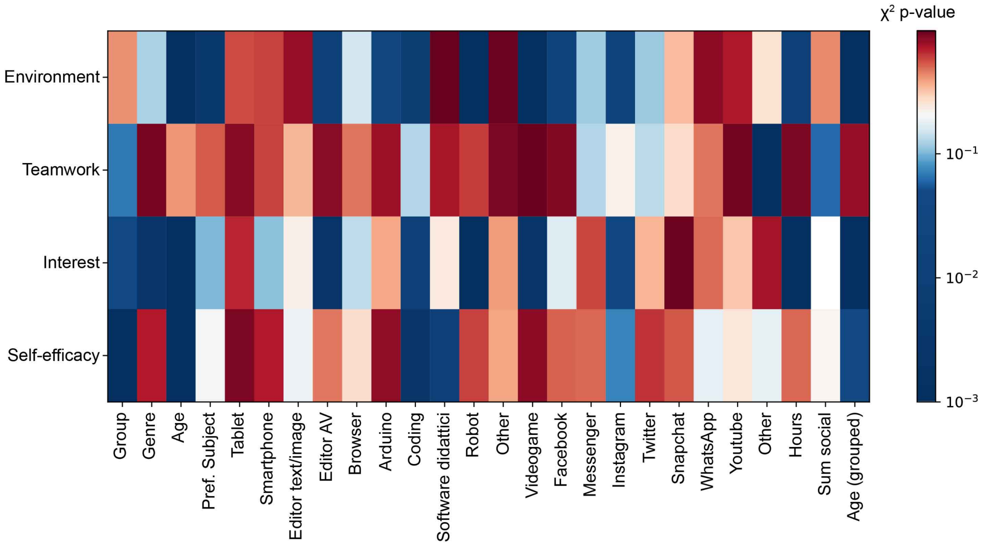


Figure 22: chi squared results for expectations vs. background variables are displayed.

Interestingly, significant association between self-efficacy and coding was found but not previous experiences with robots. Age is significantly influencing both self-efficacy, classroom climate and interest.

Another chi square test was performed to check whether there is association between the five factors in the satisfaction questionnaire and the background variables (Figure 23). Results from this second test seems to report that there is no association between control variable and final results of the test, but something change if the perspective vary to include a basic measure of change, namely the difference between post measures and measures before activities take place (Figure 24).

To check whether the values of the reported measures of the four factors (interest, teamwork, self-efficacy and classroom climate) on students are significantly different in the pre-activities condition and in the post-activity condition, the paired samples Wilcoxon test (also known as Wilcoxon signed-rank test), a non-parametric alternative to paired t-test, was performed (Wilcoxon, 1945) by means of the statistical software R (Table 7). Figure 25 shows the distributions of factors in pre-condition and in the post-condition.

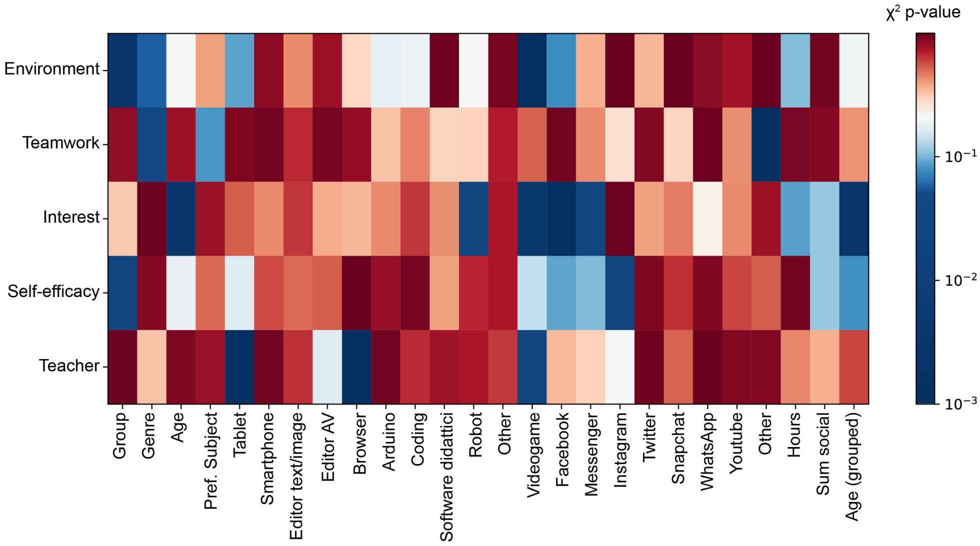


Figure 23: chi squared results for satisfaction vs. background variables are displayed.

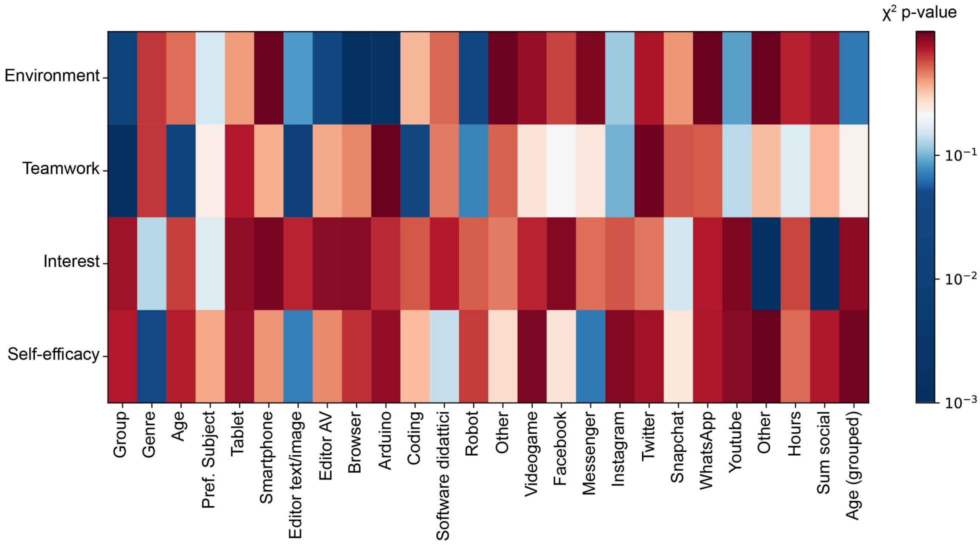


Figure 24: chi squared results for change in attitudes vs. background variables are displayed.

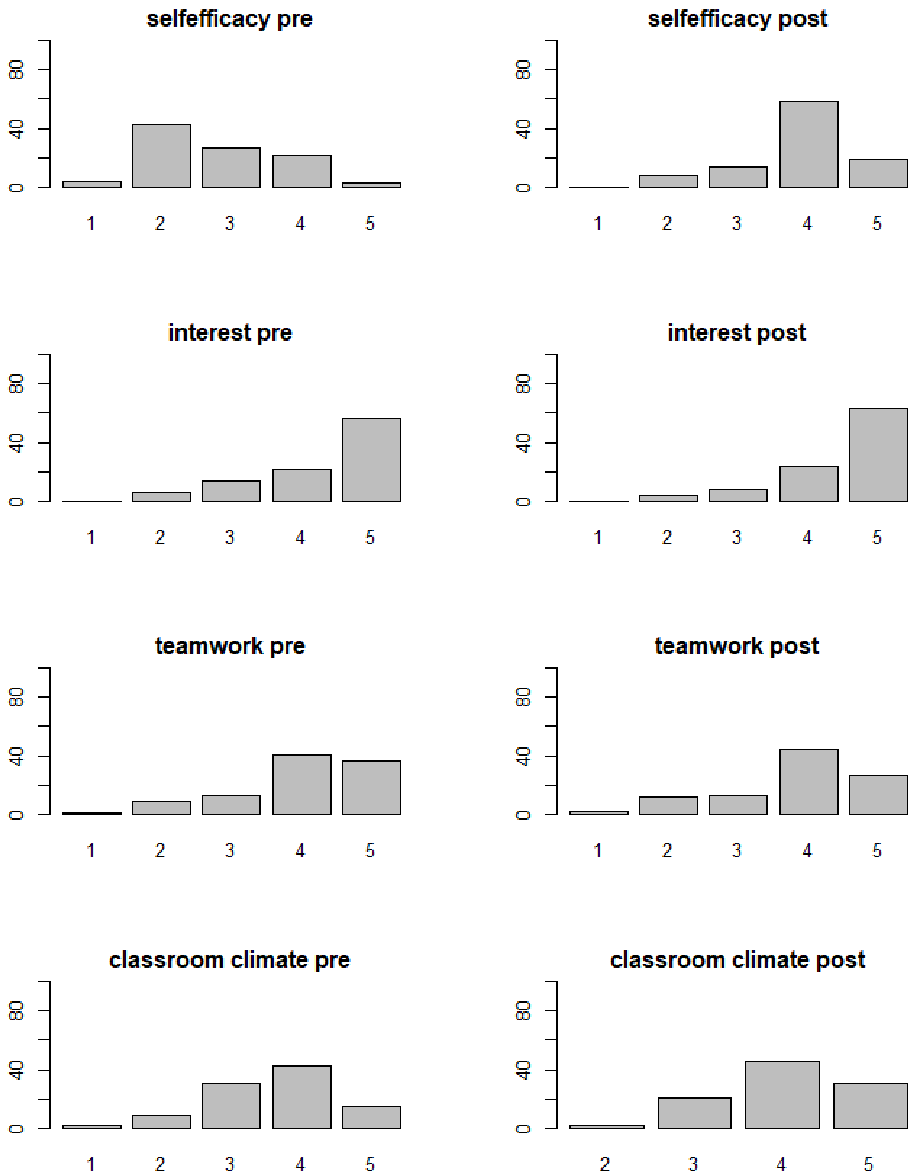


Figure 25: distributions of scores on the 4 factors in pre and post condition.

Table 7: descriptive indices and results of the paired samples Wilcoxon test.

Factor	***Median value post***	***Median value pre***	***Mean value (standard deviation) post***	***Mean value (standard deviation) pre***	***Test statistic***
Self-efficacy	4	3	3.87 (0.83)	2.77 (0.96)	7903.5**
interest	5	5	4.46 (0.85)	4.28 (0.97)	1935.5*
teamwork	4	4	3.81 (1.05)	4.04 (0.97)	1460**
classroom climate	4	4	4.04 (0.79)	3.60 (0.93)	3997**

** p-value<0.05, ** p-value<0.01*

The ER activities seem effective in promoting self-efficacy in coding and building robots, interest in robotics and classroom climate. This can be related to the methodology of bringing ER into the classroom. Activities in the classroom had a hands-on approach focusing on fundamentals of coding and robotics engaging students in playful learning (Resnick, 2004) while cooperating to solve a real and concrete problem. A typical ER activity follow the TMI approach (Martinez and Stager, 2013) where students Think about what they want to accomplish a goal, realise their idea and observe the functioning of what they made to improve it. In this process the error is not seen as something that hampers learning or a social stigma. Error is a precious part of learning. In a laboratory it is fine to make mistakes and have a trial and error approach. What empowers learner is the capacity of observing errors, think about how to fix them and learning from those error. The results of their work become evident very quickly, and they can check by themselves if they are right or wrong. By experimenting, they discover that making mistakes is part of the learning process. Error is part of the process of learning. While they learn that failure is necessary in every learning process, they develop their capacity for resilience and overcome the fear to make mistakes. They gain autonomy by creating their own robots and solving different problems by themselves, while they learn and have fun. This may greatly contribute to have a more positive classroom climate.

Teamwork attitude seems to slightly decrease form pre to post measurement. This trend can be related to the kind of questions asked: “I will have fun with my fellows” and “In my group there will be cooperation and mutual support” in the pre measurement; “Did you had fun with your fellows?” and “Did you cooperated in your group?” in the post measurement test. The implementation of ER activities is said to foster social skills as well as scientific and technological competencies. Along the process of working in groups students understand that the results they want to achieve are much more viable if they work together. They understand and experiment the importance of committing to the project they are working on, to be patient and persistent, to negotiate. All these processes interact with already existing relationships. It would be interesting to dig it deeper into this kind of issue with other tools representing the social structure of the classroom.

3.2 Case study: S-STEM questionnaire and knowledge test

Attitudes toward STEM can have a strong impact on future involvement in STEM careers and engagement in school (Daniela, Strods & Alimisis, 2017; Bagattini, Miotti, & Operto, in press). During the year 2019 summer camps and outreach activities were carried out thanks to the national call for projects “in Estate si Imparano le STEM” (2018) from the Italian Department for Gender Equality (Dipartimento per le Pari Opportunità della Presidenza del Consiglio dei Ministri – DPO) in collaboration with the Italian Ministry of Education (Ministero dell’Istruzione, dell’Università e della Ricerca - MIUR) and PON Digital Citizenship and Creativity (Cittadinanza e creatività digitale, 2017).
Three schools from the Marche region, placed in the centre of Italy, were involved with a total number of participants of 76 students described as a sample in Table 8 and Figure 26. All participants were voluntary students willing to participate in an extra-curricular activity in the summertime (from June to September), when school is closed, and usually Italian students are involved in playful activities at home and at the seaside.
The call for projects required to involve students in meaningful activities that could enhance the probabilities that they will be interested in taking a STEM career in the future and that could help developing digital citizenship and creativity. Robotics was chosen to foster attitudes toward STEM career and to increase their knowledge and competence on technology as a way to

grow digitally competent citizens that will be able to master technology to create the future job, adapt to the fast-evolving society of knowledge and interact with the digitized Italian Public Administration.

Table 8: descriptive statistics of the sample.

Age [mean (standard deviation)]	11.52(1.32)	
Age groups	Primary school	30.14%
	Secondary school	69.86%
Gender	F	53.42%
	M	46.58%
Favourite Discipline	Socio-Humanities	35.62%
	Techno-Scientific	61.64%
	Both	2.74%
Schools	A	31.58%
	B	34.21%
	C	34.21%
Distribution of subjects over schools and age groups		
	<11 years old	**>11 years old**
A	5.48	26.03
B	0	32.88
C	24.66	10.96

School A and C were involved in a two-week program (5 days a week, 4hours a day), while school B was involved in a ten meetings program. School A and C had a classroom with students from primary and lower secondary school, while school B had students only from lower secondary school.

Activities were carried out in the schools' classrooms. The playful approach was perfectly suited for the occasion and the exploration of tangible technological materials engaged students in the daily activities and challenges. Topics covered during the lessons were:

- Definition of robot
- Sensors and actuators as inputs and outputs for the centralised controller
- Comparison of human life and artificial life: similarities and differences
- Building a mechanical structure which can realise a movement

- Programming the robot to do simple movements in the environment (sequential, conditional and iterative instructions) with and without the use of sensors

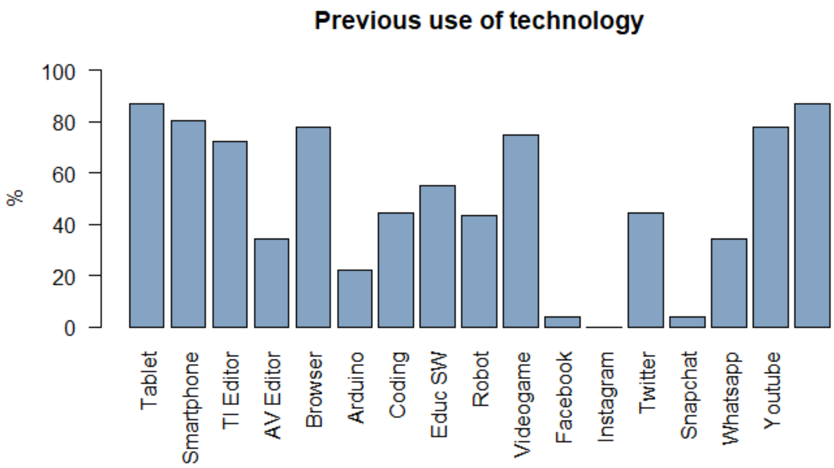


Figure 26: previous experiences with technologies and social media of participants.

The structure of the daily activities was similar in the three experiences:

- a short introduction made by the educator showing basic notions to apply during the subsequent mini-activity (for example how to set the motors parameters before programming the robot to stop at a given distance) or posing a problem to be solved with what they have learnt so far;
- a first challenge to warm up;
- debriefing with the all class and discussion on the artefact that each group created and the performance it had to learn to share and learn from common experience;
- depending on the length and complexity of the first challenge, another mini-activity was carried out with the same procedure of the first three point reported above in this list.

During the first day of the project groups were made by the teacher in the classroom following observation made during the school year for students she/he knew, or simply trying to balance groups in terms of age.

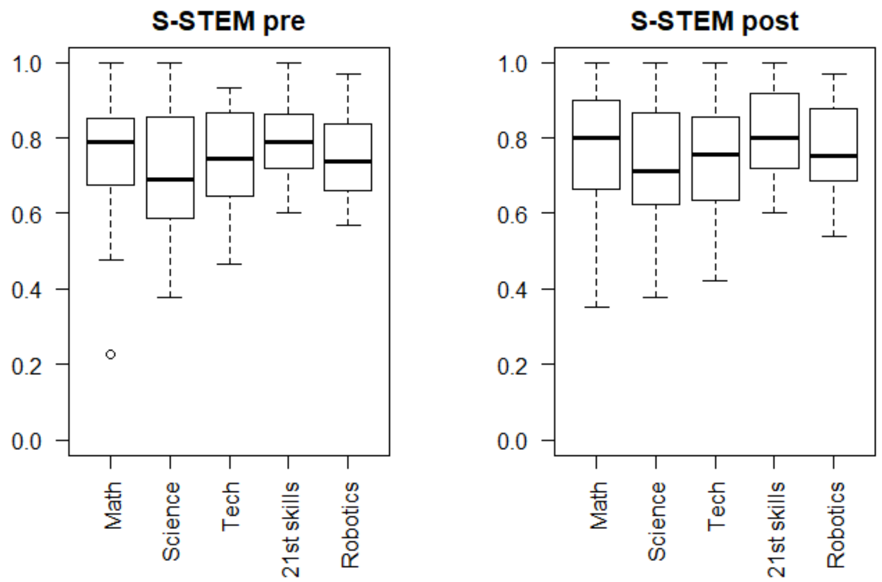


Figure 27: boxplots comparing results in the five factors of adapted S-STEM in pre and post condition.

During the first day and last day of the project, students were presented with the measurement instruments. At the pre-activity condition a survey on student's demographic and background variables, the Italian version of the S-STEM questionnaire on attitudes towards STEM studies and career (Appendix C) and a knowledge questionnaire (Appendix B) were administered. After the project had ended, during the last day of the project, another measurement was made (post-activity measurement). The S-STEM questionnaire was re-administered in the same version to students, while another knowledge questionnaire, equivalent to the first was administered (Appendix B). Data from these assessments were collected on paper by the educator in the room, who provided instruction on the structure of the questionnaires and debriefed about the importance of answering sincerely to both questionnaires; it was made clear that neither the questionnaires on attitudes, nor the knowledge test were going to be formally evaluated by school or any other person to judge them in an official manner. Paper from these instruments was digitised and data was entered in standardised tables. After the data entry phase, data from the tables were processed as follow. Missing data were discarded and the sum of the items on each of the five factors of the Friday questionnaire was computed for each case, then normalised to the highest possible score for the factor and subsequently recoded into three categories. Scores in the five constructs of the S-STEM questionnaire are reported in the boxplots of Figure 27. To capture these distributions and assess change between these scores within the factors of the same measurement condition and between the same factor in the precondition and post condition, a Kruskall Wallis test was performed (Table 9). Figure 28 shows the results from the distributions of scores on the three classes.

Table 9: results from the Kruskall-Wallis test.

	pre	**post**	**change**
Math	23.07(2)**	24.153(2)	17.139(3)**
Science	19.503(2)**	20.677(2)**	13.159(3)**
Engineering	22.753(2)**	23.267(2)**	16.226(3)**
21st century skills	23.372(2)**	23.998(2)**	18.954(4)**
Robotics	21.95(2)	22.834(2)**	15.757(3)**

* pvalue<0.05; **p-value<0.01

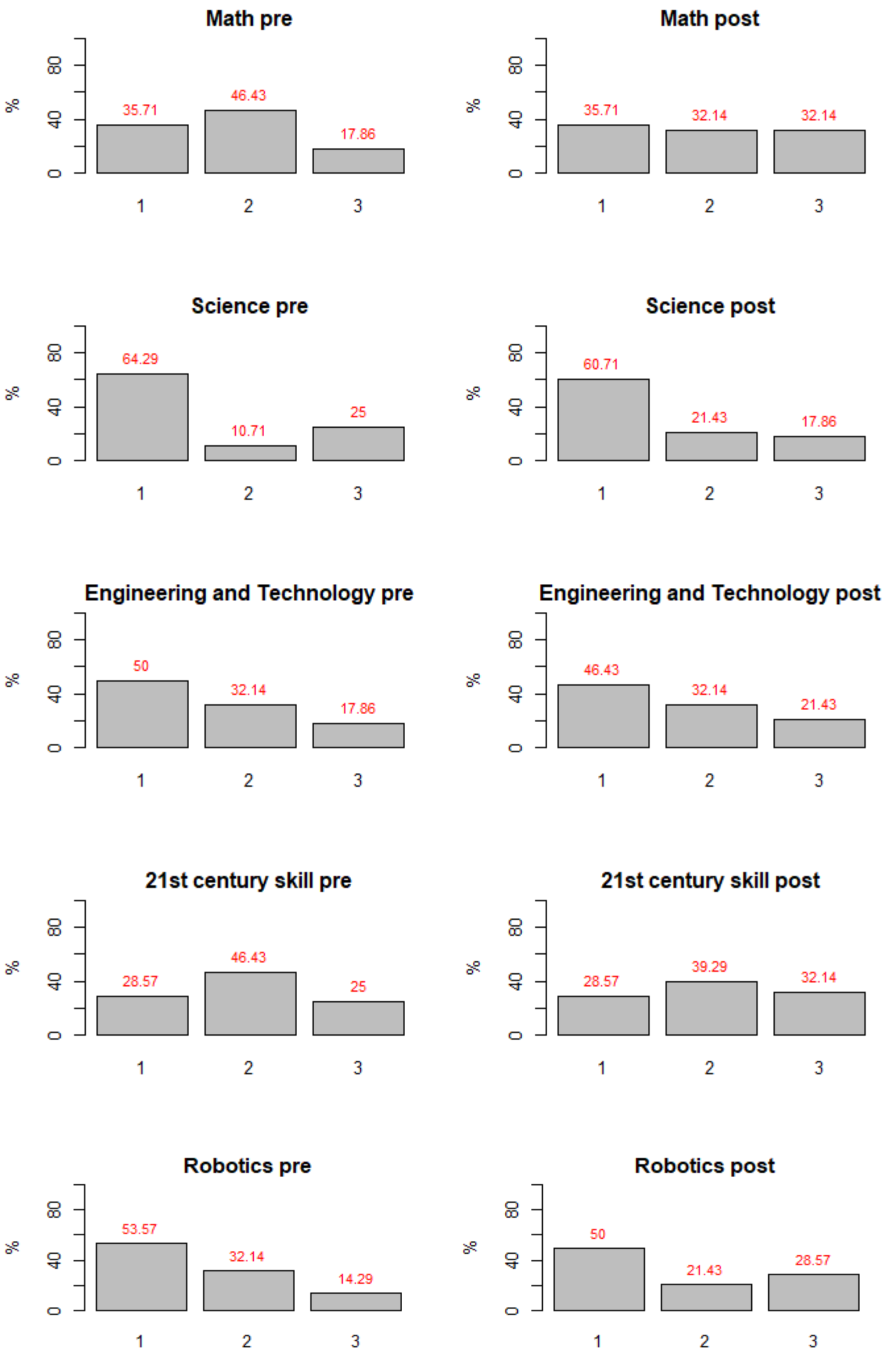


Figure 28: distribution of scores among the classes and between the factors.

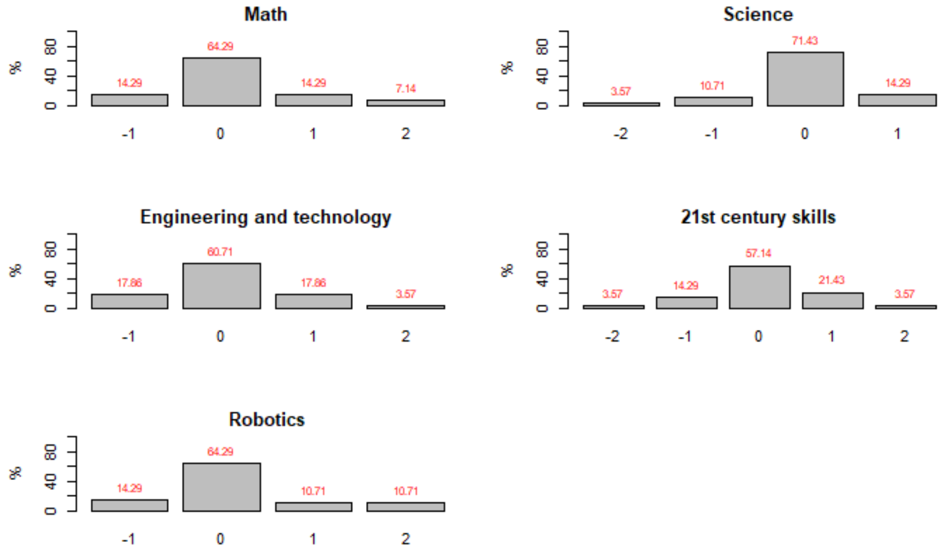


Figure 29: changes between classes from post measurement to pre measurement.

Data in Figure 29 shows that the experience with ER mostly did not change the attitude toward STEM, Robotics and 21st century skills, but there were significant improvements in some cases. The greater improvement seems to be reported in attitudes towards Robotics where the 21.42% improved his/her own attitude. This is a positive result that accounts for the effectiveness of ER on increasing interest, self-efficacy and expectancy-value in the area of Robotics. On the contrary, this result seems to not provide support for the thesis that Robotics experiences at school represent a panacea for general STEM attitudes improvement.
Data from the knowledge questionnaire was digitised and entered in a table. Each item was scored 1 point if correct, 0 if wrong. Some questions presented subparts to solve; this kind of questions were evaluated summing the correct answers to the subparts and then dividing by the number of possible answers. The total score of the questionnaire was 12. Questions from 1 to 10 aimed at testing knowledge, questions 11 and 12 aimed at testing competence. In the analysis of the questionnaire, scores were computed for two categories: knowledge and competence (Figure 30). Two examples of the answers that students provided to the competence questions are shown in Figure 31 and Figure 32. Subsequently, scores were

categorised into 3 classes of equal portion of the possible range. The result from this categorisation are reported in the boxplots of Figure 30.

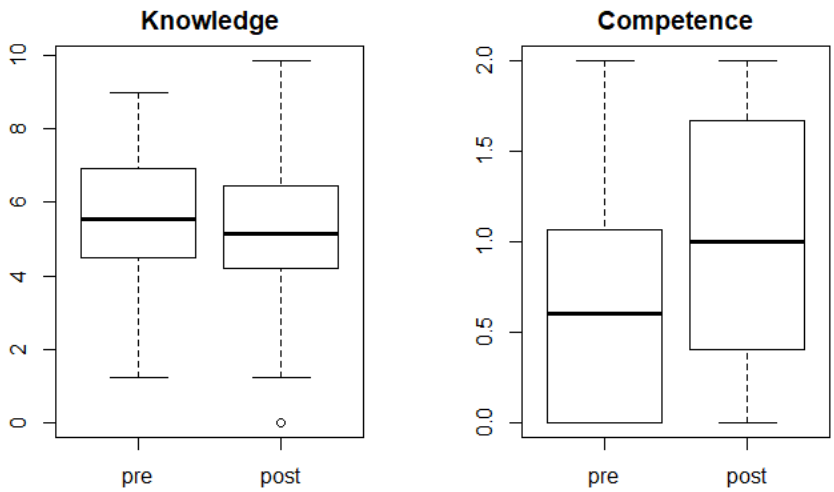


Figure 30: boxplot of scores from knowledge and competence test.

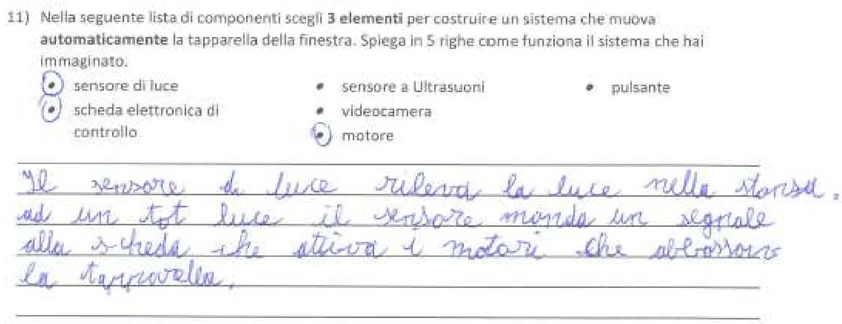


Figure 31: example of answer to question 11.

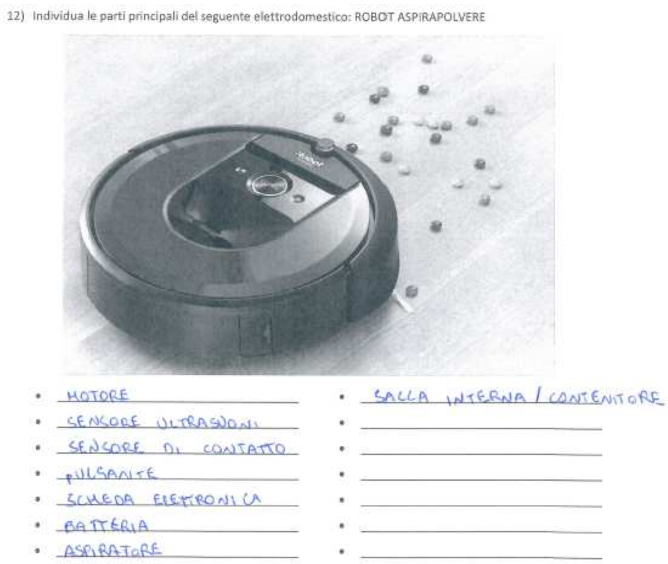


Figure 32: example of answer to question 12.

To capture these distributions and assess change between these scores within the factors of the same measurement condition and between the same factor in the precondition and post condition, a Kruskall Wallis test was performed (Table 10).

Table 10: results from the Kruskall Wallis test.

	pre	post	change
Knowledge	45.108(2)**	50.582(2)**	47.251(3)**
Competence	57.984(2)**	60.433(2)**	54.53(4)**
* p-value<0.05; ** p-value<0.01			

Surprisingly, the 28.99% worsened the performance on Knowledge, against the 15.94% that performed better after the intervention (Figure 34). On the other hand, the 34.78% of the sample performed better after the ER project on competence (Figure 34).

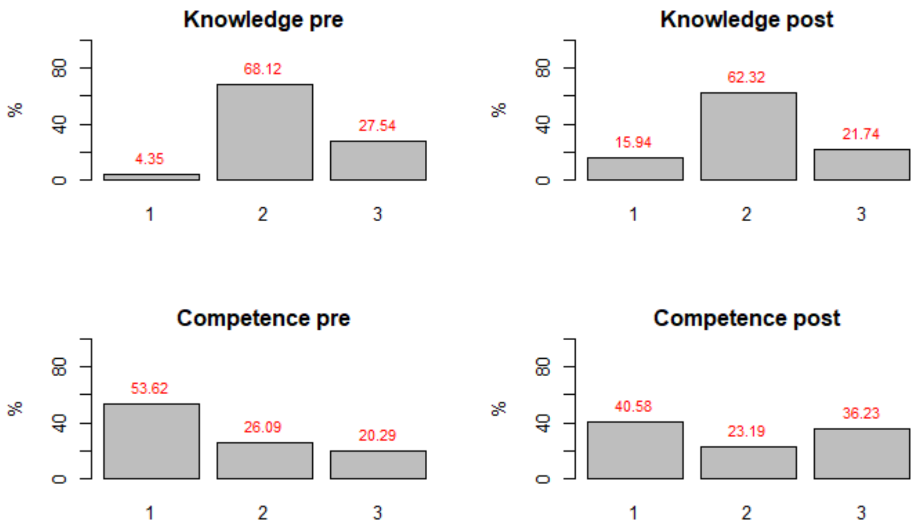


Figure 33: scores distributions**.**

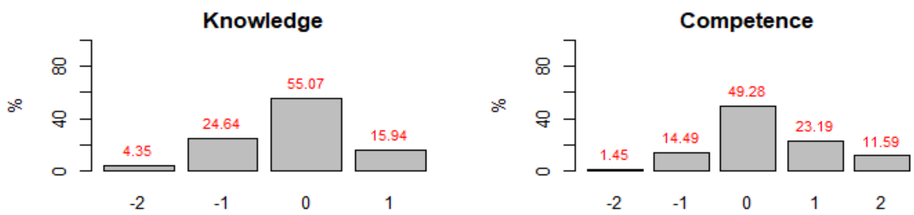


Figure 34: distribution over the simple measure of change**.**

<u>This result in the knowledge test may be due to several factors</u>. First of all, the type of activity is extra-curricular and, even if the walls of the school were the very same during the school year, the learning environment was not the actual classroom in which students work during the school year. Second, the summer camp is a non-formal occasion for learning while entertaining. Students might not have had the full focus on the test on paper. On the other hand, students that in the middle of their vacancy chose to participate in a summer camp on robotics had a strong motivation. In fact, S-STEM questionnaire shows good attitudes in all the five areas (Figure

27), in particular a mean score of 75% in the Robotics factor at the pre-activities condition.
If the present instruments were included in a feedback system to the teacher, the reasoning of an experienced observer and planner of activities, like the teacher himself/herself, would have another element to provide support and guidance to the class.

3.3 Case study: Teachers' training

Teachers' training is fundamental to provide quality education to students. New methodologies based on technologies promise to transform learning, but very often they become buzzwords rather than a real improvement in schools. To ensure that these new methodologies could be effective and sustainable, teachers should be equipped with the basic knowledge and skills to choose from their toolkit the best instrument to inspire and engage his or her students in learning. The Figure 35 represents the flow of information from experts to teachers to students.

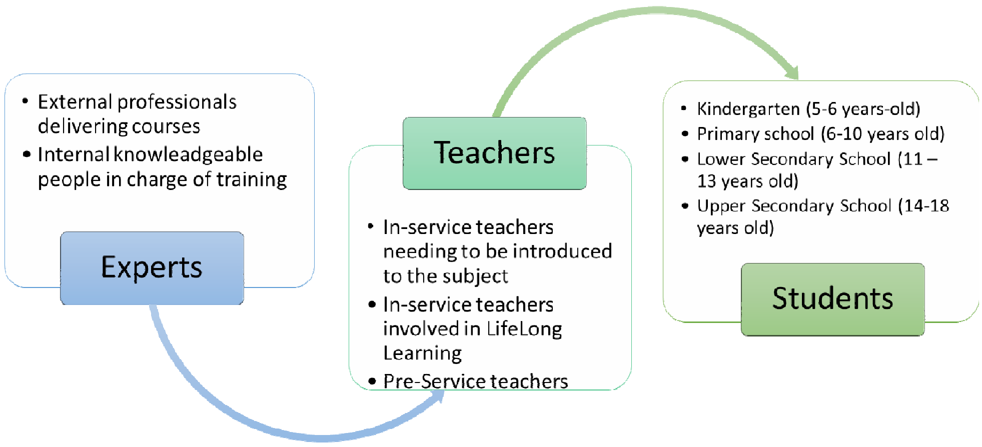


Figure 35: flow of information from experts to teachers and then to students.

In 2017 a training course about new methodologies was realised targeting in service teacher of the province of Teramo (Italy). Some findings about this

work can be found in Scaradozzi, Screpanti, Cesaretti, Storti and Mazzieri (2018).
The objectives of the training course are specified as follows:

- to provide teachers with key notions from Educational Robotics, Coding and Tinkering and examples of the tools and techniques that are proper to each one;
- to improve teachers' attitude towards technology and its usage in the classroom;
- to evaluate how much they have learnt from this course.

The course was designed to introduce teachers to each methodology, make them experience a specific activity based on that methodology and assess them anonymously.
The Research Question (RQ) is: "Is the course effective in providing knowledge and improving self-confidence on some innovative methodologies (Educational Robotics, Coding, Tinkering)?". We expect to improve both self-confidence and knowledge on the 5 innovative methodologies. We also monitor other background and demographic variables that may influence the change of self-confidence and knowledge.
The research design used in this case study is a quasi-experimental one group pretest/postest design (no control group). Sample was not chosen randomly (convenience sampling), but only people volunteering participated providing their consent and their data. Courses were delivered to almost 400 teachers in May - June 2017 in Silvi and Giulianova, Teramo (Italy) with the schedule reported in Figure 36.

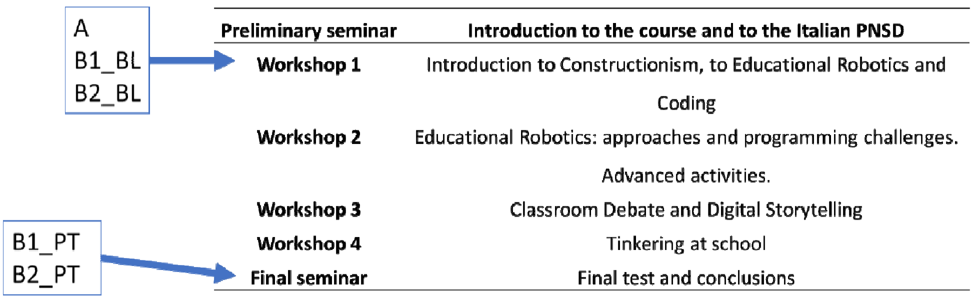


Figure 36: scheduled activities of the training.

General information was registered through a demographic questionnaire (A). Data collected from this form concerns personal information (age, sex, educational qualification, background), information regarding professional life (years spent in the career of teaching, educational stage at which he or she is currently teaching and which subject(s). Results are shown in Table 11.

Table 11: descriptive statistics of the sample.

Age (years)	Mean (dev. std.)	47.9 (8.4)
In service (years)	Mean (dev. std.)	17.6 (10.6)
Gender (%)	F	88
	M	12
Level of Education (%)	Diploma	39
	Laurea Triennale	3
	Laurea Magistrale	51
	Master I/II Livello	7
Background (%)	Science & Technology	17
	Humanities & Social Studies	83
Teaching educational stage (%)	Scuola dell'Infanzia	18
	Scuola Primaria	53
	Scuola Secondario di Primo Grado	6
	Scuola Secondaria di Secondo Grado	23
Teaching (%)	Support teaching	15
	Science and Technology (ST)	14
	Human, Arts and Soc. Studies (HASS)	40
	Both ST and HASS	10
	Other	21

This kind of data are collected to describe our sample along some meaningful dimensions in order to analyse the relation between the outcomes of this course and the composition of the sample. To avoid all possible interferences the form was delivered to teachers during the first seminar, while they were still unaware of the contents of the course. Survey A asked also about the use of technology in the classroom and results are displayed in Figure 37 where the columns represent how many teachers use that technology into his/her class. "Robot" stands for use of robots, either in kit or already assembled, in classroom for activities. "LIM" is the Italian acronym for the electronic whiteboard. "Tablet/PC" stand for the use of PC or tablet in his/her teaching activities. "Boards" mean electronic boards like Arduino, Raspberry and many other choices the market offer. "SW_Coding" is the software dedicated to computer progamming, (i.e. Scratch).

"SW_Didactical" is the software dedicated to didactical purposes (for example, learning geometry, to visualize history and geography, ...). "Text Editor" is the category for software dedicated to text editing. "AV_Editor" means the software for Audio-Video editing. "Other" collects all those technologies teachers couldn't classify satisfactorily in the previous categories.

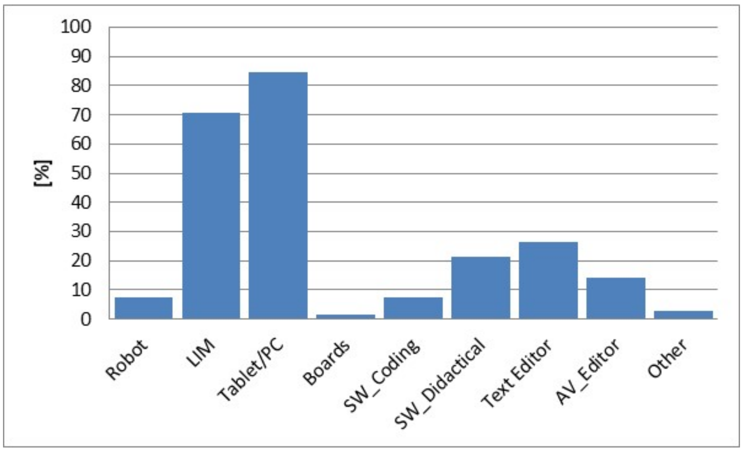


Figure 37: previous use of technology in the classroom.

The other instrument is a questionnaire (B) which will be split in two parts, B1 and B2, to highlight the difference on what they are aiming to test: B1 aims to test the attitude of teachers towards the theoretical approach, tools and design capabilities regarding the proposed methodologies, B2 the knowledge they have about them. Both of these questionnaires were presented to teachers in two different moments of the course: at the very beginning and at the end of the course, after all the workshops had ended. The first measure, namely the questionnaire proposed at the very beginning, is providing the entry level for this course and provides "Baseline" (BL) for considerations. The last measure, that at the very end of the course, provides the "Post-Training" (PT) measure.
B1 delves into the subjective perception of teachers' proficiency with respect to the content of the training course. It was presented to teachers twice: the first time, during the preliminary seminar (B1-BL), the second, during the Final seminar (B1-PT). This questionnaire explores three dimensions of each methodology (theoretic proficiency, operative competence and design skills). In fact, for each methodology the

questionnaire asked three questions: one on the proficiency on the theoretic knowledge (i.e. How much do you know about Educational Robotics' methodologies?), one on the mastery of techniques and tools (i.e. How much do you know the tools and techniques (i.e. software, kits, schemes of evaluation...) of Educational Robotics?) and one on the capability to design a didactic path employing their knowledge on the tools and techniques of that particular methodology (i.e. Do you think you can design an Educational Robotics' activity for your class?). The research interest here lied in detecting the change in the subjective perception of the teacher caused by the training they have undergone and to evaluate whether and how the accuracy of the teacher in assessing his/her own competence varied according to the participation to the course. To this end a 10-points Likert scale with 0.5 points sensitivity was chosen. This kind of answer allows to transfer the intensity of an attitude or of a perception in a numeric value which relates to a scale with a high degree of precision. So, teachers were enabled to give their opinion or position on the question expressing the degree of detail they wanted. Likert scales are widely employed with specific adaptations in various fields which are interested in measuring the attitudes or personal behaviours towards some statements, from industry (finance, marketing…) to research (medicine, biology, engineering…). However, during the data processing phase, this scale was discretized and divided in five classes: class I was associated with very low level, class II with low level, class III with medium level, class IV with good level, class V with a very good level.

B2 aimed at detecting basic knowledge about the methodologies and the tools which was designed to be the subjects of the course. It was presented to teachers twice: the first time, during the preliminary seminar (B2-BL), the second, during the Final seminar (B2-PT). Two questions per methodology were formulated: one testing teachers' knowledge on the theoretical foundations of the methodology and one testing their basic knowledge of a specific tool usually used for activities with that methodology. Each question provided three possible answers: one correct, one wrong and one partially wrong. Teachers could either chose to answer the question or not. During the processing phase the answers were classified considering different situations: two correct answers on the given methodology was classified as class III, only one correct answer on the given methodology was classified as class II, all the other situations (two wrong classes or two missing answers or a missing answer and a wrong answer) on the given methodology were classified as class I. So, class I represents the lowest level of knowledge, class III the highest and class II the medium level. The aim of the questionnaire B2 was to give an overview of how much teachers knew about the methodologies presented during the course at the beginning and how much this changed at the end of the course. This can give information both on how these methodologies are already in use, or at least known, in the world of teaching and on the changes in their basic knowledge which can be reasonably considered as an effect of the training course, thus providing a measure of the effectiveness of the training course.

Data from questionnaires B1-BL, B1-PT, B2-BL and B2-PT were analysed. Complete data were selected as not all the teachers correctly filled all the questionnaires. The total number of teachers, ranging from kindergarten to upper level secondary school, considered for this study, is 184. Data on age, gender, years of teaching service, level of education, educational background, educational level that they are working at, their teaching area were collected by means of questionnaire A and reported in Table 2. Data from B1-BL, B1-PT, B2-BL and B2-PT were divided into established classes (see the subsection Assessment Instruments for details) and tested by chi squared test to evaluate the significance of the differences (Bonferroni correction for multiple comparison was applied and significance was accepted for $p < 0.005$ for B1_BL and B1-PT, $p < 0.0167$ for B2_BL and B2-PT). To compare BL and PT values the McNemar test was applied; the significance was accepted for $p < 0.05$. Results from this procedure are reported in the following subsections divided by methodologies.
Figure 38 summarizes the outcomes of the course evaluated through questionnaires B1 and B2 at baseline (BL) and post-training (PT). The figure shows the significance of the differences between each column in each histogram (B1-BL, B1-PT, B2-BL, B2-PT), between the columns in B1-BL and B1-PT and between B2-BL and B2-PT. At baseline, all columns in B1 are significantly different from each other, showing a descending trend from class I, which is the most populated class, to class V. In post training, this trend shapes differently, showing a significant difference between each column, even if no significant differences were envisaged comparing columns in class I and class V and columns in class II and IV. All columns in B2-BL and B2-PT are significantly different from each other. Significant differences can be found in comparing B1-BL and B1-PT

considering columns for class I, III, IV and V. B2-BL and B2-PT show significantly different values for class II and III.

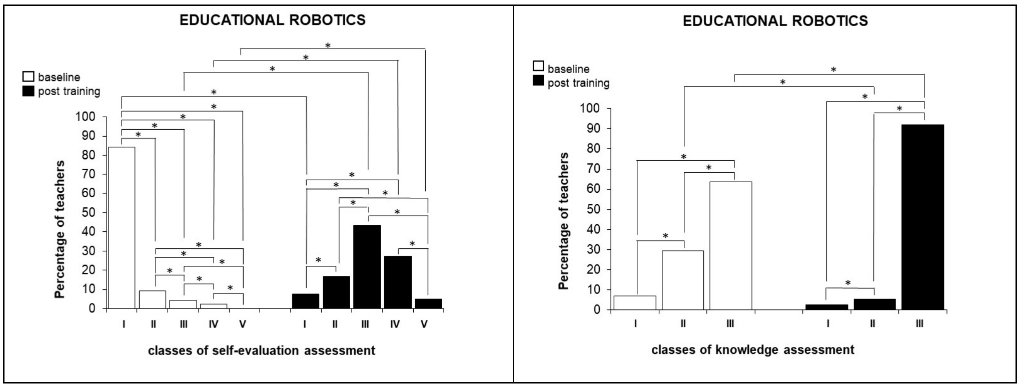


Figure 38: histograms of B1-BL and B1-PT (left), B2_BL and B2-PT (right), related to ER (B1-BL and B1-PT p-value < 0.005, B2-BL and B2-PT p-value < 0.0167 and p-value < 0.05 for comparing BL and PT values).

Figure 39 summarizes the outcomes of the course evaluated through questionnaires B1 and B2 at baseline (BL) and post-training (PT). The figure shows the significance of the differences between each column in each histogram (B1-BL, B1-PT, B2-BL, B2-PT), between the columns in B1-BL and B1-PT and between B2-BL and B2-PT. At baseline, all columns in B1 are significantly different from each other, except for class II with respect to class II and class IV to class IV; class I is significantly the most populated class of the histogram. In post training, this trend shapes differently, showing a significant difference between class III and all the other columns, between class II and class IV and between class IV and class V. All columns in B2-BL and B2-PT are significantly different from each other except for class II and class III at baseline. Significant differences can be found in comparing B1-BL and B1-PT considering columns for class I, III, IV and V. B2-BL and B2-PT show significantly different values for class II and III.

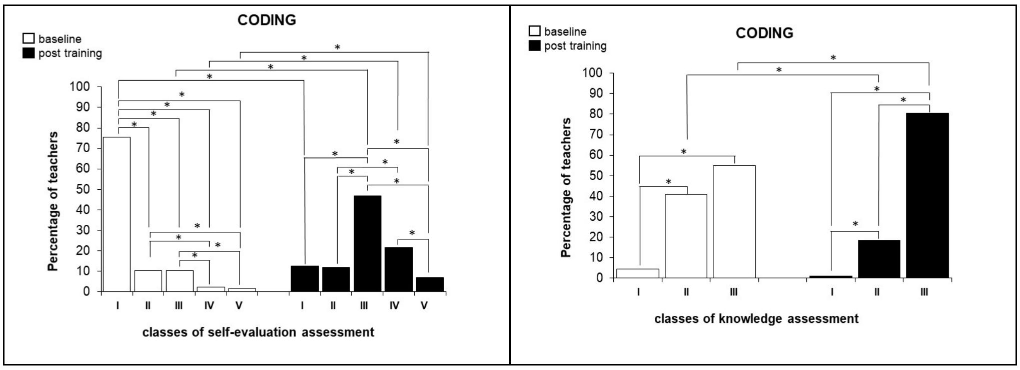


Figure 39: histograms of B1-BL and B1-PT (left), B2_BL and B2-PT (right), related to Coding (B1-BL and B1-PT p-value < 0.005, B2-BL and B2-PT p-value < 0.0167 and p-value < 0.05 for comparing BL and PT values).

Figure 40 summarizes the outcomes of the course evaluated through questionnaires B1 and B2 at baseline (BL) and post-training (PT). The figure shows the significance of the differences between each column in each histogram (B1-BL, B1-PT, B2-BL, B2-PT), between the columns in B1-BL and B1-PT and between B2-BL and B2-PT. At baseline, all columns in B1 are significantly different from each other, except for class IV and V that are also the only empty classes; class I is the class of the histogram that is significantly higher than all the others. In post training, this trend shapes differently, showing a significant difference between class V and all the other columns and between class I and class III. Columns in B2-BL show no significant difference, while B2-PT reports that class II and class III are significantly higher than class I. B1-BL and B1-PT reports significant difference between the columns of the histograms, and so do B2-BL and B2-PT.

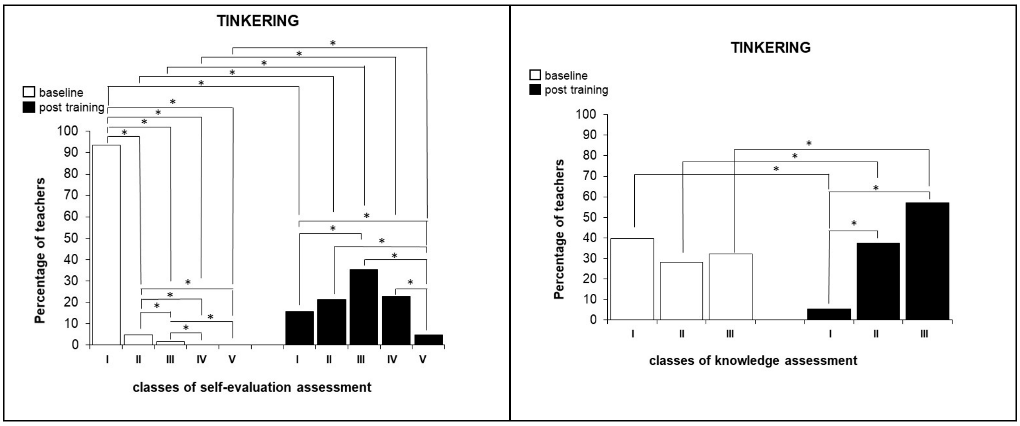


Figure 40: histograms of B1-BL and B1-PT (left), B2_BL and B2-PT (right), related to Tinkering (B1-BL and B1-PT p-value < 0.005, B2-BL and B2-PT p-value < 0.0167 and p-value < 0.05 for comparing BL and PT values).

By means of a chi squared test for association the relation between these changes in self-confidence and demographic variables collected by means of the survey A was tested (Figure 41). Significant results (p-value<0.05) can be found in all the three areas with Qualification (Level of Education) and in Educational Robotics and Tinkering in relation to Grade (teaching educational stage). Surprisingly, the primary school teachers improve more than the other category of teachers. Improvements with coding self-confidence is significantly associated with previous experience with robots. Previous studies seem to support this result because through robotics also Computational Thinking is developed (Eguchi, 2016), one of the fundamental abilities to master ability of coding. Moreover, being teachers, previous experience in the educational application with robots means that it is likely that those teachers dealt somehow with coding to program the robot and thought about how to introduce them into classroom. However, this last though is not supported by significant association of change in self-confidence in Educational Robotics with previous experience with robots.

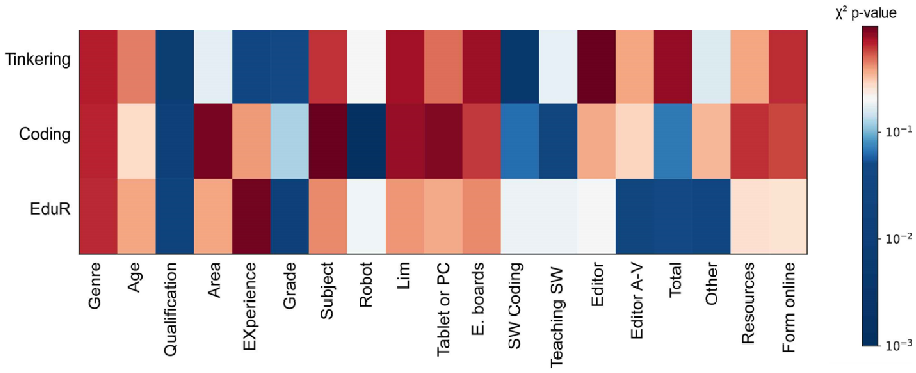


Figure 41: scores on the three methodologies compared to background variables.

Relation between changes in knowledge and the demographic variables were tested by means of a chi square test of association and results are represented in Figure 42. Even in this case coding and tinkering seem to be related to the grade (teaching educational stage). **This can be the result of school innovation policies that targeted primary school teachers more than others, thus resulting in higher degree of knowledge in the field, or maybe primary school teachers felt more motivated to learn new practical methodologies than other teachers who might have seen these methodologies more suitable for engaging primary school students in practical activities than for secondary school students.**

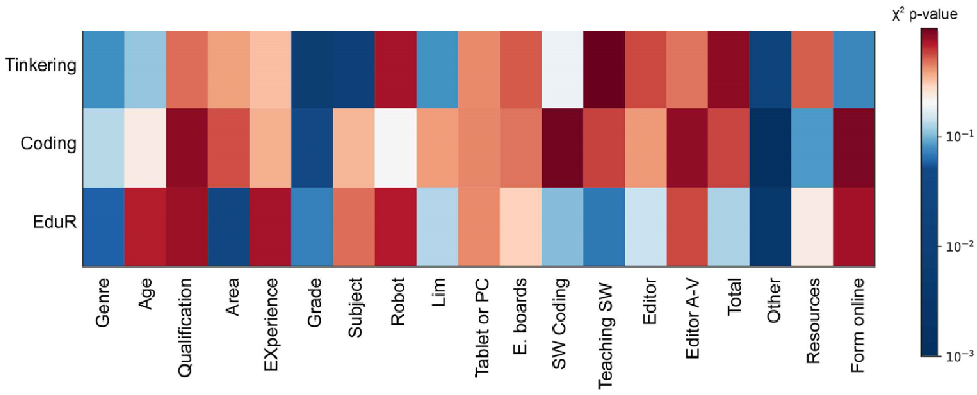


Figure 42: changes in knowledge and the demographic variables.

Regarding the relation between self-confidence and knowledge correlation analysis was performed and while in the pre-test condition the relation between the two variables seems to be linear, in the post test condition this result is not occurring (Figure 43).

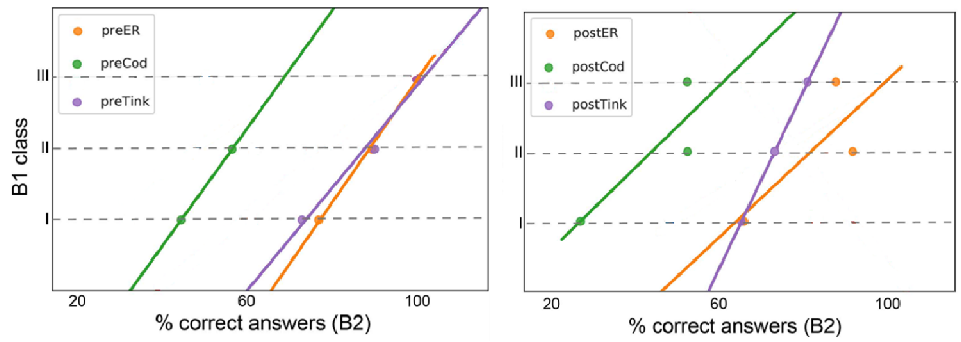


Figure 43: correlation between self-confidence and knowledge.

Analysing histograms reported in the previous section (Figure 38, Figure 39, Figure 40) some considerations can be made. First of all, similarities between histograms in all the three methodologies can be found, describing similar trends in almost all the four questionnaires. At baseline, for Educational Robotics (ER), Coding and Tinkering, we can observe that class I for B1 is significantly higher than all the columns. This suggest that the entry level for all the three methodologies was very low, reporting that teachers' attitude was very negative. Perhaps this relates to the fact that they hadn't a lot of previous experience with technologies connected with these three methodologies. On the other hand, B2-BL suggests that most of the teachers had some early general notions on ER and Coding. Nothing can be said about this on Tinkering because here all the three classes are equally represented.

At the end of course, class I for B1 is significantly lower than the corresponding class at baseline. This suggest that teachers' attitude tend to be more positive. In fact, Class III, IV and V of B1-PT are significantly higher than the corresponding classes in B1-BL, even though class V is significantly lower than class III and IV. This might be connected with the need for further exploration on the three methodologies, which could be met by adding more dedicated hours in future courses.

Anyway, B2-PT shows an increase in class III and a decrease in class II with respect to baseline for ER and Coding. The increase of class III at PT respect to BL for ER and Coding is marked, even if it should be taken into account that this class was already higher than the others at BL. At baseline

teachers seemed to have a partial knowledge on ER and Coding (Class II contains all the teachers who answered with only one correct answer to B2), at post training they seemed to improve their level of knowledge. For Tinkering we can observe a different situation at PT. Class I significantly decreases in B2-PT, while class II-III significantly increase. This can lead to think that the course effectively gave some notions on the three methodologies as we can see an increase in the classes where only one correct answer was given.
From this observation we can conclude that the course was effective and its impact on teachers' attitudes determined an increase of the positive attitude in each of the three methodologies. teachers seemed to have some previous knowledge on ER and Coding, more confused is the situation at baseline for Tinkering. This seems to suggest that teachers had previously heard about ER and Coding, while Tinkering still is not so widespread as the other two methodologies.
In the overall evaluation answers to B2 seems to be sufficiently positive from the beginning. This may relate to the fact that just two general questions are not sufficient to evaluate completely the general knowledge of a teacher on the methodology. It may be useful to empower this test with other questions or to compare it with the results from another kind of assessment instrument (for example, a design test evaluated properly). A meaningful addition to the present study could be a follow-up study to analyse if teachers had been successful in introducing the methodologies in their class and measuring the outcomes of the classrooms' activities.

3.4 Case study: Social skills

This work is part of Screpanti, Cesaretti, Storti and Scaradozzi (2019) and it focus on a way to represent the classroom's structure to provide teachers with another tool to decide on what kind of intervention should be made to make education more inclusive and learning environment more supportive. In fact, students' ability to learn is inextricably linked to the classroom environment. Developing and mastering new skills requires that students feel safe and supported. Educational Robotics (ER) brings into classroom new tools and methodologies to acquire technological, social and multidisciplinary skills and competences (Benitti,2012; Mubin et al., 2013;

Cesaretti et al., 2017; Kandlhofer and Steinbauer, 2016; Scaradozzi, Screpanti and Cesaretti, 2019; Scaradozzi et al., 2019).

During ER activities students usually work in team, being thus stimulated to collabo-rate in a small group. In this situation, students who are isolated can improve their social relationships. Sociometric tests, developed by Moreno (Scaradozzi et al., 2019; Screpanti et al., 2016, 2018a), are tools that can help teachers to examine the structure and interactions of a group, getting valuable data about the class's social relationships. Sociometric tests can help focus teacher's awareness of students who may not feel connected and need extra attention helping the creation of a supportive learning environment.

Sociometry has been applied to educational research in many contexts. Exploring this approach within ER activities is rather uncommon, but still, one example can be found Truglio, Marocco, Miglino, Ponticorvo and Rubinacci (2018).

The aim of the present work is to explore the suitability of this approach to understand classroom interactions. The Research question is: "Is ER effective in increasing connections in number and quality?" How can we understand classroom's structure and interactions between students in order to become aware of students who may not feel connected and need extra attention to create a supportive learning environment?

During the first months of 2018, two classes of an Italian primary school (ISCED 1, grade 4, age 9 years old) were involved in a path of educational robotics activities carried out in the classroom with an external educator supported by the usual teachers of those classes. Consent to participate was provided for 26 students from group A, and 22 students group B. Students were involved in activities connected with environmental issues such as differentiated waste collection and preservation of the forests. First, they explored the environmental issues; then, they wrote stories about it and brainstormed to choose elements and ideas to create a classroom's tale. After this preliminary phase students were involved in activities of robot assembly and robot programming during 4 lessons (2 hours/lesson) with Scratch and Lego WeDo 1. They were divided into groups of 3 or 4 elements and were introduced to Robotics (mechanical structure of a robot, difference between robot and machine, sensors and motors) and programming (sequential instructions only). At the end of the short course, they were able to think about how to dramatize the story about the environmental issue using robots. The third phase of the project was dedicated to the realization of the backstage scenario for the story using tinkering with waste materials. The last phase of the project was dedicated to sharing the artefacts, the story and their work with parents and other friends. The occasion was a school party, and the two groups showcased their experience with the representation of their story, short videos and pictures from laboratories. The Figure 44 sums up the four phases of the procedure.

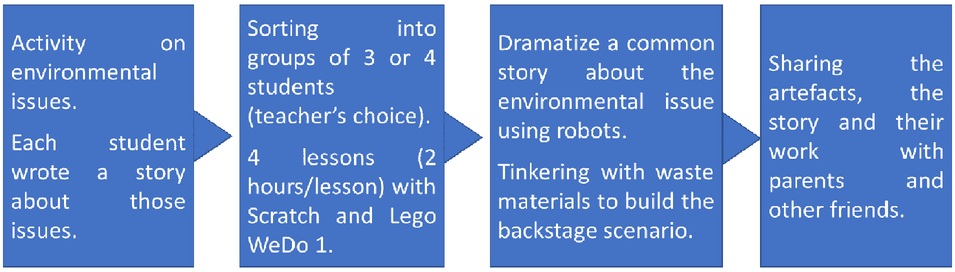


Figure 44: procedure in the classroom.

Students were asked to fill a questionnaire and a sociometric test both at the beginning (BL) of the experience and at the end of the project (PT). The items of the questionnaire were administered on paper and formulated in Italian: 3 items asked about the relationship with the instructor and methodologies, 2 questions on robot assembly and software programming, 2 on interest in this kind of activity in the future, 2 on having fun or cooperated with companions.
The sociometric test was built by four questions asking:

- Write the names and surnames of those classmates whom you would like as companion to study a school subject. You can write as many names as you like.
- Write the names and surnames of those classmates whom you would not want as companion to study a school subject. You can write as many names as you like.
- Write the names and surnames of those classmates whom you would like as companion to play with. You can write as many names as you like.
- Write the names and surnames of those classmates whom you would not want as companion to play with. You can write as many names as you like.

The sociometric test is a method to get a description of interpersonal relations within a group and to stress the social status of each member. It mainly focuses on affective-relational perspective (play) and on perspective related to group organization which aims at achieving a common purpose (study). The resulting sociogram represents members of the group as nodes and their relation as edges. If there is no relation between two nodes, no edge is shown in the graph. To examine the sociometric test data, a matrix (Adjacency matrix) is built by replacing it elements with choices (value 1), rejections (value -1) or no choice nor rejection (value 0).
Both the questionnaire and the sociometric test were developed by an expert educator and psychologist. To check whether the structure of the social network change over time, we analysed BL and PT sociometric data focusing on the density, mean indegree and mean outdegree of the network. The density of the network is the ratio of the number of links to the number of possible links; it represents the connectedness of the network. Density is connected to distance, and the distance can represent the role of the student in the network. Mean indegree and mean outdegree are measures of degree centrality appropriate for directed networks like the ones we obtained from the sociometric test. The sociometric test evaluated the students under 2 areas: play and study. For each area students had the possibility to jot down as many names as they wanted as 'choices' (people they wanted to play or study with) or as 'rejections' (people they did not want to play or study with). Paired t-tests of mean indegree and mean outdegree at BL and PT were performed in order to test the significance of the differences between the two paired samples. Change across BL and PT was assessed by means of the Wilcoxon signed rank test.
Mean values and results from the Wilcoxon signed rank test for two paired groups of data for both group A and B are shown in Table 12. In particular, results from question 4 (fun with classmates) and 5 (good collaboration with my classmates) are shown in Figure 45. Results from the processing of sociometric data are reported in Table 13. A significant change in interest towards robotics laboratory was found (Question 6) for both groups involved. No significant results were found for Questions 4 and 5, but we can qualitatively observe that scores tend to be lower for PT. This negative trend seems to find correspondence in the sociometric analysis.
Results from Table 13 shows that somehow, the project carried out increased the choices that each student received in the area "play". No significant results were found in the area "study" for both groups. The

density of the network was not tested for significance, but in both groups, it increased from BL to PT both in choices and in rejections.

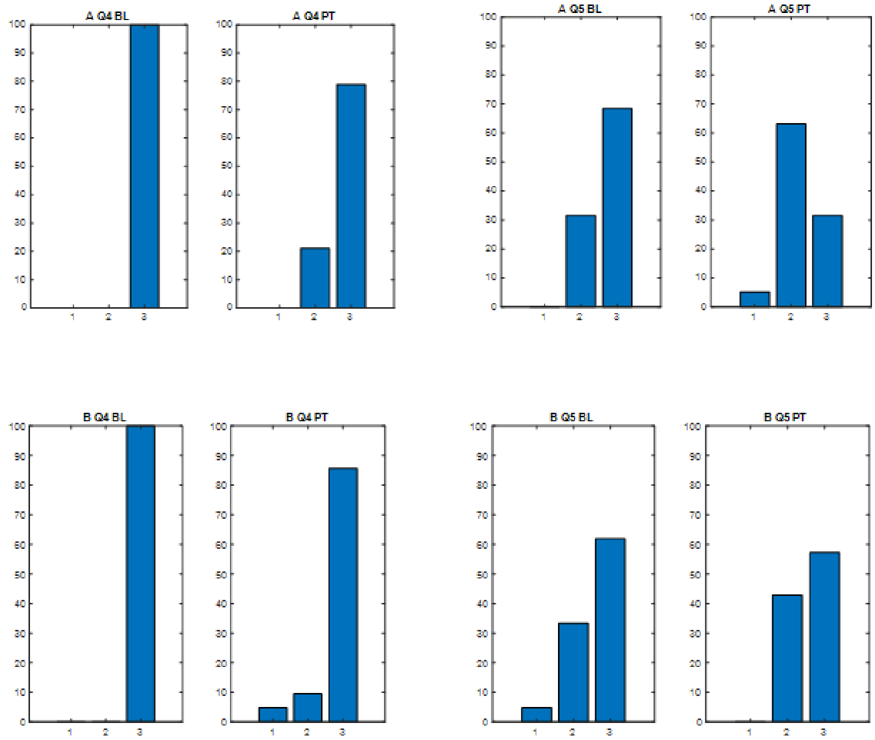


Figure 45: results from question 4 (fun with classmates) and 5 (good collaboration with my classmates).

Table 12. mean values of answers to questionnaire items. * points at the paired data showing a difference which is statistically significant (α=0.05).

	A		B	
	BL	PT	BL	PT
Q1	2.14	2.44	2.55	2.5
Q2	2.04*	2.52*	2.18	2.33
Q3	2.90	2.92	2.95	2.95
Q4	3	2.8	3	2.81
Q5	2.71*	2.32*	2.59	2.59
Q6	2.33*	2.8*	2.18*	2.91*

Table 13. Results from the analysis of sociometric data. * different points at the paired data showing a difference which is statistically significant (α=0.05).

				Mean indegree	Mean Outdegree	DENSITY
Group A	Play	Choices	BL	7.45*	7.45	0.35
			PT	8.82*	8.82	0.42
		rejections	BL	5.59	5.59	0.27
			PT	6	6	0.29
	Study	choices	BL	6.27	6.27	0.30
			PT	6.50	6.50	0.31
		rejections	BL	5.27	5.27	0.25
			PT	5.55	5.55	0.26
Group B	Play	choices	BL	4.59*	4.59*	0.21
			PT	6.45*	6.45*	0.31
		rejections	BL	4.95*	4.95*	0.24
			PT	6.5*	6.5*	0.31
	Study	choices	BL	4.82	4.82	0.23
			PT	5.32	5.32	0.25
		rejections	BL	5.18	5.18	0.25
			PT	6.23	6.23	0.30

In Figure 46 and Figure 47the sociogram are represented through a directed graph. The number of links connecting each node increases from BL to PT in both choices and rejections for both play and study. Study seem to be less affected by the intervention in the classroom. Maybe this can be connected with the playful approach to apparently extra-curricular subjects. Future development of this work would be to bring into classroom the same measurement but focusing on a subject's educational unit explored by means of ER. Moreover, a control group (a classroom not involved in the ER activities) could be introduced to cope with the maturation threat to the internal validity of the study.

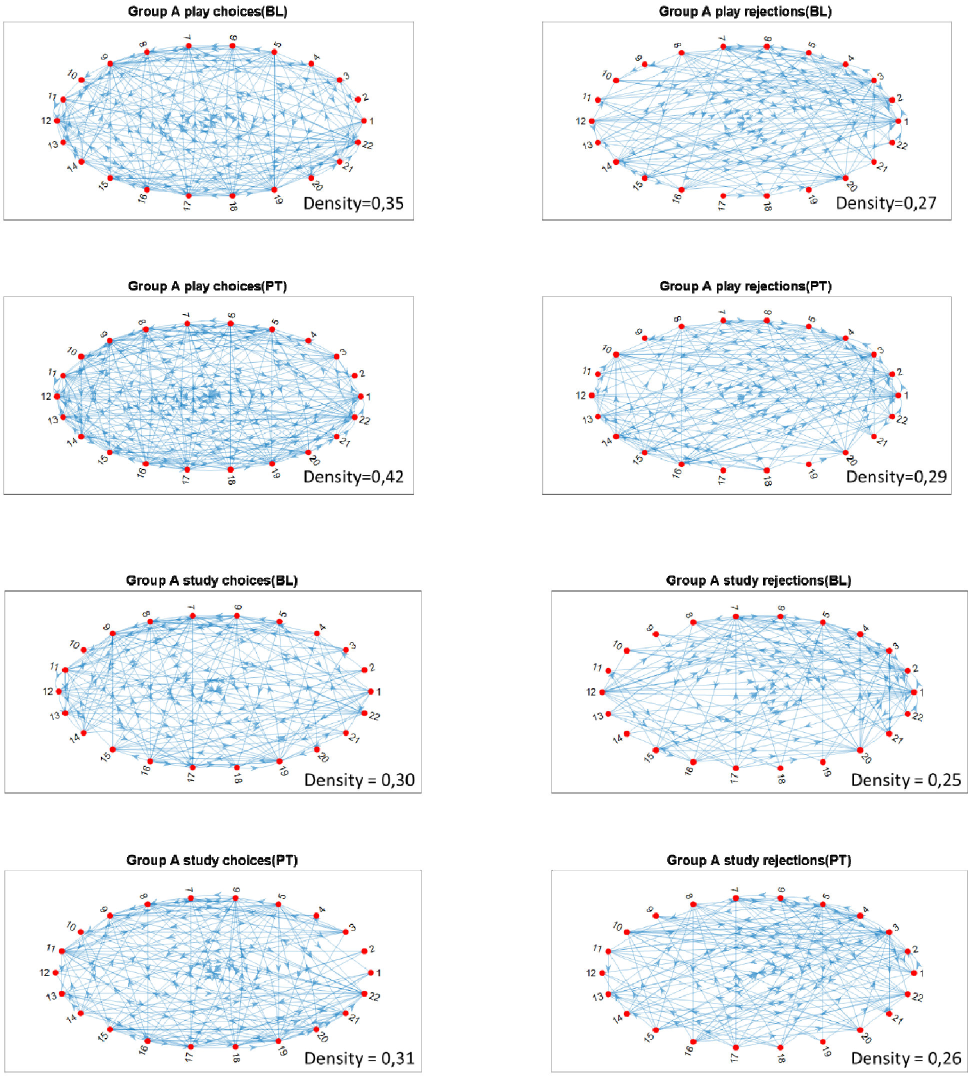


Figure 46: Group A choices (left) and rejection (right) for the area of play (top) and study (bottom).

Figure 47: Group B choices (left) and rejection (right) for the area of play (top) and study (bottom).

Overall the project reported high satisfaction in the groups involved (Table 12). Notably, the interest of the students involved in the project increased significantly (Question 6). This result seems to account for the effectiveness of short laboratories activities in increasing interest toward STEM subjects.

Answers to Questions 4 and 5 showed a decreasing or stable trend between BL and PT. On the contrary, sociometric data in the area of play showed a significant increase in the mean number of choices. Moreover, even if not significant, the same trend can also be observed in the area of study and rejections (Table 13). This is an exciting result that needs more investigation to uncover the underlying variables (for example, personal or group related variables).

The overall picture of these two classrooms showed that in group A there are more popular students (those receiving more choices than the others) and more rejected students (those receiving more rejections than others) than in group B. Density of the network was not tested for significance, but it showed an interesting trend: it increased from BL to PT showing that more relations grew out of the collaborative environment of the ER activities. This result is not unforeseen, as a part of relevant literature advocates ER's capacity to stimulate social skills (Benitti, 2012; Cesaretti et al., 2017; Kandlhofer and Steinbauer, 2016).

Nonetheless, the present research and other similar researches (Truglio et al., 2018) showed that more questions should be answered on methodological research ground and on the study of social relations within a group of study. Some students, in fact, did not improve their status after ER activities, as in (Truglio et al., 2018). This issue should be examined in-depth, maybe by including some background variables in the study like gender, race or socio-economic status. Furthermore, variables like the duration of the activities in the classroom or the way groups for activities are built may affect the creation of social relations and thus deserve more attention in future studies.

Conclusions

This thesis aimed at exploring a viable way of representing learning deriving from the ER activities carried out in the classroom. Through the lens of the Theory of Dynamical Systems the first step was to define the phenomenon which should be represented. For this reason, a definition of Educational Robotics is needed. Defining the borders of this discipline, the implications of bringing such a new methodology into classroom and the methodology employed to carry out an ER activity would greatly contribute to build a representation of the process of learning in an ER class. A classification for experiences was provided in Scaradozzi et al. (2019a) and Scaradozzi et al. (2019b) to interpret the current literature reporting experiences at school. Far from being exhaustive, these contributions aimed at laying the basis to build a shared vision on ER and on its learning objectives (Scaradozzi, 2015, 2019a). Anyway, even if the dialogue on the definition is open and ongoing (Alimisis, 2013; Angel-Fernandez and Vincze, 2018; Benitti, 2012; Scaradozzi et al., 2019b; Toh et al., 2016), it could take several years to agree on some common ground. At the same time besides no agreement on the definition of ER, another point should be clarified: what is learning. Even if there are some definition attempting to answer the question, medical, pedagogical, psychological and social research are still researching the full meaning of the word learning, each discipline through the lenses of its own ontological and epistemological stance. It is clear that modelling learning need to be approached by grounding its fundamentals in a sound scientific research method. In order to do that, both the paradigm of Systems Theory and of the Pedagogical Research should be satisfied. As expressed in the introduction and Section 1, the two paradigms are not at the opposite sides. Both System Theory and Pedagogical Research meet at the intersection of the two sciences: the empirical scientific method. Only recently (if compared to physics) pedagogy has revolutionised its world to find a way to scientifically observe and evaluate phenomena. This experimental approach could provide the structure environment to build a model of students' learning. This method in fact could provide both the demonstration for the "Causality principle" which rules the output provided by a system as a response to an input, and the way to solve the metrology issue, namely how to represent the variable learning, which is a complex construct and mostly unobservable. In fact, to represent and quantitatively describe learning, the key aspects of learning need to be identified, described, operationalised and benchmarked, providing no clues for misunderstanding or errors. Quantitative methods in the field of education could provide the sensors to capture information. Moreover, these sensors can be described in terms of validity and reliability, that inform on the best instrument to use and to measure a construct. So, research design can help demonstrating the "causality principle" and the research instruments can be the sensors that transduce a quantity into another quantity that could be processed and visualised by an external observer o by an embedded controller.

In order to test under experimental conditions this claims, the present thesis showed the instruments chosen to sense the variables chosen to represent learning: knowledge, skills, attitudes and competences. Each one represents a part of learning and they were suggested by Winterton et al. (2006) and European Union (2008) in the official document for recommendation on lifelong programme. These are not the only measures of learning that can be found in literature, but they were chosen for different reasons: not only because they were identified by an international institution and supranational government, but also because they represent the most direct way to learning. In fact, there are a number of studies reporting physical measures of learning (for example activation of specific parts of the brain, eye and posture control, …), but none of them are suitable for an ER activity or validated with respect to learning. Moreover, bringing into this kind of research this kind of evaluations could represent an ethical issue because of the young age of participants and the beneficence principle. Furthermore, biometric parameters could be mediated by several influencing factor, representing thus an indirect measure. What is considered a direct measure of learning is traditionally the assessment of the teacher in the classroom. He/She usually plan, design and evaluate the activity. The teacher or the expert carrying out an activity divides the activity into small step, small achievements for the students toward the greater goal. The teacher is the person in charge of observing the students in the classroom during all kind of activities, so the first and main researcher is the teacher that is always trying to adapt lessons and evaluation to integrate the feedback that he/she receives from the classroom. Trying to capture this feedback and provide it back to the teacher, the type of data to capture is the type of data that could be integrated by teacher. Tests, questionnaires and quizzes are traditionally accepted as a mean to evaluate learning at school. Also, these kinds of tools can be an indirect measure of the construct learning, because other variable could influence the measurement. Moreover, it is certain that not all learning will be represented by that measures.

Another consideration goes to the time of measure, the sampling period. Validated instruments could provide benchmarks to solve this problem. The sampling time ideally should remain constant during the sampling period, to allow a good collection of samples from the underlying "analogic" signal, enhancing thus the probability of a good representation, or learning trajectory. It should be noted that measurement of short-term learning trajectory (measures taken soon after the completion of the activity) could be really different from the long-term learning trajectory, and short-term measures are associated mainly with the behaviour of the single element of the system and as a result of the interaction of different elements, a new learning trajectory could emerge on the system as a whole (Steenbeek and van Geert, 2013).

The present thesis presented some experiments of assessment during an ER activity focusing on knowledge, competences, attitudes and social skills. These tools were tested for validity and used in four case studies. The ecological validity of the studies was very high because each case study involved students participating in a school program that took place in the school itself. On the other hand, longitudinal studies could enhance the external validity of those quasi-experiments. Not only external validity could be improved, but it could also be improved internal validity of the studies by including control groups, thus reducing the effects of the maturation threat to internal validity. The voluntary effect could be virtually undeletable from this kind of study, because ER activities at school and outside school are mostly promoted for voluntary participants. As a result, technology enthusiasts or students already motivated in the STEM field participates to ER projects at school and, even more, during the time spent outside school in non-formal activities.

These sensors to acquire information on learning from the system-student have as an input an information from the student and have to provide as an output a signal which is comparable with the reference signal provided by teacher, that is the controller and the reference generator of the system.

Providing training is thus important to “standardise” the input to the students, which do not mean that teachers will become a machine interpreting numbers and blindly applying corrections as a consequence. Teachers, like doctors, have the difficult role to understand humans’ phenomena and try to imagine what the system needs to grow stronger and in good health. Like in medicine a way to model some important systems have been found, while many others are still under an ongoing research, also in the educational science the systems approach can lead to a great revolution in the way we “cure” ourselves with education. The present effort to pave the way for a characterisation of sensors that could be applied to ER learning has the potential to eventually lead to the identification of the system.

References

Abbona, F., Del Re, G., & Monaco, G. (2008). *Complessità dinamica dei processi educativi. Aspetti teorici e pratici.* Franco Angeli.

Ackermann, E. (2001). Piaget's constructivism, Papert's constructionism: What's the difference. *Future of learning group publication, 5(3), 438.*

Alimisis, D. (2013). Educational robotics: Open questions and new challenges. *Themes in Science and Technology Education, 6(1), 63-71.*

Alimisis, D., Alimisi, R., Loukatos, D., & Zoulias, E. (2019). Introducing maker movement in educational robotics: beyond prefabricated robots and "black boxes". In Daniela, L. (ed.), *Smart Learning with Educational Robotics (pp. 93-115).* Springer, Cham.

Angel-Fernandez, J. M., & Vincze, M. (2018). Towards a formal definition of educational robotics. In *Proceedings of the austrian robotics workshop 2018.*

Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. *Computers & Education*, 113, 177-194.

Atmatzidou, S., & Demetriadis, S. (2016). Advancing students' computational thinking skills through educational robotics: A study on age and gender relevant differences. *Robotics and Autonomous Systems, 75, 661-670.*

Atmatzidou, S., Demetriadis, S., & Nika, P. (2018). How Does the Degree of Guidance Support Students' Metacognitive and Problem Solving Skills in Educational Robotics?. *Journal of Science Education and Technology, 27(1), 70-85.*

Bagattini, D., Miotti, B., & Operto, F. (in press). Educational robotics and gender perspective (in press). In *Proceedings of the FabLearn Italy 2019 conference*. Ancona, 20-22 November 2019, Italy.

Baker, R.S., Corbett, A.T., & Koedinger, K. R. (2004). Detecting Student Misuse of Intelligent Tutoring Systems. In *Proceedings of the 7th international conference on intelligent tutoring systems (pp. 531–540).*

Bandalos, D. L., & Finney, S. J. (2018). Factor analysis: Exploratory and confirmatory. In *The reviewer's guide to quantitative methods in the social sciences (pp. 98-122).* Routledge.

Bandura, A. (1986). *Social foundations of thoughts and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (2010). Self-efficacy. *The Corsini encyclopedia of psychology*, 1-3.

Bartlett, M. S. (1937). "Properties of sufficiency and statistical tests". *Proceedings of the Royal Statistical Society, Series A 160, 268–282*.

Beck, J. E., & Woolf, B. P. (2000). High-level student modeling with machine learning. In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), *Intelligent tutoring systems, Lecture Notes in Computer Science (pp. 584–593).* Berlin, Heidelberg: Springer.

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics, 3(21).*

Benitti, F. B. V. (2012). Exploring the educational potential of robotics in schools: A systematic review. *Computers & Education, 58(3), 978-988.*

Berland, M., Baker, R. S., & Blikstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning, 19(1-2), 205-220.*

Berland, M., Martin, T., Benton, T., Petrick Smith, C., & Davis, D. (2013). Using learning analytics to understand the learning pathways of novice programmers. *Journal of the Learning Sciences, 22(4), 564-599.*

Blikstein, P., Kabayadondo, Z., Martin, A., & Fields, D. (2017). An assessment instrument of technological literacies in makerspaces and FabLabs. *Journal of Engineering Education, 106(1), 149-175.*

Blikstein, P., & Worsley, M. (2016). Multimodal Learning Analytics and Education Data Mining: using computational technologies to measure complex learning tasks. *Journal of Learning Analytics, 3(2), 220-238.*

Blikstein, P., Worsley, M., Piech, C., Sahami, M., Cooper, S., & Koller, D. (2014). Programming pluralism: Using learning analytics to detect patterns in the learning of computer programming. Journal of the Learning *Sciences, 23(4), 561-599.*

Brennan, K., & Resnick, M. (2012, April). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the American Educational Research Association*, Vancouver, Canada (Vol. 1, p. 25).

Breazeal, C., Dautenhahn, K., & Kanda, T. (2016). Social Robotics. In: Siciliano B. & Khatib O. (eds) Springer *Handbook of Robotics (pp 1935-1972).* Springer Handbooks. Springer, Cham.

Brophy, S., Klein, S., Portsmore, M., & Rogers, C. (2008). Advancing engineering education in P-12 classrooms. *Journal of Engineering Education, 97(3), 369-387.*

Bunge M (2000). Systemism: The alternative to individualism and holism. *The Journal of Socio-Economics, 29, 147–157*. doi:10.1016/S1053-5357(00)00058-5

Caballero-Gonzalez, Y. A., Muñoz-Repiso, A. G. V., & García-Holgado, A. (2019, October). Learning computational thinking and social skills development in young children through problem solving with educational robotics. In *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality (pp. 19-23).* ACM.

Calvani A. & Menichetti, L. (2015). *Come fare un progetto didattico.* Carocci Faber.

Canuto, E. (2008). Educazione e complessità: un modello elementare dalla teoria dei sistemi di controllo. In Abbona, F., Del Re, G. and Monaco, G. (eds.), *Complessità dinamica dei processi educative. Aspetti teorici e pratici (pp. 264-284)*. Franco Angeli (Milano). ISBN: 9788856805543.

Breazeal, C., Dautenhahn, K., & Kanda, T. (2016). Social Robotics. In: Siciliano B. & Khatib O. (eds) Springer *Handbook of Robotics (pp 1935-1972).* Springer Handbooks. Springer, Cham.

Brophy, S., Klein, S., Portsmore, M., & Rogers, C. (2008). Advancing engineering education in P-12 classrooms. *Journal of Engineering Education, 97(3), 369-387.*

Bunge M (2000). Systemism: The alternative to individualism and holism. *The Journal of Socio-Economics, 29, 147–157.* doi:10.1016/S1053-5357(00)00058-5

Caballero-Gonzalez, Y. A., Muñoz-Repiso, A. G. V., & García-Holgado, A. (2019, October). Learning computational thinking and social skills development in young children through problem solving with educational robotics. In *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality (pp. 19-23).* ACM.

Calvani A. & Menichetti, L. (2015). *Come fare un progetto didattico.* Carocci Faber.

Canuto, E. (2008). Educazione e complessità: un modello elementare dalla teoria dei sistemi di controllo. In Abbona, F., Del Re, G. and Monaco, G. (eds.), *Complessità dinamica dei processi educative. Aspetti teorici e pratici (pp. 264-284)*. Franco Angeli (Milano). ISBN: 9788856805543.

Castro, E., Cecchi, F., Valente, M., Buselli, E., Salvini, P., & Dario, P. (2018). Can educational robotics introduce young children to robotics and how can we measure it*?. Journal of Computer Assisted Learning, 34(6), 970-977.*

Cattell, R. B. (1966). "The Scree Test for the number of factors". *Multivariate Behavioral Research. 1 (2): 245–276.* doi:10.1207/s15327906mbr0102_10.

Cesaretti, L., Screpanti, L., D. Scaradozzi, & Mangina, E. (*in press*). Analysis of Educational Robotics activities using a machine learning approach. *Proceedings of FabLearn Italy 2019*. Ancona, Italy.

Cesaretti, L., Storti, M., Mazzieri, E., Screpanti, L., Paesani, A., Principi, P., & Scaradozzi, D. (2017). An innovative approach to School-Work turnover programme with Educational Robotics. *Mondo Digitale, 16(72), 2017-5.*

Chomsky, N. (1980). Rules and representations. *The Behavioral and Brain Sciences, 3, 1-61.*

Cittadinanza e creatività digitale (2017). Avviso pubblico 2669 del 03 marzo 2017. https://www.istruzione.it/pon/avviso_cittadinanza-creativita.html Last accessed 28/11/2019.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78, 98-104.*

Cronbach LJ (1951). "Coefficient alpha and the internal structure of tests". *Psychometrika*. 16 (3): 297–334. doi:10.1007/bf02310555.

Cross, J. L., Hamner, E., Zito, L., & Nourbakhsh, I. (2017, October). Student outcomes from the evaluation of a transdisciplinary middle school robotics program. In *Proceedings of Frontiers in Education Conference (FIE) (pp. 1-9).* IEEE.

Council of Europe (2011). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Council of Europe. Available at https://www.coe.int/en/web/common-european-framework-reference-languages

Daniela, L. & Lytras, M.D. (2019). Educational Robotics for Inclusive Education. *Technology, Knowledge and Learning, 24, 219.* doi: 10.1007/s10758-018-9397-5

Daniela, L., & Strods, R. (2019). Educational Robotics for Reducing Early School Leaving from the Perspective of Sustainable Education. In *Smart Learning with Educational Robotics (pp. 43-61).* Springer, Cham.

Daniela, L., Strods, R., & Alimisis, D. (2017, July). Analysis of robotics-based learning interventions for preventing school failure and early school leaving in gender context. In *9th international conference on education and new learning technologies (Edulearn17) (pp. 810-818*).

Dann, R. (2014). Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice. 21 (2), p.149-166.*

Dede, C. (2010). Comparing frameworks for 21st century skills. *21st century skills: Rethinking how students learn, 20, 51-76.*

Di Lieto, M. C., Inguaggiato, E., Castro, E., Cecchi, F., Cioni, G., Dell'Omo, M., ... & Dario, P. (2017). Educational Robotics intervention on Executive Functions in preschool children: A pilot study. *Computers in human behavior, 71, 16-23.*

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53, 109-132.*

Eguchi, A. (2014, July). Robotics as a learning tool for educational transformation. In *Proceedings of 4th international workshop teaching robotics, teaching with robotics & 5th international conference robotics in education.* Padova (Italy).

Eguchi, A. (2015, March). Educational robotics to promote 21st century skills and technological understanding among underprivileged undergraduate students. In *2015 IEEE Integrated STEM Education Conference (pp. 76-82).* IEEE.

Eguchi, A. (2016, March). Computational Thinking with Educational Robotics. In G. Chamblee & L. Langub (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference (pp. 79-84).* Savannah, GA, United States: Association for the Advancement of Computing in Education (AACE). Retrieved from https://www.learntechlib.org/primary/p/172306/.

Eguchi A. (2017) Bringing Robotics in Classrooms. In: Khine M. (eds) *Robotics in STEM Education (pp. 3-31).* Springer, Cham

European Union (2008). Recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework for lifelong learning. *Official Journal of the European Union, C 111, 06 May 2008.* Available at https://op.europa.eu/en/publication-detail/-/publication/0e4412eb-2d36-4eb3-a1fb-11908e853058

European Union (2006). *Recommendation of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning*. Available at https://eur-lex.europa.eu/eli/reco/2006/962/oj Last accessed 29/11/2019.

Eurydice, 2012. *Developing Key Competences at School in Europe: Challenges and Opportunities for Policy. Eurydice Report*. Luxembourg: Publications Office of the European Union.

Faber, M., Walton, M., Booth, S., & Parker, B. (2013). *The golden LEAF STEM initiative evaluation. Year 2 Report.* Available at https://cerenc.org/wp-content/uploads/2011/11/FI_GLFSTEMYearTwoAppendices_May2013_FINAL.pdf [Last retrieved 29/11/2019].

Fabrigar, Leandre & Wegener, Duane. (2012). *Exploratory Factor Analysis*. Oxford University Press. doi: 10.1093/acprof:osobl/9780199734177.001.0001.

Firestone, W. A. (1987). Meaning in method: The rhetoric of quantitative and qualitative research. *Educational Researcher, 16(7), 16-21.*

Gardner, H. (1992). *Multiple intelligences*. Minnesota Center for Arts Education.

Gattie, D. K., Kellam, N. N., Schramski, J. R., & Walther, J. (2011). Engineering education as a complex system. *European Journal of Engineering Education, 36(6), 521-535*.

Goldman, R., Eguchi, A., & Sklar, E. (2004, June). Using educational robotics to engage inner-city students with technology. In *Proceedings of the 6th international conference on Learning sciences (pp. 214-221)*. International Society of the Learning Sciences.

Gorsuch, R. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, 19(2), 149-161.

Hair J., Anderson R. E., Tatham R. L., Black W.C. (1995). *Multivariate data analysis*. 4th ed. New Jersey: Prentice-Hall Inc.

Hattie, J. and Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77 (1), p.81-112*.

Hilpert, J. C. and Marchand, G. C. (2018). Complex Systems Research in Educational Psychology: Aligning Theory and Method. *Educational Psychologist, 53(3): 185–202*. doi:10.1080/00461520.2018.1469411.

Hogarty, K., Hines, C., Kromrey, J., Ferron, J. & Mumford, K. (2005). The Quality of Factor Solutions in Exploratory Factor Analysis: The Influence of Sample Size, Communality, and Overdetermination. *Educational and Psychological Measurement.65(2):202-26.*

Holland JH (2006). Studying Complex Adaptive Systems. *Journal of Systems Science and Complexity, 19, 1–8.* doi:10.1007/s11424-006-0001-z

Horn, J. L. (June 1965). A rationale and test for the number of factors in factor analysis. *Psychometrika. 30 (2): 179–185.* doi:10.1007/bf02289447

In estate si imparano le STEM (2018) https://www.monitordposcuola.it/2018/01/pubblicazione-bando-estate-si-imparano-le-stem-ii-edizione/ [Last accessed 29/11/2019]

INVALSI (2006). https://www.invalsi.it/invalsi/prog/ecas.php?page=ecas_it_01 [Last accessed 29/11/2019]

Jabarullah, N. H., & Hussain, H. I. (2019). The effectiveness of problem-based learning in technical and vocational education in Malaysia. *Education+ Training.*

Jacobson, M. J., Levin, J. A., & Kapur, M. (2019). Education as a complex system: Conceptual and methodological implications. *Educational Researcher, 48(2), 112-119.*

Jacobson, M. J. and Wilensky, U. (2006). Complex systems in education: educational importance and implications for learning. *Journal of the learning sciences 15 (1), 11-34.*

Jones, B. F., Rasmussen, C. M., & Moffitt, M. C. (1997). *Real-life problem solving.: A collaborative approach to interdisciplinary learning.* Washington, DC: American Psychological Association.

Jormanainen, I., & Sutinen, E. (2012, March). Using data mining to support teacher's intervention in a robotics class. In *Digital Game and Intelligent Toy Enhanced Learning (DIGITEL), 2012 IEEE Fourth International Conference on (pp. 39-46).* IEEE.

Kandlhofer, M., & Steinbauer, G. (2016). Evaluating the impact of educational robotics on pupils' technical-and social-skills and science related attitudes. *Robotics and Autonomous Systems, 75, 679-685.*

Kaiser, H. F. (1970). A Second-Generation Little Jiffy. *Psychometrika*, 35(4), 401-15.

Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark IV. *Educational and psychological measurement, 34(1), 111-117.*

Kaiser, H. F. (1959). Computer program for varimax rotation in factor analysis. *Educational and psychological measurement, 19(3), 413-420.*

Kapur, M., Hung, D., Jacobson, M., & Voiklis, J. (2007). *Emergence of learning in computer-supported, large-scale collective dynamics: A research agenda.* New Brunswick, NJ.

Kauffman, S. (1995). *At home in the universe: The search for laws of self-organization and complexity*. New York: Oxford University Press.

Kucuk, S., & Sisman, B. (2017). Behavioral patterns of elementary students and teachers in one-to-one robotics instruction. *Computers & Education, 111, 31-43.*

Lindh, J.¸ Holgersson, T. (2007). Does lego training stimulate pupils' ability to solve logical problems? *Computers & Education, 49 (4), pp. 1097-1111.*

López, X., Valenzuela, J., Nussbaum, M., & Tsai, C. C. (2015). Some recommendations for the reporting of quantitative studies. *Computers & Education, 91, 106-110.* doi: 10.1016/j.compedu.2015.09.010

Martinez, S. L., & Stager, G. (2013). *Invent to learn: Making, tinkering, and engineering in the classroom.* Torrance, CA: Constructing modern knowledge press.

Matarić, M.J., & Scassellati, B. (2016). Socially Assistive Robotics. In: Siciliano, B., & Khatib, O. (eds). *Springer Handbook of Robotics (pp. 1973-1994).* Springer Handbooks. Springer, Cham.

Mazzer, M. (2018). Technologies in School for an accessible, inclusive and competence-oriented education. *Education Sciences & Society-Open Access Journal, 9(1).*

Merceron, A., & Yacef, K. (2004). Mining student data captured from a web-based tutoring tool: Initial exploration and results. *Journal of Interactive Learning Research, 15(4), 319–346.*

Mevarech, Z. R. & Kramarski, B. (1993), Vygotsky and Papert: social-cognitive interactions within Logo environments. *British Journal of Educational Psychology, 63: 96-109*. doi:10.1111/j.2044-8279.1993.tb01044.x

Mikropoulos, T. A., & Bellou, I. (2013). Educational robotics as mindtools. *Themes in Science and Technology Education, 6(1), 5-14*.

MISO Project. Homepage. (2012) Available from: http://miso.ncsu.edu

Mitchell, M. (2009). *Complexity: A guided tour*. New York, NY: Oxford University Press.

MIUR – Ministero dell'Istruzione, dell'Università e della Ricerca (2012). Indicazioni nazionali per il curriculo della scuola dell'infanzia e del primo ciclo di istruzione. Annali della Pubblica Istruzione, LXXXVIII, Numero speciale 2012.

Mubin, O., Stevens, C. J., Shahid, S., Al Mahmud, A., & Dong, J. J. (2013). A review of the applicability of robots in education. *Journal of Technology in Education and Learning, 1(209-0015), 13*.

Navarro, M., Förster, C., González, C., & González-Pose, P. (2016). Attitudes toward science: Measurement and psychometric properties of the Test of Science-Related Attitudes for its use in Spanish-speaking classrooms. *International Journal of Science Education, 38(9), 1459-1482.*

Nicolescu, B. N., & Petrescu, T. C. (2013). Dynamical systems theory–a powerful tool in the educational sciences. *Procedia-Social and Behavioral Sciences, 76, 581-587.*

Ornelas, F., & Ordonez, C. (2017). Predicting Student Success: A Naïve Bayesian Application to Community College Data. *Technology, Knowledge and Learning, 22(3), 299-315.*

Papert, S. (1972), *Learn Think to Children,* UCLA (University of California Los Angeles).

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas.* New York: Basic Books, Inc..

Papert, S. (1991). Situating constructionism. In Papert S., & Harel I., (Eds.), *Constructionism (pp. 1-11).* Norwood, NJ: Ablex Publishing.

Papert, S. (1999). What is Logo? Who needs it. In *Logo Philosophy and Implementation (pp. 4-16).* Logo Computer Systems Inc.

Parunak, H. V. D., Savit, R., & Riolo, R. L. (1998). Agent-based modeling vs equation-based modeling: A case study and users' guide. *Lecture Notes in Computer Science, 1534, 10–25*. doi:10.1007/10692956_2

Pett, M. A., Lackey N.R., Sullivan J.J. (2003). *Making Sense of Factor Analysis: The use of factor analysis for instrument development in health care research*. California: Sage Publications Inc.

Piech, C., Sahami, M., Koller, D., Cooper, S., & Blikstein, P. (2012, February). Modeling how students learn to program. In *Proceedings of the 43rd ACM technical symposium on Computer Science Education (pp. 153-160)*. ACM.

Resnick, L.B. (1987*), Education and Learning to Think*. National Academy Press. Washington.

Resnick, M. (2004). Edutainment? No thanks. I prefer playful learning. *Associazione Civita Report on Edutainment, 14, 1-4.*

Scaradozzi, D., Cesaretti, L., Costa, D., Screpanti, L., & Zingaretti, S. (2019a). Innovative tools for teaching marine robotics, IoT and control strategies since the primary school. In Daniela, L. (ed.), *Smart Learning with Educational Robotics - Using Robots to Scaffold Learning Outcomes (pp 199-227).* Springer.

Scaradozzi, D., Cesaretti, L., Screpanti, L., & Mangina, E. (*in press*). Identification of the students learning process during Education Robotics activities. *Frontiers in Robotics and AI.*

Scaradozzi, D., Screpanti, L., & Cesaretti, L. (2019b). Towards a definition of educational robotics: a classification of tools, experiences and assessments. In *Smart Learning with Educational Robotics - Using Robots to Scaffold Learning Outcomes (pp. 63-92).* Springer, Cham.

Scaradozzi, D., Screpanti, L., Cesaretti, L., Mazzieri, E., Storti, M., Brandoni, M., & Longhi, A. (2016) "Rethink loreto: we build our smart city!" A stem education experience for introducing smart city concept with the educational robotics. In *Proceedings of the 9th annual International Conference of Education, Research and Innovation (ICERI2016), pp. 750-758*. 14-16 November 2016, Seville, Spain.

Scaradozzi, D., Screpanti, L., Cesaretti, L., Storti, M., & Mazzieri, E. (2019). Implementation and assessment methodologies of teachers' training courses for STEM activities. *Technology, Knowledge and Learning, 24(2), 247-268*. doi: 10.1007/s10758-018-9356-1

Scaradozzi, D., Sorbi, L., Pedale, A., Valzano, M., & Vergine, C. (2015). Teaching robotics at the primary school: an innovative approach. *Procedia-Social and Behavioral Sciences, 174*, 3838-3846.

Scaradozzi, D., Diamantini, C., Screpanti, L., Cesaretti, L., Mazzieri, E., & Costa, D. (2018). *Italian patent no. 10201800000963*6, Apparato di monitoraggio di operazioni di assemblaggio e programmazione. Retrieved from https://www.univpm.it/Entra/Engine/RAServeFile.php/f/Innovazione_tecnologica/brevetti/Smart_Blocks_ENG.pdf [Last accessed 29/11/2019].

Schank, R. (1982). *Dynamic Memory: A Theory of Learning in Computers and People*. New York: Cambridge University Press.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: a review. *Journal of Educational Research, 99(6), 323-338.*

Screpanti, L., Cesaretti, L., Marchetti, L., Baione, A., Natalucci, I. N., & Scaradozzi, D. (2018a, July). An Educational Robotics activity to promote gender equality in STEM education. In *Procedings of the* *eighteenth International Conference on Information, Communication Technologies in Education (ICICTE 2018) (pp. 336-346).* Chania, Crete, Greece.

Screpanti, L., Cesaretti, L., Storti, M., Mazzieri, E., Longhi, A., Brandoni, M. & Scaradozzi, D. (2018b). Advancing K12 education through Educational Robotics to shape the citizens of the future. In *Proceedings of the DIDAttica e inforMATICA conference (DidaMatica2018)*, 19-20 April 2018, Cesena, Italy. AICA.

Screpanti, L., Cesaretti, L., Storti, M., and Scaradozzi, D. (in press). Educational Robotics and social relationships in the classroom. In *Proceedings of Fablearn Italy 2019*. 20-22 November 2019, Ancona, Italy.

Steenbeek, H., & van Geert, P. (2013). The emergence of learning-teaching trajectories in education: A complex dynamic systems approach. *Nonlinear dynamics, psychology, and life sciences, 17(2), 233-267*.

Streiner, D. L. (2003). Starting at the beginning: an introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment, 80(1), 99-103.*

Strogatz SH (1994). *Nonlinear systems and chaos*. Reading, MA:Perseus.

Suthers, D. D., & Hundhausen, C. (2003). An empirical study of the effects of representational guidance on collaborative learning. *Journal of the Learning Sciences, 12(2), 183–219*.

Tabachnick, B. G., Fidell L.S. (2007). *Using Multivariate Statistics*. Boston: Pearson Education Inc.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. American Psychological Association.

Toh, L. P. E., Causo, A., Tzuo, P. W., Chen, I. M., & Yeo, S. H. (2016). A review on the use of robots in education and young children. *Journal of Educational Technology & Society, 19(2), 148-163.*

Truglio, F., Marocco, D., Miglino, O., Ponticorvo, M., & Rubinacci, F. (2018, April). Edu-cational Robotics to Support Social Relations at School. In *International Conference on Ro-botics and Education RiE 2017 (pp. 168-174).* Springer, Cham.

Turkle, S., & Papert, S. (1992). Epistemological pluralism and the revaluation of the concrete. *Journal of Mathematical Behavior, 11(1), 3-33.*

Twining, P., Heller, R. S., Nussbaum, M., & Tsai, C. (2017). *Some guidance on conducting and reporting qualitative studies. Computers & Education, 106, A1-A9.* doi: 10.1016/j.compedu.2016.12.002

Unfried, A., Faber, M., Stanhope, D. S., & Wiebe, E. (2015). The development and validation of a measure of student attitudes toward science, technology, engineering, and math (S-STEM). *Journal of Psychoeducational Assessment, 33(7), 622-639.*

Unlu, Z. K., Dokme, I., & Veli, U. N. L. U. (2016). Adaptation of the science, technology, engineering, and mathematics career interest survey (STEM-CIS) into Turkish. *Eurasian Journal of Educational Research, 16(63).*

Voogt, J., & Roblin, N. P. (2012). A comparative analysis of international frameworks for 21st century competences: Implications for national curriculum policies. *Journal of curriculum studies, 44(3), 299-321.*

Van Laar, E., van Deursen, A. J., van Dijk, J. A., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in human behavior, 72, 577-588.*

Vygotsky, L. S. (1968). *Thought and language*. [Mass]: M.I.T. Press.

Weinberg, J. B., Pettibone, J. C., Thomas, S. L., Stephen, M. L., & Stein, C. (2007, June). The impact of robot projects on girls' attitudes toward science and engineering. In *Workshop on Research in Robots for Education (Vol. 3, pp. 1-5).*

White, I., Smigiel, H., & Levin, J. (2017). Modelling the complexity of technology adoption in higher education teaching practice. In *proceedings of The Higher Education Technology Agenda (THETA 2017) Conference.* Auckland, New Zealand. Retrieved at https://dspace.flinders.edu.au/xmlui/handle/2328/37345 [Last accessed 29/11/2019]

Wiebe, E. N., & Faber, M., & Corn, J., & Collins, T. L., & Unfried, A., & Townsend, L. (2013, June), *A Large-scale Survey of K-12 Students about STEM: Implications for Engineering Curriculum Development and Outreach Efforts (Research to Practice)* Paper presented at 2013 ASEE Annual Conference & Exposition, Atlanta, Georgia. https://peer.asee.org/19073

Wiebe, E., Unfried, A., & Faber, M. (2018). The relationship of STEM attitudes and career interest. EURASIA *Journal of Mathematics, Science and Technology Education, 14(10).*

Wiener, N., (1948). *Cybernetics: or control and communication in the animal and the machine*. Technology press, Boston.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25(1), 68-68.*

Wilcoxon F. (1945) Individual comparisons by ranking methods. *Biometrics 1: 80-83.*

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine, 8(3).*

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49(3), 33-35.*

Wing, J. M. (2008). Computational thinking and thinking about computing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 366(1881), 3717-3725.*

Winterton, J., Delamare, F., Le Deist & Stringfellow, E. (2006). Typology of Knowledge, Skills, and Competences: Clarification of the Concept and Prototype. Cedefop Reference series, 64. Luxembourg: Office for Official Publications of the European Communities. Retrieved from https://www.cedefop.europa.eu/files/3048_en.pdf

Appendix A

Crossword puzzle for primary school from Scaradozzi et al., 2016 and Screpanti et al., 2018

CRUCISMART

Prova a risolvere questo cruciverba utilizzando le parole che hai incontrato durante questo percorso di Robotica e Smart City!

Orizzontali

4 Una città innovativa, ecologica e accessibile...insomma, intelligente!
5 È l'elemento naturale che fa muovere la pala eolica
6 Quello che serve quando da solo non riesci a risolvere un problema
9 Una fonte di energia che non si esaurisce mai
12 Il blocco su Scratch che fa andare il robot in un verso o nell'altro
15 L'energia che abbiamo utilizzato per far muovere i robot
16 Il "cervello" del robot
18 Il ruolo nel gruppo nel quale cerchi e gestisci i pezzi del kit

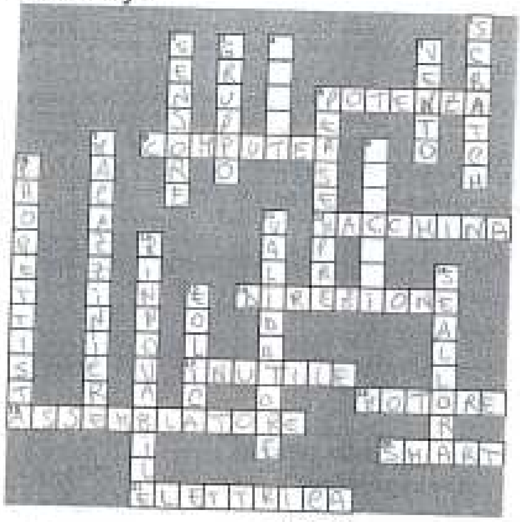
Verticali

1 Generano il movimento del robot
2 I sensi del robot
3 Un insieme di meccanismi che svolge un lavoro ma privo di intelligenza
7 Una macchina che non serve a niente
8 Il blocco su Scratch che permette di ripetere delle istruzioni all'infinito
10 Il ruolo nel gruppo in cui usi le mani per costruire
11 Il blocco su Scratch che fa muovere più o meno veloce il robot
13 I "motori" nell'uomo
14 L'energia che possiamo simulare soffiando
17 Macchina autonoma programmabile e intelligente

Crossword puzzle for lower secondary school from Scaradozzi et al., 2016 and Screpanti et al., 2018

CRUCISMART

Prova a risolvere questo cruciverba utilizzando le parole che hai incontrato durante questo percorso di Robotica e Smart City!



Orizzontali

6 Il blocco su Scratch che fa muovere più o meno veloce il robot
8 Il "cervello" del robot
12 Un insieme di meccanismi che svolge un lavoro ma privo di intelligenza
16 Il blocco su Scratch che fa andare il robot in un verso o nell'altro
17 Una macchina che non serve a niente
18 Generano il movimento del robot
19 Il ruolo nel gruppo in cui usi le mani per costruire
20 Una città innovativa, ecologica e accessibile...insomma, intelligente!
21 L'energia che abbiamo utilizzato per far muovere i robot

Verticali

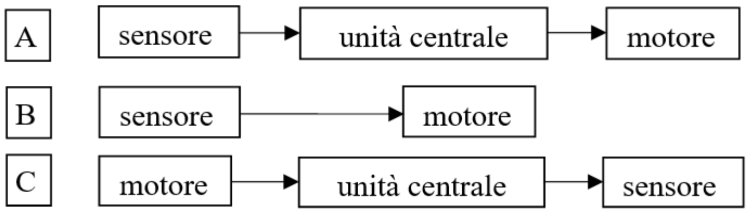
1 Il programma che serve a costruire le sequenze per far muovere il robot
2 I sensi del robot
3 Quello che serve quando da solo non riesci a risolvere un problema
4 Macchina autonoma programmabile e intelligente.
5 È l'elemento naturale che fa muovere la pala eolica
6 Il blocco su Scratch che permette di ripetere delle istruzioni all'infinito
7 Il ruolo nel gruppo nel quale cerchi e gestisci i pezzi del kit
9 I "motori" nell'uomo
10 Il ruolo nel gruppo nel quale osservi e spieghi le istruzioni per costruire il robot
11 Il ruolo nel gruppo nel quale devi tenere gli occhi aperti
13 Una fonte di energia che non si esaurisce mai
14 Il blocco su Scratch che si usa per accendere il motore del robot quando la distanza misurata dal sensore è < 10
15 L'energia che possiamo simulare soffiando

Appendix B

Knowledge questionnaire for students of the Lower Secondary School in the pre-activities phase

1) Come possiamo usare il sensore a Ultrasuoni?
   a. per misurare la distanza tra il robot e l'ostacolo.
   b. per misurare i suoni dentro la stanza.
   c. per azionare gli altoparlanti del robot.

2) Quale di queste configurazioni consente al robot di controllare il suo movimento di fronte a un ostacolo?



3) Associa le parti di un robot (riportate nella lista a sinistra) con le parti del corpo umano (riportate nella lista a destra). Associa le parti che svolgono una funzione simile nel robot e nel corpo umano.

Parti di un robot	**Parti del corpo umano**
A. Motori	□ Ossa/scheletro
B. Videocamera	□ Bocca/corde vocali
C. Microfono	□ Cuore
D. Unità centrale di elaborazione delle informazioni	□ Muscoli

E. Batteria

F. Struttura meccanica

G. Altoparlante

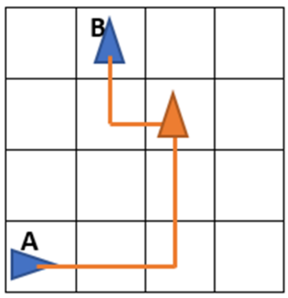
☐ Udito

☐ Vista

☐ Cervello

4) Metti in ordine le seguenti istruzioni in modo tale che la freccia può arrivare al punto B partendo dal punto A (inserisci i numeri vicino la corrispondente istruzione).

	Avanti di 1 posizione
	Avanti di 1 posizione
	Avanti di 2 posizioni
	Avanti di 2 posizioni
	Gira a sinistra di 90°
	Gira a sinistra di 90°
	Gira a destra di 90°
	fine
	inizio

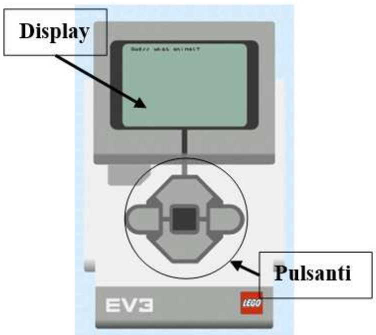


5) Un ambiente di programmazione visuale contiene:
- ☐ comandi per i robot rappresentati da solo testo
- ☐ comandi per i robot rappresentati da dei blocchi colorati
- ☐ comandi per i robot rappresentati solo da frecce e pulsanti

6) Che cos'è un motore?
   - ☐ È un componente del robot responsabile di acquisire dati
   - ☐ È un componente del robot responsabile del movimento del robot
   - ☐ È un componente del robot responsabile dell'elaborazione dei dati

7) Voglio che il robot non tocchi il muro, quindi programmo il robot affinché
   - ☐ i motori si spengano quando la distanza rilevata dal sensore è uguale a 30 cm
   - ☐ i motori si spengano quando la distanza rilevata dal sensore è maggiore di 30 cm
   - ☐ i motori si spengano quando la distanza rilevata dal sensore è minore di 30 cm

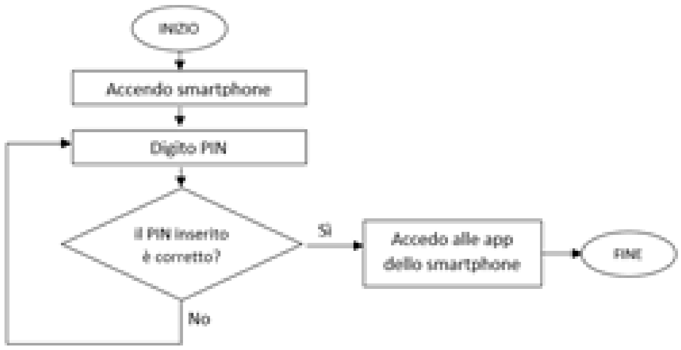
8) Individua gli input/ingressi (IN) e output/uscite (OUT) sul mattoncino intelligente



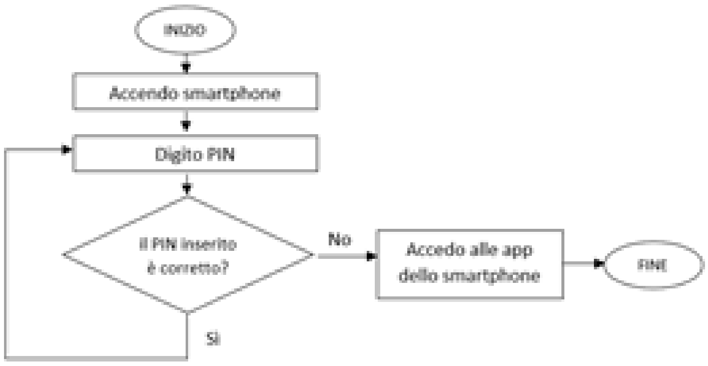
	Ingresso/Input	Output/Uscita
Motore a elevata potenza	☐	☐
Sensore a Ultrasuoni	☐	☐
Sensore di luce	☐	☐
Pulsanti	☐	☐
Display	☐	☐
Sensore di colore	☐	☐
Motore a potenza media	☐	☐
Sensore di contatto	☐	☐
Giroscopio	☐	☐

9) Voglio usare lo smartphone, ma devo inserire il PIN di sblocco. Quale tra i seguenti è il diagramma di flusso che meglio rappresenta le azioni che devo compiere?

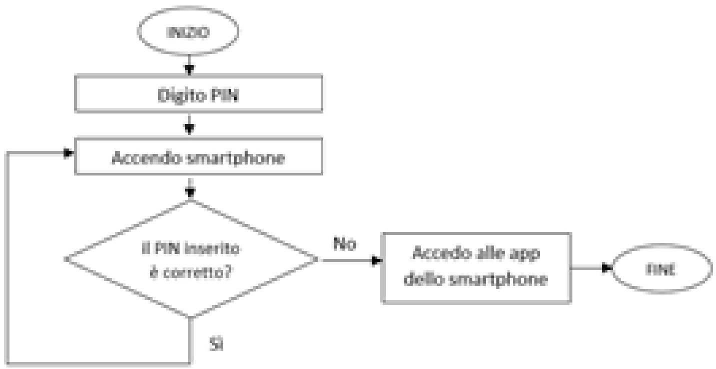
A


INIZIO
Accendo smartphone
Digito PIN
Il PIN inserito è corretto?
Sì
No
Accedo alle app dello smartphone
FINE

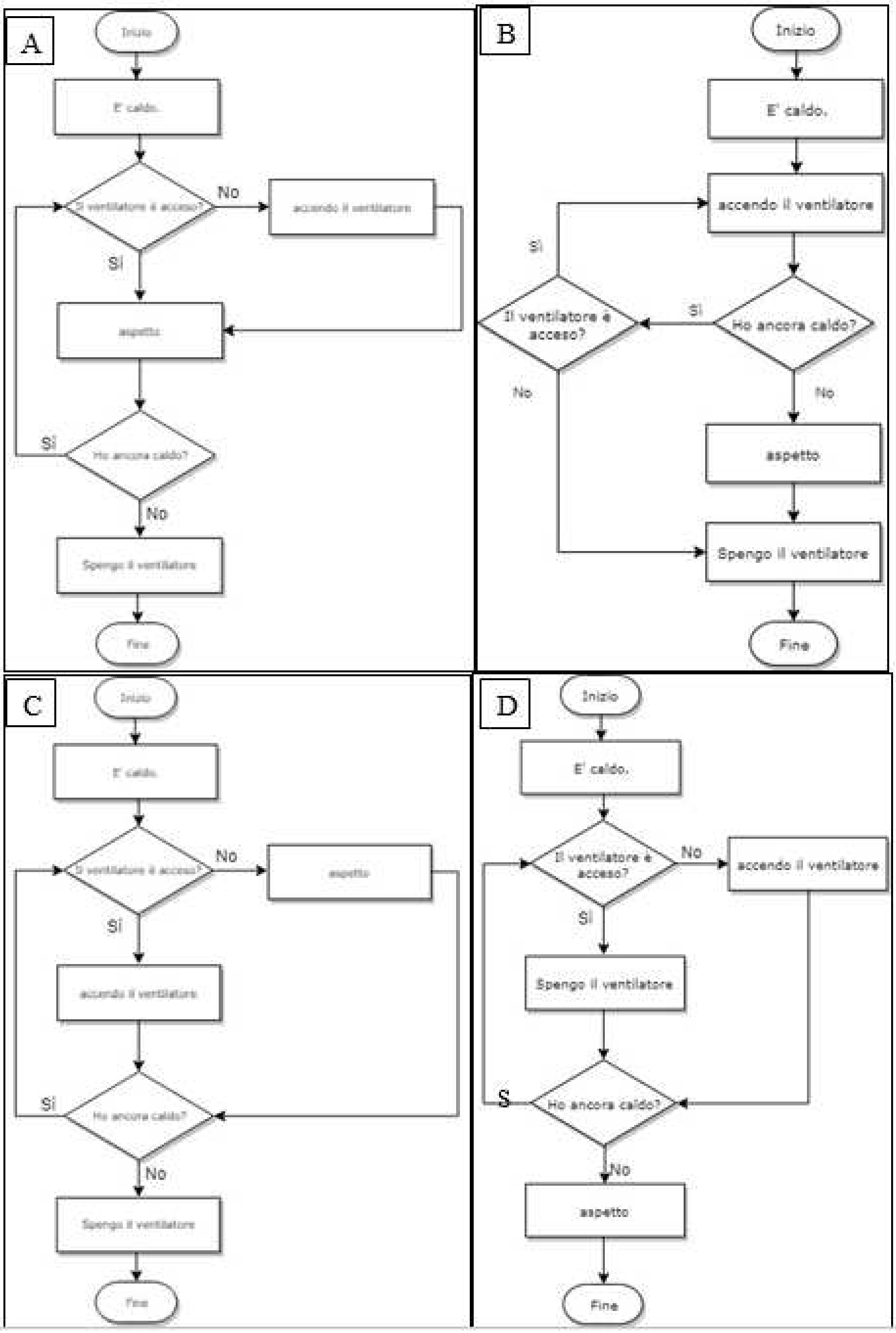
B


INIZIO
Accendo smartphone
Digito PIN
Il PIN inserito è corretto?
No
Sì
Accedo alle app dello smartphone
FINE

C


INIZIO
Digito PIN
Accendo smartphone
Il PIN inserito è corretto?
No
Sì
Accedo alle app dello smartphone
FINE

10) Segna la risposta che corrisponde al diagramma CORRETTO.



11) Nella seguente lista di componenti scegli 3 elementi per costruire un sistema che muova automaticamente la tapparella della finestra. Spiega in 5 righe come funziona il sistema che hai immaginato.

- sensore di luce
- scheda elettronica di controllo
- sensore a Ultrasuoni
- videocamera
- motore
- pulsante

12) Individua le parti principali del seguente elettrodomestico: ROBOT ASPIRAPOLVERE

Knowledge questionnaire for students of the Lower Secondary School in the post-activities phase

1) Il sensore a Ultrasuoni
    a. fa avvicinare il robot all'ostacolo, toccarlo e tornare indietro.
    b. fa fare un percorso al robot senza che vada a sbattere.
    c. permette al robot di seguire una linea nera sopra un telo bianco.

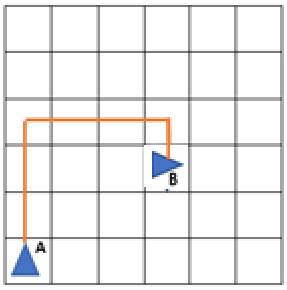
2) Affermazione 1: il motore è un ingresso per l'unità centrale.
   Affermazione 2: il sensore è un'uscita per l'unità centrale.
    a. Affermazione 1 è vera e Affermazione 2 è falsa.
    b. Affermazione 1 è falsa e Affermazione 2 è falsa.
    c. Affermazione 1 è falsa e Affermazione 2 è vera.

3) Le affermazioni che vedi sono vere o false?

	VERO	FALSO
Il microfono è la voce del robot.	☐	☐
La batteria serve al robot per pensare.	☐	☐
Le ossa del robot sono i blocchi che costituiscono la sua struttura meccanica.	☐	☐
L'udito del robot è l'altoparlante.	☐	☐
Il blocchetto intelligente è il cervello del robot.	☐	☐
Il robot vede lo spazio che lo circonda grazie ai motori.	☐	☐
Il robot si muove grazie ai motori.	☐	☐

4) Metti in ordine le seguenti istruzioni in modo tale che la freccia può arrivare al punto B partendo dal punto A (inserisci i numeri vicino la corrispondente istruzione).

	Gira a sinistra di 90°
	Avanti di 1 posizione
	Avanti di 3 posizioni
	Avanti di 3 posizioni
	Gira a destra di 90°
	Gira a destra di 90°
	Fine
	Inizio



5) Riempi gli spazi vuoti con la parola che trovi nell'elenco sottostante.

"Un ambiente di programmazione ____________ consente di costruire un _____________ tramite dei blocchetti colorati. Ogni _____________ rappresenta un comando per il ______________. Ci sono per esempio i blocchi per _____________ i motori, attraverso alcuni parametri."

- robot
- programma
- blocchetto
- visuale
- controllare

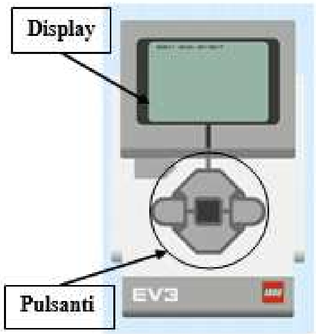
6) Che cos'è un sensore?

☐ È un componente del robot responsabile di acquisire dati
☐ È un componente del robot responsabile del movimento del robot
☐ È un componente del robot responsabile dell'elaborazione dei dati

7) Voglio che il robot non cada dal tavolo, quindi programmo il robot affinché

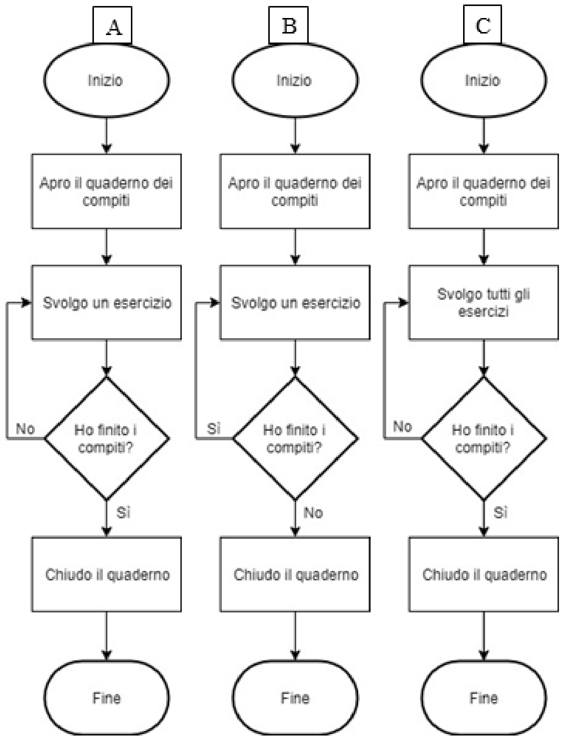
☐ si fermi quando la distanza rilevata dal sensore è uguale a 20 cm
☐ si fermi quando la distanza rilevata dal sensore è maggiore di 20 cm
☐ si fermi quando la distanza rilevata dal sensore è minore di 20 cm

8) Individua gli input/ingressi (IN) e output/uscite (OUT) sul mattoncino intelligente.



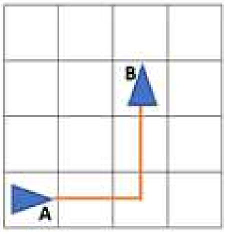
	Ingresso/Input	Output/Uscita
Motore a elevata potenza	☐	☐
Sensore a Ultrasuoni	☐	☐
Sensore di luce	☐	☐
Pulsanti	☐	☐
Display	☐	☐
Sensore di colore	☐	☐
Motore a potenza media	☐	☐
Sensore di contatto	☐	☐
Giroscopio	☐	☐

9) Quale tra i seguenti è il diagramma di flusso che meglio rappresenta le azioni che devo compiere per finire i compiti?

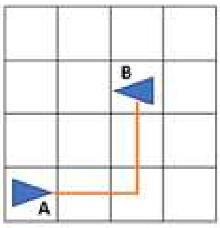


10) Quale immagine rappresenta le seguenti istruzioni?
Dalla posizione A, la freccia gira a sinistra di 90 gradi.
Poi la freccia va avanti di due caselle.
Poi la freccia gira a destra di 90 gradi.
Infine, la freccia va avanti di due caselle e raggiunge così la posizione B.

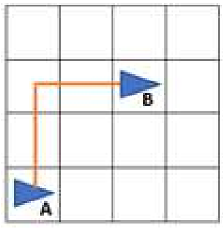
A



B



C



11) Nella seguente lista di componenti scegli **3 elementi** per costruire un sistema che muova **automaticamente** la una **carrozzina**. Spiega in 5 righe come funziona il sistema che hai immaginato. Ecco alcune cose che potrebbero esserti utili.

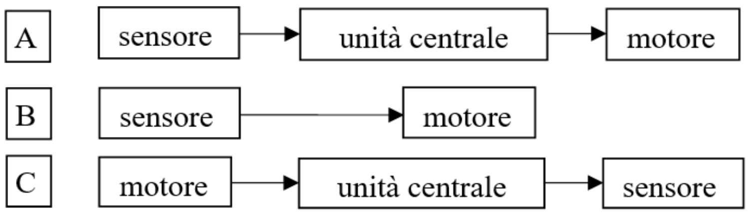
- sensore di luce
- scheda elettronica di controllo
- sensore a ultrasuoni
- videocamera
- motore
- pulsante
- joystick

12) Individua le parti principali del seguente elettrodomestico: ROBOT TAGLIAERBA.

Knowledge questionnaire for primary school's students in the pre-activity phase

1) Come possiamo usare il sensore a Ultrasuoni?
    a. per misurare la distanza tra il robot e l'ostacolo.
    b. per misurare i suoni dentro la stanza.
    c. per azionare gli altoparlanti del robot.

2) Quale di queste configurazioni consente al robot di controllare il suo movimento di fronte a un ostacolo?

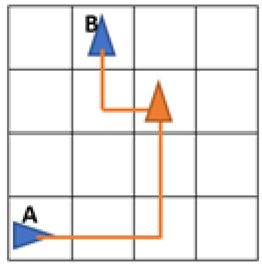


3) Associa le parti di un robot (riportate nella lista a sinistra) con le parti del corpo umano (riportate nella lista a destra). Associa le parti che svolgono una funzione simile nel robot e nel corpo umano.

Parti di un robot	**Parti del corpo umano**
H. Motori	□ Ossa/scheletro
I. Videocamera	□ Bocca/corde vocali
J. Microfono	□ Cuore
K. Unità centrale di elaborazione delle informazioni	□ Muscoli
L. Batteria	□ Udito
M. Struttura meccanica	□ Vista
N. Altoparlante	□ Cervello

4) Costruisci la sequenza che ti permette di arrivare dal punto A al punto B (inserisci i numeri vicino la corrispondente istruzione).

	Avanti di 1 posizione
	Avanti di 1 posizione
	Avanti di 2 posizioni
	Avanti di 2 posizioni
	Gira a sinistra di 90°
	Gira a sinistra di 90°
	Gira a destra di 90°
	fine
	inizio



5) Un ambiente di programmazione visuale contiene:
- ☐ solo testo
- ☐ comandi per i robot rappresentati da dei blocchi colorati
- ☐ solo frecce e pulsanti

6) Che cos'è un motore?
- ☐ È un componente del robot responsabile di acquisire dati
- ☐ È un componente del robot responsabile del movimento del robot
- ☐ È un componente del robot responsabile dell'elaborazione dei dati

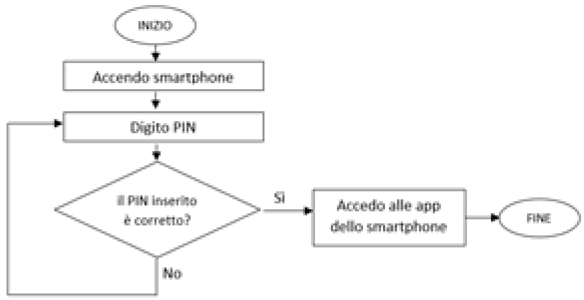
7) Voglio che il robot non tocchi il muro, quindi programmo il robot affinché:
   - ☐ i motori si spengano quando la distanza rilevata dal sensore è uguale a 30 cm
   - ☐ i motori si spengano quando la distanza rilevata dal sensore è maggiore di 30 cm
   - ☐ i motori si spengano quando la distanza rilevata dal sensore è minore di 30 cm

8) Individua gli input/ingressi (IN) e output/uscite (OUT) sul mattoncino intelligente.

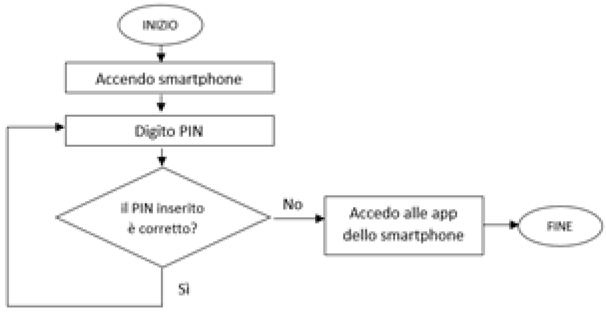
	Ingresso/Input	Output/Uscita
Motore a elevata potenza	☐	☐
Sensore a Ultrasuoni	☐	☐
Sensore di luce	☐	☐
Sensore di colore	☐	☐
Sensore di contatto	☐	☐
Giroscopio	☐	☐

9) Voglio usare lo smartphone, ma devo inserire il PIN di sblocco. Quale tra i seguenti è il diagramma di flusso che meglio rappresenta le azioni che devo compiere?

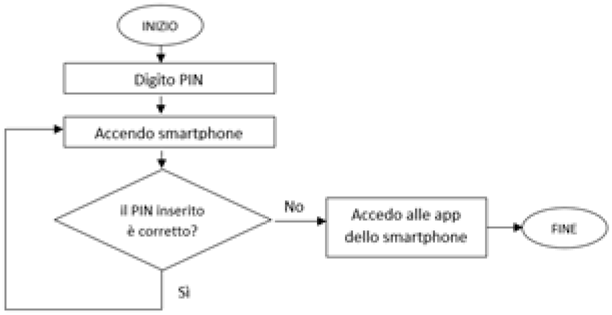
A


INIZIO
Accendo smartphone
Digito PIN
il PIN inserito è corretto?
Sì
No
Accedo alle app dello smartphone
FINE

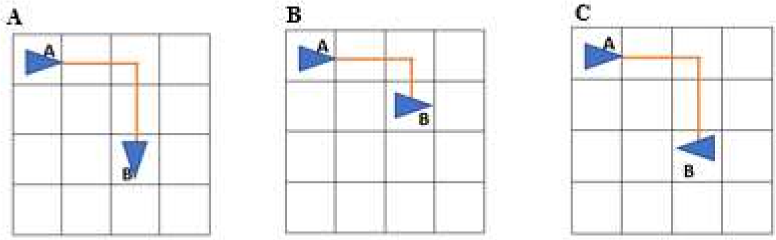
B


INIZIO
Accendo smartphone
Digito PIN
il PIN inserito è corretto?
No
Sì
Accedo alle app dello smartphone
FINE

C


INIZIO
Digito PIN
Accendo smartphone
il PIN inserito è corretto?
No
Sì
Accedo alle app dello smartphone
FINE

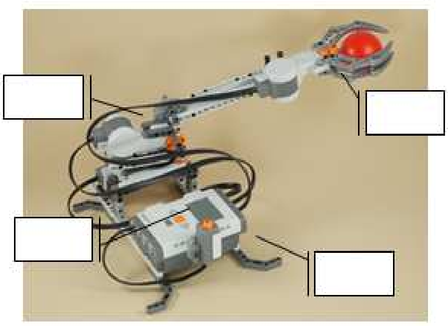
10) Quale immagine rappresenta le seguenti istruzioni?
Dalla posizione A, la freccia va avanti di due caselle.
Poi la freccia gira a destra di 90 gradi.
Poi la freccia va avanti di due caselle.
Infine, la freccia gira a destra di 90 gradi e raggiunge così la posizione B.



11) Come possiamo rendere un oggetto intelligente? Pensa ad un oggetto comune e a come renderlo "automatico". Potrebbe essere un ventilatore, uno spazzolino da denti, una forchetta, etc. oppure un oggetto che piace a te! Scegli tre elementi dalla lista e spiega in 5 righe come funziona il sistema che hai immaginato. Ecco alcune cose che potrebbero esserti utili:

- sensore di luce
- scheda elettronica di controllo
- sensore a ultrasuoni
- videocamera
- motore
- pulsante

12) Individua le parti principali del seguente robot:



Ecco alcuni suggerimenti:
- motore
- motore
- scheda di controllo
- sensore di contatto

Knowledge questionnaire for primary school's students in the post-activities phase

1) Il sensore a Ultrasuoni
    a. fa avvicinare il robot all'ostacolo, toccarlo e tornare indietro.
    b. fa fare un percorso al robot senza che vada a sbattere.
    c. permette al robot di seguire una linea nera sopra un telo bianco.

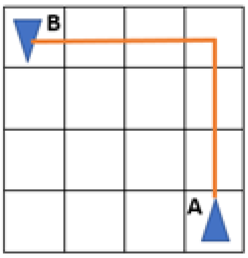
2) Affermazione 1: il motore è un ingresso per l'unità centrale.
   Affermazione 2: il sensore è un'uscita per l'unità centrale.
    a. Affermazione 1 è vera e Affermazione 2 è falsa.
    b. Affermazione 1 è falsa e Affermazione 2 è falsa.
    c. Affermazione 1 è falsa e Affermazione 2 è vera.

3) Le affermazioni che vedi sono vere o false?

	VERO	FALSO
Il microfono è la voce del robot.	☐	☐
La batteria serve al robot per pensare.	☐	☐
Le ossa del robot sono i blocchi che costituiscono la sua struttura meccanica.	☐	☐
L'udito del robot è l'altoparlante.	☐	☐
Il blocchetto intelligente è il cervello del robot.	☐	☐
Il robot vede lo spazio che lo circonda grazie ai motori.	☐	☐
Il robot si muove grazie ai motori.	☐	☐

4) Costruisci la sequenza che ti permette di arrivare dal punto A al punto B (inserisci i numeri vicino la corrispondente istruzione).

	Avanti di 3 posizioni
	Avanti di 3 posizioni
	Gira a sinistra di 90°
	Gira a sinistra di 90°
	fine
	inizio



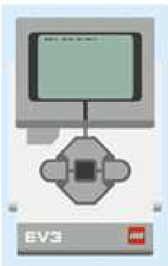
5) Riempi gli spazi vuoti con la parola che trovi nell’elenco sottostante.
“Un ambiente di programmazione _______________ consente di costruire un _____________ tramite dei blocchetti colorati. Ogni ________________ rappresenta un comando per il ______________. Ci sono per esempio i blocchi per _____________ i motori, attraverso alcuni parametri.”

- robot
- programma
- blocchetto
- visuale
- controllare

6) Che cos’è un sensore?
- ☐ È un componente del robot responsabile di acquisire dati
- ☐ È un componente del robot responsabile del movimento del robot
- ☐ È un componente del robot responsabile dell’elaborazione dei dati

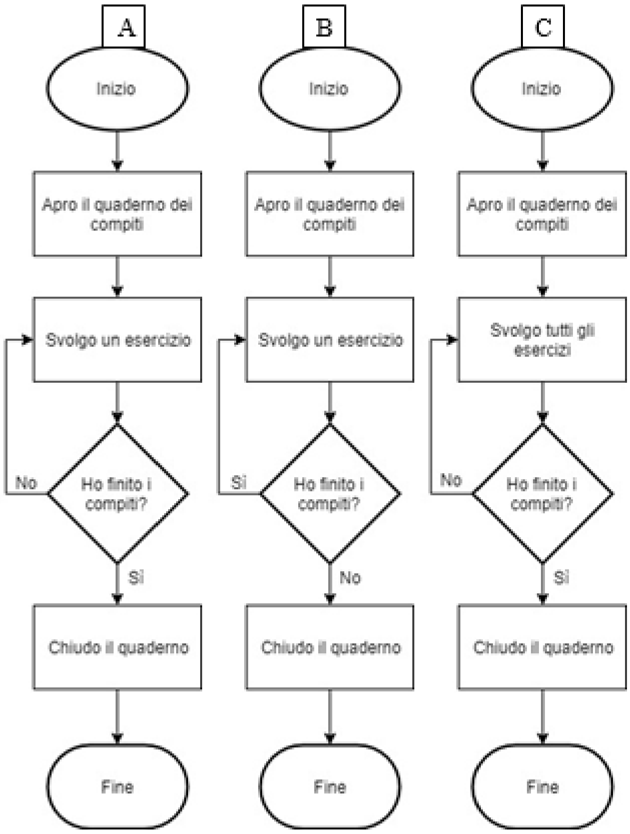
7) Voglio che il robot non cada dal tavolo, quindi programmo il robot affinché
- ☐ si fermi quando la distanza rilevata dal sensore è uguale a 20 cm
- ☐ si fermi quando la distanza rilevata dal sensore è maggiore di 20 cm
- ☐ si fermi quando la distanza rilevata dal sensore è minore di 20 cm

8) Individua gli input/ingressi (IN) e output/uscite (OUT) sul mattoncino intelligente.

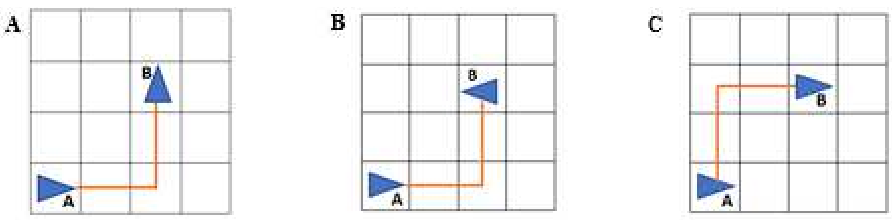


	Ingresso/Input	Output/Uscita
Motore a elevata potenza	☐	☐
Sensore a Ultrasuoni	☐	☐
Sensore di luce	☐	☐
Sensore di colore	☐	☐
Sensore di contatto	☐	☐
Giroscopio	☐	☐

9) Quale tra i seguenti è il diagramma di flusso che meglio rappresenta le azioni che devo compiere per finire i compiti?



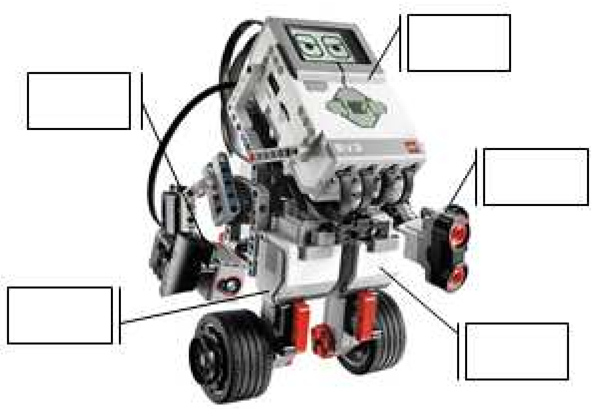
10) Quale immagine rappresenta le seguenti istruzioni?
Dalla posizione A, la freccia gira a sinistra di 90 gradi.
Poi la freccia va avanti di due caselle.
Poi la freccia gira a destra di 90 gradi.
Infine, la freccia va avanti di due caselle e raggiunge così la posizione B.



11) Come possiamo rendere un oggetto intelligente? Pensa ad un oggetto comune e a come renderlo "automatico". Potrebbe essere un ventilatore, uno spazzolino da denti, una forchetta, etc. oppure un oggetto che piace a te! Scegli tre elementi dalla lista e spiega in 5 righe come funziona il sistema che hai immaginato. Ecco alcune cose che potrebbero esserti utili:

- sensore di luce
- sensore a ultrasuoni
- videocamera
- motori
- pulsanti
- joystick
- scheda elettronica di controllo

12) Individua le parti principali del seguente robot.



Ecco alcuni suggerimenti:
- motore
- motore
- scheda di controllo
- sensore di luce
- sensore a ultrasuoni

Appendix C

S-STEM Questionnaire – Italian modified version

Istruzioni:
Nelle pagine che stai leggendo troverai una lista di frasi. Scegli la tua risposta tra le opzioni presenti selezionando quella che più rispecchia il tuo pensiero sulle frasi.
Esempio:

	Molto in disaccordo	In disaccordo	Né d'accordo, né in disaccordo	D'accordo	Molto d'accordo
Mi piace Ingegneria	○	○	○	○	○

Nel leggere la frase saprai se sei d'accordo o non sei d'accordo con quanto stai leggendo. Riempi o metti una croce sul cerchietto sotto la categoria che meglio descrive quanto sei d'accordo o non sei d'accordo.
Anche se alcune frasi sono molto simili tra loro, rispondi ad ogni frase.
Non ci sono limiti di tempo per rispondere al questionario. Lavora velocemente, ma con attenzione. Non ci sono risposte giuste o sbagliate! Le uniche risposte giuste sono quelle che sono vere per te. Dove possibile, lascia che qualcosa che ti è capitato ti aiuti a decidere quale risposta è la più giusta per te.
Dai una sola risposta per ogni frase.

<u>Matematica:</u>
Q1. Matematica è la materia in cui vado peggio
Q2. Farei studi o lavori che usino la matematica
Q3. La matematica è difficile per me
Q4. Io sono il tipo di studente che in matematica va bene
Q5. Vado bene nella maggior parte delle materie ma in matematica non riesco bene
Q6. Sono sicuro/a di poter fare un lavoro che usi la matematica a livello avanzato

Q7. Posso ottenere buoni voti in matematica
Q8. Sono bravo/a in matematica

<u>Scienze:</u>
Q9. Mi sento sicuro/a di me stesso/a quando faccio scienze
Q10. Prenderei in considerazione un lavoro nel campo delle scienze
Q11. Mi aspetto di usare le scienze una volta che avrò terminato il mio percorso di studio
Q12. Conoscere le scienze mi aiuterà a guadagnarmi da vivere
Q13. Mi servirà conoscere le scienze per il mio futuro lavoro
Q14. So di poter esser bravo/a in scienze
Q15. Le scienze saranno importanti nel lavoro che vorrei fare
Q16. Vado bene nella maggior parte delle materie, ma in scienze non riesco bene
Q17. Sono sicuro/a di poter fare un lavoro che usi le scienze a livello avanzato

<u>Ingegneria e tecnologia:</u>
Prima di rispondere alle domande leggi le righe riportate qui di seguito.
Gli ingegneri usano la matematica, la scienza e la creatività per risolvere dei problemi che migliorano la vita di tutti e per inventare nuovi prodotti. Ci sono molti tipi diversi di ingegneria, come per esempio ingegneria chimica, elettronica, meccanica, informatica, civile, ambientale e biomedica. Gli ingegneri progettano e migliorano le cose come i ponti, le auto, i tessuti, il cibo e la realtà virtuale per i parchi divertimento. I tecnici realizzano i progetti che gli ingegneri sviluppano. I tecnici costruiscono, verificano e mantengono i prodotti e i processi.

Q18. Mi piace immaginare di creare nuovi prodotti
Q19. Se studierò ingegneria, potrò migliorare le cose che la gente usa tutti i giorni
Q20. Sono bravo/a a costruire o riparare le cose
Q21. Mi interessa sapere come funziona una macchina, un dispositivo, un apparato, etc.
Q22. Progettare prodotti o strutture sarà importante nel mio lavoro futuro
Q23. Sono curioso/a di sapere come funziona l'elettronica
Q24. Mi piacerebbe usare la creatività e l'innovazione nel mio futuro lavoro

Q25. Sapere come usare la matematica e le scienze insieme mi consentirà di inventare cose utili
Q26. Credo che io posso avere successo in un lavoro nel campo dell'ingegneria

Abilità del 21esimo secolo:
Q27. Sono sicuro/a che posso guidare altri al raggiungimento di un obiettivo
Q28. Sono sicuro/a che posso incoraggiare gli altri a dare il meglio di sé
Q29. Sono sicuro/a che posso produrre lavori di alta qualità
Q30. Sono sicuro/a che posso rispettare le differenze dei miei compagni
Q31. Sono sicuro/a che posso aiutare i miei compagni
Q32. Sono sicuro/a che posso tenere in considerazione i punti di vista degli altri quando prendo una decisione
Q33. Sono sicuro/a che posso fare dei cambiamenti quando le cose non vanno come previsto
Q34. Sono sicuro/a che posso stabilire i miei obiettivi di studio da solo/a
Q35. Sono sicuro/a che posso organizzare bene il mio tempo quando lavoro da solo/a
Q36. Quando ho molti compiti, sono in grado di scegliere quali devono esser fatti per primi
Q37. Sono sicuro/a che posso lavorare bene con altri studenti che hanno conoscenze, competenze ed esperienze differenti dalle mie

Robotica:
Prima di rispondere alle domande leggi le righe riportate qui di seguito.
La Robotica si occupa di costruire robot che aiutino l'uomo nello svolgere i lavori più faticosi o difficili. I robot possono avere grandi dimensioni (come un robot industriale per sollevare carichi pesanti) o piccole (come i robot che aiutano in casa per le pulizie o i piccoli droni), di qualsiasi forma (umanoide come NAO, a blocchetti come i LEGO, con le ruote come mBot) e utili per qualunque lavoro (andare sottacqua, esplorare Marte, compiere un lavoro ripetitivo, guidare un cliente verso il giusto scaffale del supermercato, aiutare i medici durante un'operazione). Un robot ha una struttura meccanica, dei sensori e degli attuatori per interagire con l'ambiente esterno e un "cervello" che può essere programmato. Costruire un robot significa lavorare in squadra per poterlo costruire e programmare.

Q38. Mi piace costruire i robot
Q39. Mi piace programmare i robot
Q40. Se costruirò robot, potrò aiutare le persone nel loro lavoro e nella loro vita
Q41. Sono bravo/a a costruire o riparare un robot
Q42. Mi interessa sapere come funziona un robot
Q43. Conoscere la robotica sarà importante nel mio lavoro futuro
Q44. Sono curioso/a di sapere come funziona un robot
Q45. Mi piacerebbe usare i robot nel mio futuro lavoro
Q46. Mi piacerebbe costruire i robot nel mio futuro lavoro
Q47. Mi piacerebbe programmare i robot nel mio futuro lavoro
Q48. Conoscere la robotica mi consentirà di inventare cose utili
Q49. Credo che io possa avere successo in un lavoro del campo della robotica
Q50. Mi piacerebbe fare ancona attività di robotica in classe

Il tuo futuro

Fisica: è lo studio delle leggi di base che governano il moto, l'energia, la struttura e le interazioni della materia, incluso lo studio della natura dell'universo. (ingegnere aeronautico, tecnico delle energie alternative, fisico, astronomo)

Lavoro per l'ambiente: include imparare i processi fisici e biologici che governano la natura e lavorare per migliorare l'ambiente.Questo include trovare e progettare una soluzione a problemi come l'inquinamento, riuso dei materiali di scarto e riciclo.(tecnico della protezione idrogeologica e del suolo, tecnico per il controllo dell'inquinamento, ingegnere ambientale, ingegnere dei sistemi energetici, tecnico manutentore)

Biologia e Zoologia: include lo studio degli organismi viventi (come le piante o gli animali) e i processi associati alla vita. Questo include lavorare in fattorie con gli animali e in aree come nutrizione e allevamento. (tecnico biologo, biologo, agronomo, tecnico di laboratorio per colture, zoologo)

Veterinaria: include le scienze che prevengono o curano le malattie degli animali. (assistente veterinario, veterinario, produttore di carni, custode di animali)

Matematica: è la scienza dei numeri e delle operazioni. Include calcoli, algoritmi e teorie usate per risolvere i problemi e riassumere i dati. (ragioniere, matematico applicato, economista, analista finanziario, matematico, statistico, analista di mercato, analista di borsa)

Medicina: include il mantenimento della salute, la prevenzione e cura delle malattie. (assistente medico, infermiere, dottore, nutrizionista, tecnico di medicina di emergenza, fisioterapista, dentista)
Scienze della terra: è lo studio della terra in senso ampio, inclusi aria, terra e oceani. (geologi, meteorologo, archeologi, geoscienziati)
Informatica: consiste nello sviluppo e nella verifica di sistemi di calcolatori, progettando nuovi programmi e aiutando gli altri a usare il computer. (tecnico dei computer, sviluppatore di programmi informatici, tecnico di reti di computer, sviluppatore di videogiochi, ingegnere informatico, specialista IT)
Scienze Mediche: include il campo della ricerca nell'area delle malattie umane e lavorare per trovare soluzioni ai problemi di salute umani. (tecnico di laboratorio biomedico, scienziato nell'ambito della medicina, ingegnere biomedico, specialista di epidemiologia, farmacologo)
Chimica: usa la matematica e gli esperimenti per cercare nuove sostanze chimiche e per studiare la struttura della materia e come si comporta. (tecnico chimico, chimico, ingegnere chimico)
Energia: include lo studio e la produzione di energia, come il calore o l'elettricità. (elettricista, ingegnere elettrico; tecnico per il riscaldamento, la ventilazione e il condizionamento dell'aria; ingegnere nucleare; ingegnere energetico; installatore o tecnico di sistemi di energie rinnovabili)
Ingegneria: include la progettazione, verifica e manifattura di nuovi prodotti (come macchine, ponti, edifici e elettronica) attraverso l'uso della matematica, scienza e computer. (ingegnere civile, industriale, agrario, meccanico; saldatore, meccanico, tecnico ingegnere, capocantiere)

Quanto ti aspetti di andar bene quest'anno in:
Italiano, Inglese ed educazione artistica
Matematica
Scienze
Qualcosa di più su di te
Conosci qualche persona adulta che di lavoro fa lo/la scienziato/a
Conosci qualche persona adulta che di lavoro fa l'ingegnere/ingegnera?
Conosci qualche persona adulta che di lavoro fa il/la matematico/a?
Conosci qualche persona adulta che di lavoro fa il/la tecnico/a?

Appendix D

Expectations questionnaire

Quiz robotico iniziale

- On a 5-points Likert scale (1= “per niente”, 5= “moltissimo”)

1. Sarà facile costruire i robot in gruppo
2. Sarà facile usare il programma al computer
3. Non vedo l’ora di frequentare le attività del laboratorio
4. Durante le attività in classe il clima sarà sereno
5. Nel mio gruppo ci sarà collaborazione e sostegno
6. Lavorando in gruppo migliorerò la mia relazione con uno o più compagni
7. Questa proposta della scuola mi sembra divertente e appassionante

8 Mi piacerebbe prendere parte ad altre attività di robotica in futuro

- Free text answer:

Che cosa imparerai in questo laboratorio?
Che cosa ti potrebbe piacere di questo laboratorio
Che cosa potrebbe non piacerti di questo laboratorio?
Che cosa vorresti fare durante questo laboratorio?

Satisfaction questionnaire

Quiz robotico finale

- On a 5-points Likert scale (1= “per niente”, 5= “moltissimo”)

1. Ho capito le istruzioni che mi ha dato l'istruttore
2. L’istruttore è stato disponibile e attento alle mie domande e bisogni
3. Il metodo usato dall’istruttore è stato coinvolgente
4. È stato facile costruire i robot
5. È stato facile utilizzare i software al computer
6.Ho frequentato volentieri le attività del laboratorio
7. il clima in classe è stato sereno
8. Il lavoro nel mio gruppo è stato di collaborazione e sostegno
9. La mia relazione con uno o più compagni è migliorata
10. Ho trovato divertente e appassionante partecipare a questo tipo di proposta della scuola
11. Mi piacerebbe prendere parte ad altre attività di Robotica in futuro

- Free text answer:

12. Che cosa hai imparato in questo laboratorio?
13. Qual è stata la cosa che ti è piaciuta di più in questo laboratorio?
14. Qual è stata la cosa che ti è piaciuta di più in questo laboratorio?
15. C’è qualcos’altro che avresti voluto fare nel laboratorio?