Thèse n° 9563

Towards the alignment of educational robotics learning systems with classroom activities

Présentée le 19 novembre 2020

à la Faculté des sciences et techniques de l'ingénieur
Laboratoire de biorobotique
Programme doctoral en robotique, contrôle et systèmes intelligents

pour l'obtention du grade de Docteur ès Sciences

par

Christian GIANG

Acceptée sur proposition du jury

Prof. S. Micera, président du jury
Prof. F. Mondada, Dr A. Piatti, directeurs de thèse
Prof. R. Nagpal, rapporteuse
Prof. J. Cascalho, rapporteur
Prof. T. Käser Jacober, rapporteuse

École
polytechnique
fédérale
de Lausanne

2020

If you think you are too small to make a difference, try sleeping with a mosquito.

— African proverb quoted by the Dalai Lama

Acknowledgements

The past three years have been a unique experience for personal growth and learning and in retrospect, I can say that pursuing this doctoral degree was one of the best decisions I have made so far. I enjoyed every single moment and I am very grateful for all the people that have made this dissertation possible.

First, I would like to thank my thesis advisors Francesco and Alberto. Thank you Francesco for hosting me in your lab, for providing me with valuable input for my research, for allowing me to participate in so many interesting projects and for always giving me the freedom that I needed. Thank you Alberto for believing in me from the very beginning, for inspiring and encouraging me to explore new ideas, for all your supportive advice and for always being available when I needed help. Both of you made this experience very enjoyable and I wish for any doctoral student to have such a great duo of advisors.

I would also like to thank the jury members of my thesis committee. Thank you Silvestro, Tanja, Radhika and José for taking the time to evaluate my thesis work and for giving me your much appreciated feedback.

Thank you also to all my colleagues from the MOBOTS, LEARN-GE and TME groups. I am fortunate to have met and collaborated with such amazing groups of people. Thank you everyone for all the interesting conversations and discussions. Special thanks go out to: Frank, for being an inspiring PostDoc and for always sharing your expertise and knowledge. Dany and Norbert, for all your help with the IT and mechanical questions in my projects. Eliot for your Friday sandwich delivery and anti-foodwaste services. Vaios, for introducing me to the culinary delights of Greece. Genia, for teaching me so many things about Thymio and crafting. Amaury, for being another inspiring PostDoc that was also my bodyguard during the Swiss Digital Day. Frédérique for taking me to Evian, which marked the beginning of the Thymio Escape Game. Morgane, for introducing me to the world of pedagogy and for the great collaborations in our projects. Sophia, for always making the sun shine, even when the sun is not shining outside. Melissa, for making my American dream come true. Yves, for all your help related to the VPL platform. Jessica for always ensuring a good air quality in our shared office space. Laila, for enriching the supervision of the student projects with your valuable inputs and your great personality. Lucio, for helping me to find my bearings in Locarno and for always sharing your pedagogical expertise. Thierry, Luca and Masiar for your support in the research studies at DFA. Claudia, Sandrine, Corinne, Sandra, Donatella, Monica and Alessia for all your precious support to get me through the jungle of paperwork. Michael, Fanny, Christophe and Vincent from Mobsya for participating in my research studies and for all the help related to Thymio. Thank you Ran for your feedback on the first version of the Thymio Escape Game.

Personally, I found the most enjoyable part of the past three years was the supervision of student projects. Thank you Andrea, Noé, Julien, Aditya and Anthony for all your great work. I have learned so much during these projects and I really appreciate the contributions you made to my projects. I was certainly inspired by my fantastic supervisors from the time when I was doing my master thesis. Thank you Nawal, Elvira, Martina and Camilla for supporting and encouraging me to pursue a career in research. Thank you also Leonid for sparking my interest and passion for research in the first place.

I would also like to thank my parents Tam and Bac for paving the way towards this thesis by teaching me the relevant skills and attitudes from early age on. Thanks to my brother Max and Monika for providing me with a roof over my head (actually multiple) between all the train rides. Thanks to the crazy pilgrims, Jacob, Ludo, Dimitri, Daniel, Beatriz, Mirabela, Nicoleta, Oussama and Evan for also making my leisure time adventurous, exciting and fun. Thanks of course to Isabela, who made this whole journey even more wonderful. Thank you for keeping me grounded when I start to float too far and for teaching me so many important things about life (volleyball included).

Finally, I would like to thank Tracy for providing me with my first (and best) collocation in Lausanne. Thank you Alessandro and Linda for the German-Italian-French tandems that got me perfectly prepared for this position. And last but not least, thank you Helio and Shania for testing the first educational robotics activity I ever designed - the one that eventually initiated this journey.

Lausanne, 11 October 2020 C. G.

Abstract

The technological advances of the past years have impressively demonstrated that the digital age is no longer just a science fiction vision of the future - we are in the midst of it. The digital transformation is affecting many areas of our lives, including the educational system. For instance, recent years have shown an increased adoption of educational robotics activities in classrooms. Although educators and researchers have acknowledged the potential of educational robotics, further efforts are needed to effectively integrate them into formal education. As for any kind of educational tool, one key factor for the successful integration of educational robots in classrooms is the proper alignment of such tools with the classroom activities and vice versa. In practice, however, this may pose challenges. On the one hand, it still appears that only few educators have already developed the required know-how allowing them to leverage educational robotics for classroom activities. On the other hand, the issue of poor alignment may already arise in the design stages of educational robotics tools, due to the limited experience of developers with the pedagogy and learning theories related to such tools.

This thesis, therefore, seeks to study how instructional alignment can be attained in the context of educational robotics. To this end, an alignment framework will be devised aimed at supporting developers in the design of educational robotics tools, as well as educators in the design and implementation of classroom activities involving educational robotics. In this context, we will introduce the notion of Educational Robotics Learning Systems (ERLS), a model that conceptualizes alignment for educational robotics activities. In different studies, we will illustrate how the devised alignment framework can be used to guide the development of educational robotics tools and activities. The *PaPL* and *CreroBot* projects will consider the alignment issue from the developers' perspective. The former is concerned with the development of a tangible programming interface based on accessible materials such as paper and cardboard. The latter will then expand these ideas to the development of a do-it-yourself educational robot. In two other projects, we will then illustrate how the devised framework can be used to guide the development of educational robotics activities. We will first discuss the *Thymio lawnmower mission*, an activity devised to foster the development of students' computational thinking competencies. Finally, we will present the *Thymio Escape Game*, an immersive learning activity that has been inspired by escape room experiences.

The findings from these studies are intended to make a contribution to the field of educational robotics, specifically by addressing the global research question of this thesis: *How can instructional alignment be attained for educational robotics activities?*

Keywords: computational thinking, constructionism, educational robotics, game-based learning, heuristic evaluation, instructional alignment, socio-constructivism, tangible programming

Zusammenfassung

Die technologischen Fortschritte der vergangenen Jahre haben eindrucksvoll gezeigt, dass das digitale Zeitalter nicht mehr nur eine Science-Fiction Zukunftsvision ist - wir befinden uns mittendrin. Der digitale Wandel wirkt sich auf viele Bereiche unseres Lebens aus, das Bildungssystem mit eingeschlossen. So werden in letzter Zeit auch immer mehr Lern- und Lehraktivitäten durchgeführt, in denen Bildungsroboter eingesetzt werden. Pädagogen und Forscher haben das Potenzial der Bildungsrobotik anerkannt - allerdings braucht es weitere Bemühungen, um sie wirkungsvoll in den Schulunterricht zu integrieren. Wie für jedes pädagogische Instrument, ist ein Schlüsselfaktor für die erfolgreiche Integrierung von Bildungsrobotern, die richtige Anpassung (auf englisch *alignment*) solcher Instrumente an den Schulunterricht und umgekehrt. In der Praxis kann dies jedoch Herausforderungen mit sich bringen. Einerseits scheint es immer noch, dass nur wenige Pädagogen bereits das erforderliche Knowhow entwickelt haben, welches es ihnen ermöglicht, die Bildungsrobotik wirkungsvoll im Schulunterricht einzusetzen. Andererseits kann das Problem mangelnden Alignments bereits in der Entwicklung solcher Instrumente auftauchen, da Entwickler oftmals nur begrenzte Erfahrung mit Pädagogik und mit denen der Bildungsrobotik zugrundeliegenden Lerntheorien haben.

Daher ist das Ziel dieser Dissertation zu untersuchen, wie Alignment im Kontext der Bildungsrobotik erreicht werden kann. Zu diesem Zweck werden wir ein theoretisches Gerüst entwickeln, welches Entwicklern beim Design von Instrumenten für die Bildungsrobotik helfen soll. Zudem soll es gleichermassen Pädagogen bei der Konzeption und Umsetzung von Aktivitäten der Bildungsrobotik unterstützen. In diesem Zusammenhang werden wir das Konzept von Educational Robotics Learning Systems (ERLS) einführen, welches ein Modell für das Alignment von Aktivitäten der Bildungsrobotik beschreibt. Wir werden dann in verschiedenen Studien veranschaulichen, wie die erarbeitete Methodik verwendet werden kann, um die Entwicklung von Instrumenten der Bildungsrobotik sowie die Gestaltung von Lern- und Lehr-Aktivitäten zu begleiten. Die Projekte *PaPL* und *CreroBot* werden die Alignment-Frage aus der Entwicklerperspektive betrachten. Ersteres befasst sich mit der Entwicklung einer physischen Programmierplattform, welche auf leicht verfügbaren Materialien wie Papier und Karton basiert. Das zweite Projekt wird den Ansatz dann auf die Entwicklung eines

Do-it-yourself-Bildungsroboters ausweiten. In zwei weiteren Forschungsprojekten werden wir dann veranschaulichen, wie die erarbeitete Methodik zur Gestaltung von Aktivitäten der Bildungsrobotik genutzt werden kann. Zunächst werden wir die *Thymio Rasenmäher-Mission* vorstellen, eine Lernaktivität, die konzipiert wurde, um die Entwicklung von Computational Thinking Kompetenzen zu fördern. Schliesslich werden wir das *Thymio Escape Game* präsentieren, eine immersive Lernaktivität, welche durch sogenannte Escape Rooms inspiriert wurde.

Die Erkenntnisse aus diesen Studien zielen darauf ab, einen wissenschaftlichen Beitrag im Bereich der Bildungsrobotik zu leisten, insbesondere in Bezug auf die übergreifende Forschungsfrage dieser Dissertation: *Wie kann Alignment für Aktivitäten der Bildungsrobotik erreicht werden?*

Stichwörter: Alignment, Bildungsrobotik, Computational Thinking, Heuristische Evaluierung, Konstruktionismus, Physische Programmierplattformen, Sozio-Konstruktivismus, Spielbasiertes Lernen

Résumé

Les avancées technologiques de ces dernières années ont démontré de façon impressionnante que l'ère numérique n'est plus seulement une vision de science-fiction du futur - nous sommes au milieu de celle-ci. La transformation numérique touche de nombreux domaines de notre vie, y compris le système éducatif. Par exemple, ces dernières années ont montré une adoption accrue des activités robotiques éducatives dans les salles de classe. Bien que les éducateurs et les chercheurs aient reconnu le potentiel de la robotique éducative, des efforts supplémentaires sont nécessaires pour l'intégrer efficacement dans l'éducation scolaire. Comme pour tout type d'outil éducatif, un facteur clé pour l'intégration réussie des robots éducatifs dans les salles de classe est l'alignement approprié de ces outils avec les activités en classe et vice versa. Dans la pratique, cependant, cela peut poser des défis. D'une part, il semble encore que peu d'éducateurs ont déjà développé le savoir-faire nécessaire leur permettant de tirer parti de la robotique éducative pour les activités en classe. D'autre part, le problème du mauvais alignement peut déjà se poser dans les phases de conception des outils pour la robotique éducative, en raison de l'expérience limitée des développeurs en matière de pédagogie et de théories d'apprentissage liées à ces outils.

Cette thèse vise donc à étudier comment l'alignement pédagogique peut être atteint dans le contexte de la robotique éducative. À cette fin, un cadre d'alignement sera conçu pour aider les développeurs à concevoir des outils pour la robotique éducative, ainsi que les éducateurs à concevoir et à mettre en œuvre des activités en classe utilisant la robotique éducative. Dans ce contexte, nous introduirons la notion de systèmes d'apprentissage de la robotique éducative (ERLS), un modèle qui conceptualise l'alignement pour les activités de robotique éducative. Dans différentes études, nous illustrerons comment le cadre d'alignement conçu peut être utilisé pour guider le développement d'outils et d'activités de robotique éducative. Les projets *PaPL* et *CreroBot* examineront la question de l'alignement du point de vue des développeurs. Le premier concerne le développement d'une interface de programmation tangible basée sur des matériaux accessibles tels que le papier et le carton. Le deuxième élargira ensuite ces idées à la mise au point d'un robot éducatif à réaliser soi-même. Dans deux autres projets, nous illustrerons ensuite comment le cadre conçu peut être utilisé pour guider le développement d'activités de robotique éducative. Nous parlerons tout d'abord de la

Thymio tondeuse, une activité conçue pour favoriser le développement des compétences de pensée computationnelle des élèves. Enfin, nous présenterons le *Thymio Escape Game*, une activité d'apprentissage immersive qui s'inspire des expériences vécues dans les salles d'évasion.

Les résultats de ces études sont destinés à apporter une contribution au domaine de la robotique éducative, en particulier en abordant la question de recherche globale de cette thèse : *Comment peut-on parvenir à un alignement pédagogique pour les activités de robotique éducative ?*

Mots clefs: alignement pédagogique, apprentissage par le jeu, constructionnisme, évaluation heuristique, pensée computationnelle, programmation tangible, robotique éducative, socio-constructivisme

Abstract

I progressi tecnologici degli ultimi anni hanno dimostrato in modo impressionante che l'era digitale non è più solo una visione fantascientifica del futuro - siamo nel bel mezzo di essa. La trasformazione digitale sta influenzando molte aree della nostra vita, compreso il sistema educativo. Ad esempio, gli ultimi anni hanno mostrato una maggiore adozione delle attività di robotica educativa nelle aule scolastiche. Sebbene educatori e ricercatori abbiano riconosciuto il potenziale della robotica educativa, sono necessari ulteriori sforzi per integrarla efficacemente nella formazione scolastica. Come per qualsiasi tipo di strumento educativo, un fattore chiave per il successo dell'integrazione dei robot educativi nelle aule è il corretto allineamento di tali strumenti con le attività in classe e viceversa. In pratica, tuttavia, questo può rappresentare una sfida. Da un lato, sembra che solo pochi educatori abbiano già sviluppato il know-how necessario per poter sfruttare la robotica educativa per le attività in classe. Dall'altro lato, il problema dello scarso allineamento può già sorgere nelle fasi di progettazione degli strumenti di robotica educativa, a causa della limitata esperienza degli sviluppatori con la pedagogia e le teorie di apprendimento relative a tali strumenti.

Questa tesi, quindi, cerca di studiare come l'allineamento pedagogico può essere raggiunto nel contesto della robotica educativa. A tal fine, sarà elaborato un quadro di allineamento volto a supportare gli sviluppatori nella progettazione di strumenti di robotica educativa, così come gli educatori nella progettazione e nella realizzazione di attività didattiche in classe che coinvolgono la robotica educativa. In questo contesto, introdurremo il concetto di sistemi di apprendimento della robotica educativa (ERLS), un modello che concettualizza l'allineamento per le attività di robotica educativa. In diversi studi, illustreremo come il quadro di allineamento ideato possa essere utilizzato per guidare lo sviluppo di strumenti e attività di robotica educativa. I progetti *PaPL* e *CreroBot* prenderanno in considerazione la questione dell'allineamento dal punto di vista degli sviluppatori. Il primo riguarda lo sviluppo di un'interfaccia di programmazione tangibile basata su materiali accessibili come carta e cartone. Il secondo estenderà poi queste idee allo sviluppo di un robot educativo fai-da-te. In altri due progetti, illustreremo poi come il quadro ideato possa essere utilizzato per guidare lo sviluppo di attività di robotica educativa. Per prima cosa parleremo del *Thymio tosaerba*, un'attività concepita per favorire lo sviluppo delle competenze di pensiero computazionale degli allievi.

Infine, presenteremo il *Thymio Escape Game*, un'attività di apprendimento immersiva che si ispira alle esperienze delle escape room.

I risultati di questi studi mirano a dare un contributo al campo della robotica educativa, in particolare affrontando la domanda di ricerca globale di questa tesi : *Come si può raggiungere l'allineamento pedagogico per le attività di robotica educativa ?*

Parole chiave : allineamento pedagogico, apprendimento basato su giochi, costruttivismo, pensiero computazionale, programmazione tangibile, robotica educativa, socio-costruttivismo, valutazione euristica

Contents

List of Figures

List of Tables

1 Introduction

Past research has suggested that Educational Robotics (ER) has a potential as an innovative learning and teaching tool for STEM education (Alimisis, 2013; Eguchi, 2017; Karim et al., 2015). Systematic reviews on the use of ER have concluded that this potential exists not only for topics related to robotics, but can also be extended to other subjects and disciplines (Benitti, 2012; Toh et al., 2016). A recent review (Anwar et al., 2019) has also emphasized the multi-disciplinarity of ER and it has suggested that ER can support the education of students "who do not display immediate interest in academic disciplines related to science and technology". Furthermore, previous work has shown that ER activities can be used to foster the development of 21st century skills (Eguchi, 2014; Khanlari, 2013) as well as computational thinking competencies (Atmatzidou & Demetriadis, 2016; Bers et al., 2014; Chen et al., 2017).

Although educators and researchers have acknowledged the potential of educational robots, they are still not as widespread in classrooms as this potential suggests (Mondada et al., 2017). As for any kind of educational tool, the successful integration of ER tools in classrooms requires them to be well aligned with the classroom activities and vice versa. In practice, however, this question has not been studied extensively.

In a recent work, Antonenko et al. (2017) have highlighted that only little conceptual research has been performed to construct useful and usable frameworks helping educational technology developers to align their systems with the needs of the target users (i.e., teachers and students). With respect to the specific case of ER, Pachidis et al. (2018) have made similar suggestions, proposing that "future research needs to deal with the more effective design of robots to align with the educational need, in terms of hardware and software". As a matter of fact, developers of ER tools are usually researchers or engineers, who often have limited experience with the pedagogy and the learning theories related to such tools. Conceptual frameworks guiding the development of ER tools could therefore make an important contribution to the alignment of ER tools with classroom activities. Indeed, striving for alignment already in the design stages, would promote the development of ER tools, that are inherently aligned with classroom needs and are thus educationally meaningful.

But even perfectly designed ER tools are worthless, if they are not appropriately adopted for the classroom activities. Hence alignment does not only have to be considered in the development stages of ER tools, but also in the design and final implementation of the classroom activities. However, the use of ER in formal education is a relatively recent trend, and not many educators have already developed the required know-how allowing them to effectively integrate ER into their classroom activities. Indeed, except for some pioneer teachers that have already successfully adopted ER into their teaching, the majority of educators have barely even come into contact with educational robots. At the same time, there is still a lack of research related to the pedagogy and the underlying learning theories of ER (Alimisis, 2013; Jung & Won, 2018). However, for ER tools to be successfully integrated into classroom activities, it is crucial that educators develop an understanding of their qualities to be able to consciously design and implement learning and teaching activities that effectively leverage them. Also in this context, a conceptual framework on the use of ER in classrooms could help educators in better planning, designing and implementing their classroom activities. Moreover, such a framework could support educators in identifying suitable ER tools for their classroom activities and provide guidance about how to effectively use them.

Instructional alignment is an established principle that has served as a framework to guide curriculum design for many years. As described by Cohen (1987), it considers the interplay of three instructional components: intended outcomes, instructional processes and instructional assessment. Previous works have shown that deliberately aiming for alignment of these components, can significantly improve students' achievements (Mitchell, 1999; Wonder-McDowell et al., 2011). The traditional alignment framework, however, may not be considered sufficient, when classroom activities are augmented with educational technology, such as ER. The integration of ER tools in the learning system adds a new dimension to the alignment issue, that has to be taken into consideration. The new challenges but also opportunities that are accompanying these tools, need to be appropriately captured, hence implying an adaptation of the traditional alignment framework to account for the integration of ER tools in the learning system.

To this end, the present dissertation aims at developing a conceptual alignment framework, that is specifically adapted to the use of ER in classroom environments. This framework may guide developers in the design of new ER tools as well as educators, who would like to design and implement ER activities in their classrooms. The framework has been applied in different case studies to assess its usability and utility for the design of ER tools as well as for the design and implementation of ER activities.

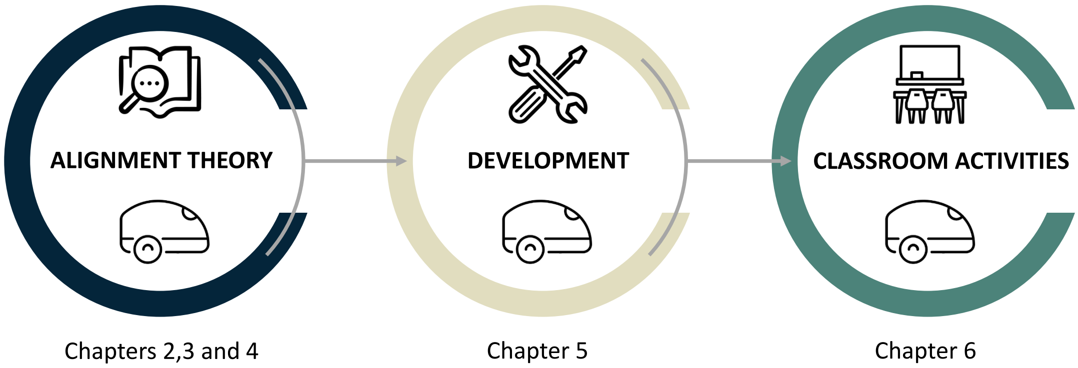

Figure 1.1 – Three main stages of alignment addressed in this thesis.

1.1 Organization of this thesis

A main objective of this thesis is to provide diverse perspectives on the alignment issue, making contributions to support developers in the design of ER tools as well as educators in the implementation of ER classroom activities. To this end, alignment of ER tools with classroom activities has been addressed in three main stages (Fig. 1.1).

In the first stage, we will elaborate how the theory of instructional alignment can be applied as a guiding principle in the context of ER classroom activities:

- In Chapter 2, we will present a literature review on previous works that have discussed instructional alignment, while putting a particular emphasis on studies about the alignment of educational technologies. Based on the current state of the art we will then formulate the main objectives of this thesis work.
- In Chapter 3, we will first specify the term "ER activities" as referred to in this thesis. Following this definition, we will identify the main elements of ER activities and discuss relevant learning theories, that are then used to develop a conceptual model for the alignment of Educational Robotics Learning Systems (ERLS).
- In Chapter 4, we will outline the elaboration of HERLS, a set of design heuristics for ERLS, that represent a more operational way to guide developers and educators in aligning their tools and activities within the ERLS framework.

The second stage of this thesis will be fully dedicated to the development of ER tools:

- In Chapter 5, we will illustrate, how the ERLS framework and the HERLS heuristics have been applied to guide the development of ER tools. In two examples, we will showcase how alignment can already be considered by developers in the design of their ER tools.

The first example will present the development of Thymio PaPL, a tangible programming interface for the educational robot Thymio. In the second example, we will outline the development of CreroBot, an accessible do-it-yourself educational robot.

The third stage of this thesis will then address how instructional alignment with ER tools can be achieved by appropriately designing and implementing the learning activities:

- In Chapter 6, we will discuss two examples of learning activities that have been aligned using the ERLS framework and the HERLS heuristics as guiding tools. In the first example we will present the Thymio lawnmower mission, a learning activity aimed at fostering students' computational thinking competencies. Specifically, we will illustrate how two different ways of implementing the same activity may result in varying degrees of alignment, eventually altering students' learning experiences. In the second example, we will discuss the development of the Thymio Escape Game, a learning activity inspired by escape room experiences. Here we will outline the development of an activity that has been specifically adapted for classroom use and we will illustrate how proper alignment can promote favorable learning experiences.

Finally, this dissertation will be closed with a general conclusions chapter:

- In Chapter 7, we will present a synthesis of the results obtained from the different studies and discuss the main contributions of this thesis. The work will be closed with a discussion on the limitations of the studies and a presentation of directions for future works.

2 Literature review

2.1 Summary

In this chapter, we will present a literature review on previous works that have discussed the principle of instructional alignment, with a particular emphasis on studies that involved educational technologies. We will illustrate how the principle of alignment has been utilized to guide the development of educational technologies as well as the design of classroom activities using educational technologies. Moreover, we will present previous works that have specifically discussed alignment in the context of educational robotics. Finally, based on the current state of the art, we will present the main objectives of this thesis.

2.2 Instructional alignment

Instructional alignment is an established principle in curriculum planning, considering the interplay of three components: intended outcomes, instructional processes and instructional assessment. The main idea is that during the process of course planning, educators should select and design each component in a way that they support each other (Fig. 2.1). For instance, misalignment would occur if the instructional processes would not address the topics or skills that have been originally set as the intended outcomes. Likewise, assessing topics that have not been covered in the instructional processes represents another example of poor alignment. While good educators may sometimes intuitively achieve well-aligned course designs, alignment can also be attained through the explicit design of the curriculum (Wiggins et al., 2005).

Indeed, several works have shown that deliberately aiming for alignment, appeared to be more effective. In his work, Cohen (1987) has emphasized how comparatively small efforts of instructional alignment have shown large effect sizes in different studies (Elia, 1988; Fahey, 1988; Koczor, 1986; Tallarico, 1985). These studies also suggested that alignment may especially benefit low achievers as compared to high achievers.

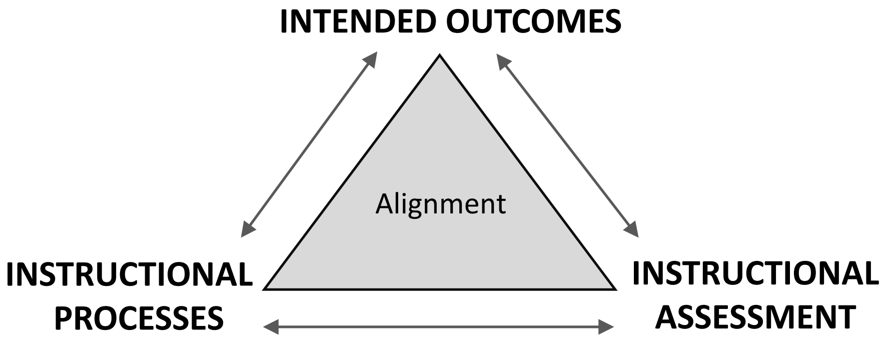

Figure 2.1 – Traditional model of instructional alignment.

A similar finding was presented by Wonder-McDowell et al. (2011), who investigated the effect of aligning classroom core reading instructions with supplementary reading instructions provided to second graders with reading difficulties. Students were separated in two treatment conditions, characterized by aligned and unaligned instruction. Though students in both conditions improved their reading abilities after the 20-week study, students in the aligned condition evidenced higher gains.

In a large-scale study with over 4000 third graders, Mitchell (1999) showed that implementing curriculum alignment for one year, resulted in statistically significant improvements of student achievement in mathematics. In addition, Mitchell concluded that instructional alignment may act as a "curriculum equalizer", since it appeared to weaken the predictability of variables such as socioeconomic status, ethnicity, gender or school size.

By linking the ideas of instructional alignment with the learning theory of constructivism, Biggs (1996) introduced the concept of "constructive alignment". To this day, constructive alignment has had a strong influence on curriculum design, especially in higher education. The main idea is to use the constructivist learning approach as a guiding framework at all stages of the instructional design and alignment. As described by Biggs, this refers to the principle that learning activities should help students in autonomously constructing meaning for the learning, rather than having the meaning imparted from the teacher to the students (Biggs, 2003). The role of the teacher is to "set up a learning environment that supports the learning activities appropriate to achieving the desired learning outcomes". Moreover, he emphasized the importance of considering teaching and learning as a system, in which ideally "all aspects of teaching and assessment are tuned to support high-level learning". To achieve a "good system", Biggs (2003) proposed a backwards design approach based on four major steps: the design process should start with the definition of the desired learning outcomes. Based on the defined learning outcomes, educators should then choose teaching and learning activities that would likely lead to achieving these learning outcomes. Finally the students' actual learning outcomes are assessed to see how well they match with what was intended. Based on this assessment students could then be graded.

When Biggs introduced his ideas of constructive alignment, he also suggested the use of educational technology as a favorable approach. However, with respect to traditional instructional alignment, the integration of educational technology implies a new dimension of complexity and should therefore be carefully considered by both developers and educators. In the following we will therefore discuss previous efforts made to achieve alignment of educational technology with educational needs.

2.3 Alignment of educational technology

The technological advances of the past decades have given rise to an increased integration of educational technology in classrooms. Consequently, researchers have started to study how alignment can be attained when this new class of educational tools is integrated into the learning system. Previous works have studied how alignment of educational technology with educational needs can be achieved by both developers and educators (Antonenko et al., 2017; Bower, 2008; Bower & Sturman, 2015; P. Kirschner et al., 2004; Mishra & Koehler, 2009; Osborne, 2014; Seau Yoon et al., 2005). In this context, the notion of affordances has served many works as a concept to assess the alignment of educational technology with educational needs.

2.3.1 Affordances of educational technologies

Originally introduced as a psychological concept, Gibson (1979) defined affordance in the context of relations between an animal and its surrounding environment:

> *"The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. The verb to afford is found in the dictionary, the noun affordance is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment."* (Gibson, 1979)

At the time when Gibson first introduced the term, affordances were indeed not included in dictionaries. However, nowadays a definition has been added to many English dictionaries, for instance, the Cambridge Online Dictionary:

> *"affordance: a use or purpose that a thing can have, that people notice as part of the way they see or experience it"* (Cambridge Online Dictionary, accessed 4th August 2020)

This definition is a result of the evolution that the term has made ever since it was first introduced by Gibson. As a matter of fact, the notion of affordances has been adopted by many fields, including the human-computer interaction and the design community. A great influence on today's understanding of affordances is rooted in the definition suggested by Norman (1988):

> *"the term affordance refers to the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used."* (Norman, 1988)

Following Norman's definition, affordances can be interpreted as design aspects of objects, that suggest how they should be used. Norman particularly emphasized, that affordances have to be perceived by the potential user to be useful. It is based on this definition, that several domains, including the educational technology community, have referred to affordances as "action possibilities", i.e., properties of the technology, for which users can perform the necessary actions to utilize them. Based on this understanding of affordances, the term has served many research works studying how educational technology designs can be aligned with educational needs. Among others, affordance analyses have been used as a design approach for developers and likewise, as a method for educators in the selection of appropriate educational technologies for their teaching purposes.

2.3.2 Alignment through affordance analysis

Different research works have used the notion of affordances to study the alignment of educational technology designs with educational needs (Antonenko et al., 2017; Bower, 2008; P. Kirschner et al., 2004). Affordance analyses have been used to inform the design of educational technology as well as to support educators in the selection of appropriate tools for their instructional activities.

P. Kirschner et al. (2004) for instance, have presented a framework for the design of computer supported collaborative learning (CSCL) environments. The authors first introduced and discussed three types of affordances that they considered to be relevant for the development of CSCL environments, namely *technological, educational* and *social affordances.* Furthermore, they argued that affordances are not just independent features that can simply be designed for a system. Instead, they are by definition (as presented in the previous section), highly dependent on the relationship between the object and the user. To account for this relationship, the authors suggested that the design of CSCL systems should consider interaction design approaches, rather than simple interface design. The authors further argued, that the usefulness of a system is formed by the combination of its utility (whether it provides what is needed) and its usability (how easy and pleasant it is to use). While utility can be related to the social and educational affordance of CSCL systems, their usability is defined by the technological affordances of the environment. As a matter of fact, interaction design concerns itself with both usability and utility and consequently, the authors proposed an interaction design approach for CSCL environments based on a six-stage model. In a first step, designers should determine what students actually do in the concerned learning activities. The authors suggested to observe students and collaborating groups even before designing and developing the system. Following this step, designers should then determine what can be done to support the students identifying required affordances of the system. This is followed by a constraint analysis, considering all kind of constraints that may play a role in the learning activities (i.e., technological, educational and social constraints). The final three stages are then concerned with the iterative improvement of the system. Designers should ask themselves how students perceive the support provided by the system and how they actually use it. Finally, designers should also evaluate what students have actually achieved using the devised CSCL system. The proposed methodology was then applied in three different research projects to highlight the importance of the presented affordances as well as the need for interaction design to address both usability and utility of CSCL environments.

In the same vein, Bower (2008) proposed a methodology to support educators in the design of e-learning tasks based on an affordance analysis. The authors first compiled a list of eleven e-learning technology affordances, referring to them as "action possibilities" offered to the user: *media affordances, spatial affordances, temporal affordances, navigation affordances, emphasis affordances, synthesis affordances, access-control affordances, technical affordances, usability, aesthetics* and *reliability*. While the authors acknowledged that the last four categories may not directly concern the educational point of view per se, they also highlighted their importance for the quality of the learning experience and related them to the usability of a system. Based on this list of affordances, the authors then discussed how an affordance anaylsis can be performed for e-learning technology as part of the design process. Similar to the backwards design approach presented by Biggs (2003), the design process starts with the identification of the educational goals. Based on these goals, educators should then postulate suitable tasks, that may lead to attaining these goals. Following this step, an affordance anaylsis is performed based on the categories proposed by the author. On the one hand, educators should determine what kind of affordances are needed for the intended tasks. At the same time, the affordances provided by the technology available should be considered. The synergistic consideration of available and required affordances would then eventually result in the final e-learning task design. Bower especially emphasized that the consideration of required and provided affordances should be performed simultaneously, due to two main reasons. First, only determining technological affordances without considering the needs of the task, can result in unnecessary and hence inefficient analysis. Second, only determining the affordance requirements of the task without considering the affordances provided by the technology, can result in unfeasible solutions. In the remainder of the work, Bower illustrated the utility of the methodology using the example of an online graduate degree program, for which appropriate online tools had to be determined. He concluded that the proposed affordance analysis could be useful to educators, helping them to understand how the underlying attributes of technologies can support collaboration and cognition and eventually assisting them in the identification of appropriate educational tools for their instructional activities.

In a recent work, Antonenko et al. (2017) pointed out that the field of educational technology would benefit from more explicit frameworks, allowing to align user needs with technological affordances. The authors argued that, often, development of educational technology has been driven by what can be developed rather than what should be developed. According to the authors, one issue is that many educational technology developers do not perform analyses with respect to their target population, especially when the group of developers is rather small, like for instance, in the design of mobile applications. Another issue they have pointed out, is that developers who do collaborate with instructional designers, often do not succeed to actually match the needs of the target users with the affordances of the technology they are developing. To address this issue, the authors therefore proposed an alignment framework consisting of ten educational technology affordances: *media affordances, spatial affordances, temporal affordances, navigation affordances, emphasis affordances, synthesis affordances, metacognitive affordances, personalization affordances,* and *socialization affordances*. The categories have been developed by adopting the affordances presented by P. Kirschner et al. (2004) and Bower (2008), and additionally integrating the multimodal text affordances presented by Gall and Breeze (2005). In a case study, the authors then illustrated how this framework can be used to evaluate the alignment of user needs and affordances of educational technologies. In this context, the authors especially emphasized the importance of addressing the users' needs, an element that would make their approach distinguishable from the framework of Bower (2008). Instead of aligning the technology with common learning tasks, the authors aimed at an alignment with the needs of a particular learning group. As an example, the authors presented the case of a technology for dyslexic learners. Based on existing literature, they first defined the user needs. Subsequently, they performed an affordance analysis using the list of educational technology affordances presented in their work, to identify and assess a potentially useful educational technology. The authors concluded that the proposed framework could help both educators and developers in aligning educational technology with classroom use. While designers could use the framework as a checklist for the design of educational technology, it could also support educators in the selection of appropriate tools. However, the authors also acknowledged, that the framework may not be usable by developers with little background in psychology and education, as well as by other stakeholder groups such as parents, tutors and informal educators.

2.3.3 Alignment of classroom activities

The affordance analyses presented in the previous subsection, may guide developers in aligning their designs with educational needs and it may also support educators in choosing the appropriate tools for their teaching. However, such approaches cannot guarantee that educational technology will finally also be used appropriately for classroom activities. Hence alignment does not stop with the devised educational technology - it continues in the classroom, where educators use such technologies to design and implement classroom activities. A perfectly designed and adequately selected educational technology tool is worthless, if the educator does not know how to appropriately use it to implement teaching and learning activities.

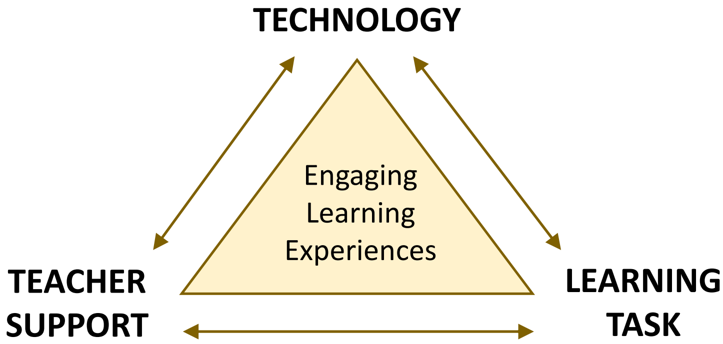

Figure 2.2 – Tripartite model for engaging learning experiences with ICT adapted from Seau Yoon et al. (2005).

The importance of alignment for successful teaching with educational technology has been illustrated in the work of Seau Yoon et al. (2005). In a case study with six teachers, they studied how the teachers designed and implemented learning activities with information communication technologies (ICT). The teachers taught lessons in core subjects (such as math, English, geography, sciences or social studies) to pupils of varying grades at primary and secondary schools. In two lessons, each teachers used one or more ICT tools (such as discussion forums, online polls, robotics dataloggers or geographic information systems) for their classroom activities. The first lesson was planned by the teachers alone, while for the second one they received a reflection and feedback session with the researchers beforehand. The researchers analyzed the lessons by collecting data through observation forms, interviews and video analyses and defined pupils' engagement as the main observable of the study. Based on their findings, the authors suggested, that teachers may achieve higher engagement of pupils when they consciously consider the alignment of three elements: learning tasks, teacher support and the technology used (Fig. 2.2). For instance, lessons appeared to be more engaging when teachers designed the learning task with the pupils' ability in mind, appropriately chose the ICT tools considering their affordances and provided appropriate teacher support. Lessons of teachers who neglected one part of the tripartite model, appeared to be less engaging. For instance, one teacher, who used a robotics datalogger for sciences classes, planned well the learning task and the use of the ICT tool, providing students with both worksheets for the task and manuals for the tool. However, not much thought was given to the intended teacher support, leaving pupils struggling with the set-up. The authors finally concluded that the proposed tripartite model could serve as a useful framework for educators, helping them to understand how to effectively design and implement ICT-supported lessons, and hence leading to more engaging learning activities.

As a matter of fact, it is always possible for educators to adapt technologies to fit their needs - this applies even to those technologies that were not primarily designed for educational purposes. In their work, Mishra and Koehler (2009) showed three examples of how technologies such as microblogging, specialized search engines or even DJ software can be repurposed

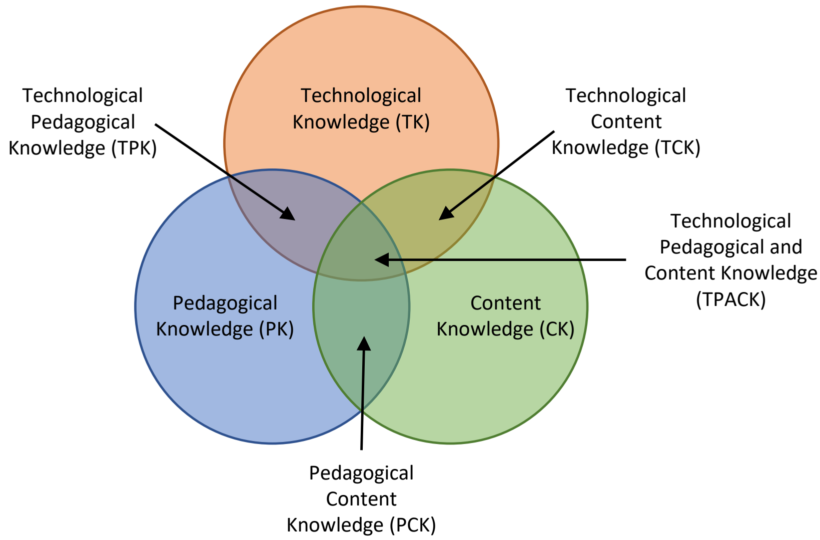

Figure 2.3 – Schematic of the TPACK framework adapted from Koehler and Mishra (2009).

to align with educational needs. However, in order to so, educators require a set of skills, that has been introduced in the Technological Pedagogical and Content Knowledge (TPACK) framework (Koehler & Mishra, 2009; Mishra & Koehler, 2006). The framework was built on the notion of Pedagogical Content Knowledge (PCK) introduced by Shulman (1986) and extended it with a Technological Knowledge (TK) component (Fig. 2.3). Shulman described PCK as the knowledge enabling educators to understand how to appropriately and effectively teach a specific subject matter. It is hence the "content knowledge for teaching". As exemplified by Mishra and Koehler (2009), this specialized knowledge is what distinguishes a good math teacher from a highly trained mathematician. By introducing technological knowledge (TK) as a third component to the model, two new sub-classes of knowledge have emerged from the TPACK model of Koehler and Mishra (2009):

- **Technological Content Knowledge (TCK)**: This sub-class describes the knowledge of how the applied technology and the content to be taught influence and constrain one another. Understanding this relationship helps educators to decide which technology can be used to address the teaching and learning of particular content.
- **Technological Pedagogical Knowledge (TPK)**: This sub-class describes the knowledge of how a particular use of a technology can influence teaching and learning. For educators to build TPK, it is important that they have a thorough grasp of the constraints and affordances of the applied technology, in order to appropriately harness them for pedagogical designs.

For teachers to successfully integrate technology in their teaching activities, they would require all sub-classes of knowledge, summarized as technological pedagogical content knowledge (TPACK).

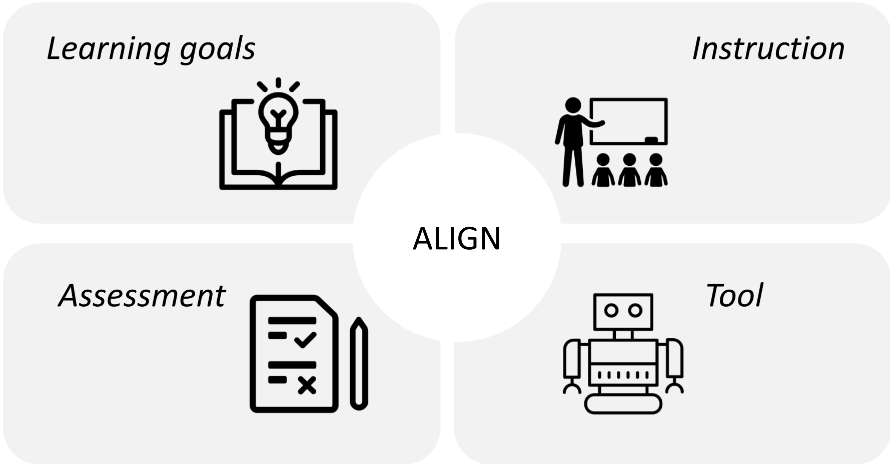

Figure 2.4 – Extended alignment model adapted from Lauwers (2010).

2.4 Alignment of educational robotics

While the existing body of literature has discussed the alignment of educational technology in terms of screen-based software, the specific case of ER appears to be less extensively explored. However, it can be assumed that frameworks developed for screen-based educational technology may not be unreservedly applicable to ER, since the integration of educational robots would imply a reorientation of the alignment analysis. Consequently new frameworks are required to guide such analyses that could help both developers in the design of ER tools as well as educators in the design and implementation of ER classroom activities.

An important step towards a conceptual framework specifically aimed at the alignment of ER with educational needs was presented by Lauwers (2010). Classifying ER as one example of "configurable embodied interfaces", he suggested that the use of such tools would require an adaptation of the traditional alignment model. Lauwers (2010) argued that integrating these tools in the instruction component of the traditional alignment model would not be sufficient, since their use may enable new ways of assessment and even new learning goals. Likewise, learning goals, instructional and assessment activities would also have a strong influence on the design and use of the tool. Therefore, Lauwers proposed to extend the traditional alignment model by a fourth component, that he denoted as the *tool* (Fig. 2.4).

In line with Biggs' propositions, Lauwers also emphasized the importance of considering the entire learning system when designing new tools. He then argued that the main difference between traditional engineering approaches and design processes that are guided by alignment are the constraints that are being imposed. Indeed, there are many different ways of designing technological tools, however, in order to create educationally relevant technology, designers should also be aware of the constraints imposed by the learning goals, assessment and instructional processes used in the course. In the remainder of his work, he then discussed

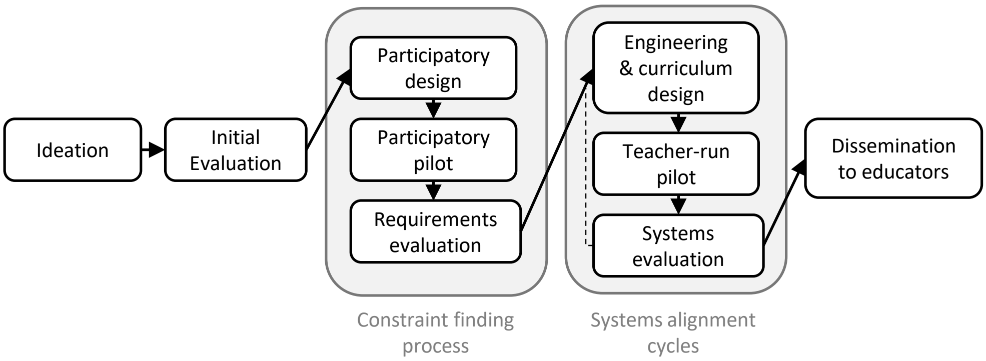

Figure 2.5 – Alignment-centered design process adapted from Lauwers (2010).

participatory design and design-based research as two methods to discover those constraints. Lauwers then suggested an alignment-centered design process based on his findings from three different projects concerned with the design of an electronic Braille writing tutor as well as robotics learning experiences at school and university level. Based on the experiences gained from these projects, Lauwers proposed an alignment-centered design process involving different steps (Fig. 2.5).

According to Lauwers, the design process should start with an idea, that could possibly be interesting to be implemented. This step is followed by an initial evaluation, that can be performed in various ways. In the projects presented by Lauwers, for instance, interviews and focus groups were used for this purpose. The initial evaluation serves as a first step to guide the design, which is followed by the constraint-finding process. It is composed by three elements, that according to Lauwers can often be overlapping: participatory design, participatory pilots and requirements evaluation. The process starts with the participatory design of the tool, involving both teachers and students. This first version of the tool is then tested in participatory pilot studies with the target user group. The process is completed with the requirements evaluation, with the main purpose of identifying the constraints that are inherent to the educational context and the target users. According to Lauwers, "these can be learning goals, required hardware features, instructional activities, or interactions between the hardware or software and the student". In the examples presented by Lauwers, the evaluation was made through the analysis of observation logs, surveys and assessments presented to the students. In fact, the constraint finding process is to some extent, comparable with what we have referred to as an affordance analysis in the previous sections. Following the constraint finding process, the systems alignment cycles are initiated. In this phase, designers take the constraints found from the initial evaluation and the constraint finding process to align the design of their tool, possibly through several iterations. Therefore, the designers start to build the tool and eventually design a curriculum for its use, in case they do not yet exist. The tools are then being tested through teacher-run pilot studies, in which the design team is much less involved compared to the participatory pilots. This is followed by a systems evaluation, which again, can be performed using various methods, both quantitative and qualitative. Based on these evaluations, the designers can then decide, whether a new alignment cycle should be initiated or in case the system is considered satisfactory, whether it can finally be disseminated to educators for use. In this regard, Lauwers highlighted two main types of dissemination, one being the dissemination of ideas and the other the dissemination of the tool. While the first one mainly refers to traditional ways such as academic papers or conference presentations, Lauwers also pointed out that dissemination efforts outside the academic world are required to have an impact on educational practice. Moreover, he also emphasized the importance of disseminating the tool, since it has a potential to further increase the impact on educational practice, though he also acknowledged, that this type of dissemination is usually much more difficult to achieve.

While the work of Lauwers provided some important contributions with regard to alignment of ER (and ER-like) tools from the designer's perspective, it did not explore how such tools could finally be adopted by educators to implement classroom activities. As a matter of fact, the lessons presented in his work were mostly co-designed with the design team. Though undoubtedly favorable, the possibility to be involved in co-design projects is not always available or feasible. Educators without these possibilities would therefore benefit from guiding tools helping them to independently design and implement classroom activities with existing ER tools.

As an approach to address this issue, Catlin and Blamires (2010), compiled a list of ten principles for ER applications, called the ERA principles. According to the authors, these principles could, among other things, serve as a framework for those who want to design educational robots as well as for those who want to develop activities using them. The authors grouped the principles according to the three parties that are involved in the interaction with ER: students, teachers and the technology itself (Table 2.1).

The authors emphasized that the proposed principles are not necessarily independent ideas, however, considered all together they would provide a holistic set of values related to ER. The authors also acknowledged, that the principles as presented in the work, were rather of hypothetical nature and that further work would be necessary to transform them into verified principles. Nevertheless, the principles postulated by the authors can be considered interesting, since they provide a thorough discussion on the underlying research for each principle and presented possible implications for educators and developers. With respect to our current discourse, especially the pedagogy principle appears to be appealing. In the description of this principle, the authors referred to constructionism and constructivsm as the underlying learning theories that would justify the role of ER in education. Moreover, they also briefly introduced a set of strategies (in the form of keywords) that could help educators in creating and analyzing ER activities.

In a later work, Catlin (2017) then further elaborated on this idea, re-branding these strategies as pedagogical tags that could be used to describe ER activities. The work entitled *"29 effective ways you can use robots in the classroom"* presented a set of 29 different tags: *catalyst, chal-*

Table 2.1 – The Educational Robotics Applications (ERA) principles adapted from Catlin and Blamires (2010)

No.	Description
	Technology
1	Intelligence - Educational Robots can have a range of intelligent behaviours that enables them to effectively participate in educational activities.
2	Interaction - Students are active learners whose multimodal interactions with educational robots take place via a variety of appropriate semiotic systems.
3	Embodiment - Students learn by intentional and meaningful interactions with educational robots situated in the same space and time.
	Student
4	Engagement - Through engagement Educational Robots can foster affirmative emotional states and social relationships that promote the creation of positive learning attitudes and environments, which improves the quality and depth of a student's learning experience.
5	Sustainable learning - Educational Robots can enhance learning in the longer term through the development of meta-cognition, life skills and learner self-knowledge.
	Teacher
6	Pedagogy - The science of learning underpins a wide range of methods available for using with appropriately designed educational robots to create effective learning scenarios.
7	Curriculum and assessment - Educational Robots can facilitate teaching, learning and assessment in traditional curriculum areas by supporting good teaching practice.
8	Personalisation - Educational robots personalise the learning experience to suit the individual needs of students across a range of subjects.
9	Equity - Educational robots support principles of equity of age, gender, ability, race, ethnicity, culture, social class, life style and political status.
10	Practical - Educational robots must meet the practical issues involved in organising and delivering education in both formal and informal learning situations.

lenge, coding, conceptualizing, cooperative task, creativity, deductive thinking, demonstration, design, engagement, experience, experimentation, exploration, focused task, games, group work, inductive thinking, links, memorization, modeling, pacifier, problem solving, projects, provocateur, puzzle, relational artifact, topic work and *transfer*. For each tag, Catlin then presented a description and an example of an ER activity that was devised by a teacher. While the proposed tagging scheme could help to better classify different ER activities, it remains open to what extent it could help educators in designing and implementing new activities. The examples presented by Catlin could undoubtedly inspire educators to develop new ideas for ER activities, however, we suggest that more conceptual frameworks are needed to effectively support educators in the application of design principles (such as the ERA principles), by providing guidance in a more structured and operational way.

2.5 Objectives of this thesis

The literature review presented in this chapter has shown that in the past, several efforts have been made to study the alignment of educational technology with classroom needs. Both the viewpoints of designers and educators have been considered and different frameworks and methods have been proposed to attain instructional alignment when integrating educational technology. However, in the specific case of ER, the alignment issue appears to be less extensively explored.

The work of Lauwers (2010) has proposed an extended alignment model as well as an alignment-centered design process that could guide the development of new ER tools. It is undisputed that the contributions made in this work have significantly advanced the question of alignment of ER tools with classroom needs. However, though the proposed alignment model accounts for the integration of the ER tool, it appears that it could still be further elaborated. The use of ER activities is strongly linked with learning theories such as constructionism and socio-constructivism (a discussion on these theories will follow in the next chapter). To appropriately address the question of alignment in the context of ER, we therefore argue that the alignment model has to embody these underlying learning theories. Moreover, it is important that such a model incorporates a formal conceptualization of ER activities that helps developers in identifying and understanding what kind of objects are involved in such activities, which roles they assume and which affordances they provide. Indeed, as we will outline in the next chapter, ER activities may involve more objects than just educational robots. Finally, the work of Lauwers has mainly focused on the alignment issue from the developers' perspective. However, as reasoned earlier in this thesis, alignment should also be considered from the educators' perspective, since ultimately, it is the educators who use the ER tools to design and implement classroom activities. With respect to this question, the work of Catlin and Blamires (2010) has provided some valuable input that could help educators in appropriately designing their ER classroom activities. Especially the pedagogy category of the ERA principles raised some interesting ideas, describing how ER activities could be designed to attain well-aligned instruction. However, we argue that more conceptual and explicit frameworks are needed to effectively assist educators in this matter. To help educators understand which elements and which affordances have to be considered to attain well-aligned instruction with ER tools, we believe that models similar to the one developed by Seau Yoon et al. (2005) for ICT tools, could provide them with more targeted guidance.

One main objective of this thesis is it therefore, to develop an alignment model that embodies all elements considered essential for ER activities as well as the underlying learning theories. Such a model could serve as a general framework for the instructional alignment of activities involving ER tools, that could be equally useful to developers who want to design ER tools, as well as educators who want to design and implement classroom activities with ER tools. While the development of such a model may provide a good framework to address the question of alignment from the theoretical point of view, it might also still be too abstract for some developers and educators to put the ideas into practice, especially when they have limited experience with ER.

Another objective is it hence to develop a methodology that concretizes the ideas of the proposed alignment model, so that developers and educators can put them into practice. We will therefore present a backwards design approach, that outlines different steps that can be followed by developers and educators when designing ER tools and ER activities, respectively. Furthermore, special emphasis will be put on the development of a time- and cost-efficient methodology to support the backwards design approach, in order to allow developers and educators identify crucial design changes without the need of tedious procedures. As highlighted by Antonenko et al. (2017), one limitation of their alignment framework for educational technologies is that it may not be usable by developers with little or no background in psychology and education. Aiming for a more explicit methodology that complements the alignment model may therefore also allow developers with little or no background in education to perform alignment analyses. The same may also apply to those educators who have not yet developed much technological pedagogical knowledge (TPK). We believe that the devised framework could represent a valuable support to complement existing methodologies, such as the alignment-centered design process presented by Lauwers.

A final objective of this thesis is to provide some evidence about the usefulness of the devised alignment model and the related methodologies. Therefore, different studies will be conducted to evaluate the proposed approaches. By applying the methodologies in authentic design contexts, this thesis finally also aims at developing examples of ER tools and ER activities for classroom use that are educationally meaningful. In this context, we will also discuss how the developed tools and activities are going to be disseminated in order to also have an impact outside of the academic realm.

2.5.1 Research questions

Aiming at achieving the presented study objectives, the global research question addressed in this thesis can be formulated as:

- How can instructional alignment be attained for educational robotics activities?

With the goal of addressing this global research question, we will also encompass different sub-questions, each studied in one of three main stages of this thesis:

- How can instructional alignment in the context of ER be conceptualized?
- How can developers align the designs of their ER tools?
- How can educators align their instructional activities in the context of ER?

3 An alignment model for Educational Robotics Learning Systems

Disclaimer
The content of this chapter has been adapted from the following works - with permission of all co-authors and publishers:

- Giang, C., Piatti, A., & Mondada, F. (2020). Aligning the design of educational robotics tools with classroom needs. *Manuscript submitted for publication.*

  As the main author of this publication, my contribution to this work involved: conceptualization, methodology, formal analysis, investigation, data curation, visualization and writing - original draft preparation.

3.1 Summary

In this chapter, we will introduce the notion of Educational Robotics Learning Systems (ERLS). It represents the conceptual framework that can guide developers and educators in attaining instructional alignment for ER activities. However, before addressing the question of alignment, we will begin this chapter by defining the meaning of ER activities as used in this thesis. Based on this definition, we will conceptualize ER activities and determine the different elements that are involved in the activities. Finally, a pedagogically meaningful alignment of these elements is discussed, while considering learning theories relevant to ER. The proposed ERLS model represents the centerpiece of this dissertation and it will be used as a reference framework for the studies presented in the subsequent chapters.

3.2 Conceptualizing ER activities

Before discussing instructional alignment of ER, we would like to start this chapter by defining the term *educational robotics* in the context of this dissertation. As illustrated by Angel-Fernandez and Vincze (2018), the number of articles containing the search query "educational robotics" has steadily increased over the past years. Yet, according to the authors, the definition of the field has remained vague. In another recent study, Scaradozzi et al. (2019) have argued that until today, there is little agreement among researchers on which features should be considered essential for ER. Different works have considered different criteria when studying ER. For instance, in the review of Benitti (2012), activities involving robots to teach robotics were excluded, while the review of Jung and Won (2018) included them. On the other hand, Jung and Won (2018) did not consider social robots, which in contrast, were included in the work of Mubin et al. (2013). In order to appropriately address instructional alignment of ER, it is therefore necessary to first specify the nature of the ER activities that are considered in this thesis.

3.2.1 ER activities in the context of this thesis

One way to classify ER activities, is to consider the role that the robots take in the activities. As suggested by Mubin et al. (2013), there are three main roles that robots can assume in learning activities: they either can act as tutors, peers or tools. When they are used as tutors or peers, they can be classified as social robots, delivering the learning experience through social interaction with the learners (Belpaeme et al., 2018). Research on the design of such robots usually falls within the domain of Human-Robot Interaction (HRI), or in the specific case of child education, within the field of Child-Robot Interaction (CRI/cHRI).

For the remainder of this dissertation, however, we will focus on ER activities that involve robots in the latter suggested role, i.e., the use of robots as learning tools. When used in this context, the robots do not primarily focus on social interactions with the learners. They, instead, represent pedagogical instruments that are mainly used as cognitive artifacts, i.e., “human-made physical objects that functionally contribute to performing a cognitive task” (Heersmink, 2013). In fact, such “objects-to-think-with”, represent a key element of Papert’s constructionism (Papert, 1980), which has been considered the historical origin of ER. It is within this paradigm, that we will discuss instructional alignment of ER.

Apart from the different roles that robots may take in ER activities, Mubin et al. (2013) also discussed the different subject domains for which ER activities can be implemented. According to the authors, there are two main domains for which robots are used in education. On the one hand, there is the technical domain, involving topics such as robotics and computer education. On the other hand, robots can also be used for non-technical education, involving subjects such as sciences or languages. A similar classification was used by Jung and Won (2018), who distinguished between ER to teach robotics and ER to teach other subjects. For the studies performed in this dissertation, the domain in which the ER activities are used has rather secondary importance. Although most of the studies conducted in the context of this dissertation involve ER activities with a primary focus on technical education, the theories and findings may equally apply to non-technical education activities, as long as the robot is used as a tool. As a matter of fact, most ER activities have been considered to be highly multi-disciplinary either way, often incorporating technical as well as social topics (Anwar et al., 2019). Finally, another domain for which ER activities have increasingly been implemented, is the development of computational thinking (CT) competencies. Popularized by Wing (2006), CT describes a set of problem-solving skills, that in the past years have gained more and more attention in education. For instance, Shute et al. (2017) have referred to CT as "the conceptual foundation required to solve problems effectively and efficiently (i.e., algorithmically, with or without the assistance of computers) with solutions that are reusable in different contexts". ER activities implemented to foster the development of CT competencies therefore represent another application, which is highly multi-disciplinary and which leverages robots as tools to support cognition. Later on in this thesis, we will present an example of an ER activity specifically designed for this purpose.

A final aspect that should be considered, is the location in which the ER activities take place. As classified by Mubin et al. (2013), ER activities can be either intra- or extra curricular. While the former refers to the use of ER in formal education as part of the school curriculum, the latter refers to informal settings, such as after-school activities, robotics clubs or workshops. This distinction is important, since teaching and learning practices in both environments may differ (Anwar et al., 2019). In this dissertation, we will focus our discussions on ER activities in the context of formal education. As compared to informal settings, the use of ER in formal settings poses specific challenges particularly requiring frameworks to facilitate the instructional alignment of ER. Such challenges may for instance, include teachers' lack of technical pedagogical knowledge (TPK) with respect to ER, as well as practical questions such as the infrastructure, the availability of materials and above all, the compatibility with existing school curricula. Such factors usually tend to be less of a problem in informal settings, where the whole purpose of the instruction is oriented at ER-related topics and the activities are hence inherently aligned. However, the increased interest of educational institutions and policy makers to integrate ER in classroom teaching, calls for frameworks that could specifically support instructional alignment of ER in formal education.

3.2.2 ER activities as situated cognitive systems

Now that we have specified that ER activities, in the context of this thesis, refer to activities in formal education where robots are used as cognitive artifacts to teach both technical and non-technical topics, we can begin to approach the question of alignment for such activities.

In their highly cited work from the late eighties, Brown et al. (1989) have argued that knowledge is *situated*, "being in part a product of the activity, context, and culture in which it is developed and used". Consequently, design and implementation of ER activities, as an approach for knowledge development, should also adopt a situated perspective on learning. Indeed, Heft (1989) has emphasized that also affordances are situated, i.e., they exist in a specific context between the agent and the environment. Based on this notion, Osborne (2014) has argued that in the context of digital educational technology, the situated perspective of learning should not only include the real environment but also the digital one.

Along these lines and as an approach to taxonomize cognitive artifacts, Heersmink (2013) has introduced the notion of *situated cognitive systems*. In his work, Heersmink presented three possible configurations for such systems: they can be composed by only human agents, by human agents and non-artifactual objects, or by human agents and cognitive artifacts. In fact, the ER activities considered in this thesis can be characterized as the latter: in such activities, human agents are called to solve cognitive tasks, using cognitive artifacts to support them.

As already mentioned earlier, one type of the artifacts used in ER activities are obviously the robots. However, many times, ER activities also involve other objects supporting the human agents in performing the task: apart from the robots, they also often involve programming or interaction interfaces and playgrounds for task accomplishment. According to the taxonomy of Heersmink (2013), each of these objects can be classified as a specific type of cognitive artifact.

The robots in this system can be classified as *structural ecological* artifacts. This kind of cognitive artifacts embody a manipulable physical or virtual structure that can help performing a cognitive task by providing new information during or after the manipulation. Heersmink presented the example of the Tetris blocks, that can be rearranged or rotated, providing the player with new information to continue the game. Another example given by Heersmink are Scrabble tiles, that can be constantly reorganized to spot new words. Likewise, in ER activities where robots are used as tools, the states and behaviors displayed by the robot provides the learner with new information to progress with the cognitive task.

The second cognitive artifact is the programming or interaction interface, a *symbolical representational* artifact. Such artifacts are symbolical representations of other objects, which can take arbitrary forms, acquiring their meaning through logical rules and conventions. Examples for symbolical representational artifacts are algebras or alphabets. Likewise, in ER activities, programming/interaction interfaces are a symbolical representation of the robots' properties. Sometimes, the interfaces may also embody iconical representations (for example iconical programming blocks). In these cases, the interfaces may be considered a combination of iconical and symbolical representational artifacts.

Finally, there are the playgrounds used for task completion, representing *spatial ecological* artifacts. Such artifacts store information in their spatial structure, that can be used to support the completion of the cognitive task. The playground is an essential element of many ER activities, since it usually embodies a representation of the cognitive task (e.g. when a robot has to be controlled to move from one position to another). Previous work has emphasized that it is important for ER activities to provide large spaces for the pupils to work, allowing them to "play around, and test different kind of solutions for each kind of project they face". (Lindh & Holgersson, 2007). Indeed, in many cases, it is through the combination of the robot and playground that the learner can evaluate if the cognitive task was completed successfully.

As reasoned by Heersmink, identifying the components of situated cognitive systems is an important step to better understand how to conceptualize their contributions for cognitive performance. Nevertheless, in line with Biggs' view of "teaching and learning as a system", he also highlighted, that ultimately, situated cognitive systems should be studied as a whole. If we are aiming at enhancing instructional alignment of ER, it is therefore not enough to only consider how the robots can fit in classroom activities. We have to look at it from a situated perspective that considers the entirety of the cognitive artifacts and their affordances and understand how they can be used in teaching and learning activities in classrooms. In the next section we will therefore introduce the notion of Educational Robotics Learning Systems (ERLS), representing a conceptual framework for the instructional alignment of ER.

3.3 Introducing Educational Robotics Learning Systems (ERLS)

While the extended alignment model devised by Lauwers (2010) represents a good point of departure, this section will suggest further modifications, to make it more appropriate for the type of ER activities discussed in this thesis. The general structure of Lauwers' model (see Fig. 2.4) allows for a systematic and situated view on ER activities, since it accounts for the interplay of the different instructional components. However, its characterization of the fourth component as the "tool", appears to be too limiting. Indeed, what defines ER activities, is not only the insertion of a tool to the learning system. First of all, as outlined in the previous section, ER activities involve a variety of different elements (i.e., the robots, the interfaces and the playgrounds). Furthermore, as described by Heersmink's taxonomy, each of these elements embodies a specific cognition-aiding property and hence provides different affordances. To emphasize this more holistic view on ER activities, we therefore introduce the notion of Educational Robotics Learning Systems (ERLS), representing an conceptual framework for the instructional alignment of ER (Fig. 3.1). The model was devised by appropriately adapting the alignment model proposed by Lauwers: the first three components of the ERLS model (i.e., intended outcomes, instruction and assessment), remain mostly unchanged and are therefore, coherent with the traditional model for instructional alignment (see Fig. 2.1). The last component, however, has been redefined as "ER artifacts" to account for the entirety of the cognitive artifacts and their distinct affordances. Moreover, we decided to use the original notion of "intended outcomes" presented by Cohen (1987) rather than "learning goals" as presented by Lauwers. This choice was motivated by the fact that ER activities are sometimes also used to attain objectives that are broader than learning goals as defined by the taxonomy of Bloom et al. (1956). For instance, one goal of ER activities is often to elicit students' interest and excitement about the studied topics. While this may not be considered a learning goal in the traditional sense, the term "intended outcomes" allows to capture this objective.

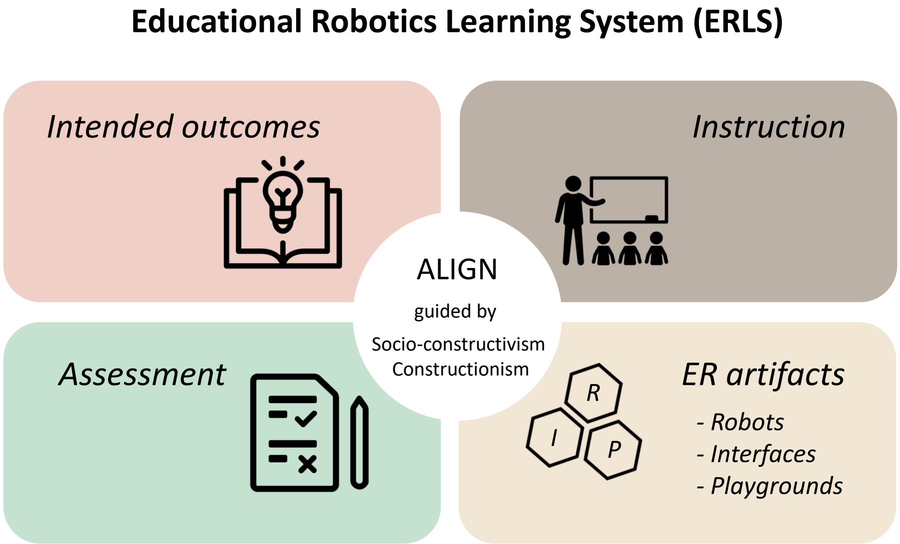

Figure 3.1 – Alignment model for Educational Robotics Learning Systems (ERLS).

In his work, Lauwers further discussed participatory and design-based research as possible approaches to attain alignment. Though the proposed methods appeared to be suitable to create well-aligned designs, the work missed to further elaborate on the learning theories that should guide the alignment. However, in order to enhance the relevance of the alignment model, appropriate learning theories need to be identified, in order to provide appropriate guidance for the design of ER tools and ER activities.

One main learning theory underlying ER activities is *constructionism*, introduced by Seymour Papert (1980), a former student of Jean Piaget. Papert's constructionism can, in a sense, be considered an extension of Piaget's *constructivism* (Piaget & Cook, 1954), that among other things, has also strongly influenced Biggs' idea of constructive alignment. Piaget argued that knowledge is constructed by the learners themselves rather than being imparted from the teachers to learners. Knowledge is built by students through active and personal learning processes, that help them creating the links between the new topics they are learning and their prior knowledge and experiences. Building on the theory of constructivism, Papert added the idea that constructivist approaches can be even more favorable, when the learner is "consciously engaged in constructing a public entity, whether it's a sand castle on the beach or a theory of the universe" (Papert & Harel, 1991). The inclusion of such public entities, assuming the role of cognitive artifacts, is one of the key ideas of Papert's constructionism. To promote constructionist learning approaches, Papert advocated for the use of educational technology in classrooms. Indeed, his use of the Logo turtle for the exploration of programming concepts, can, so to say, be considered the origin of today's ER. It is no coincidence, that one of the most used ER kits worldwide, the Lego *Mindstorms* series, still bears the name of Papert's influential book from the early eighties (Papert, 1980).

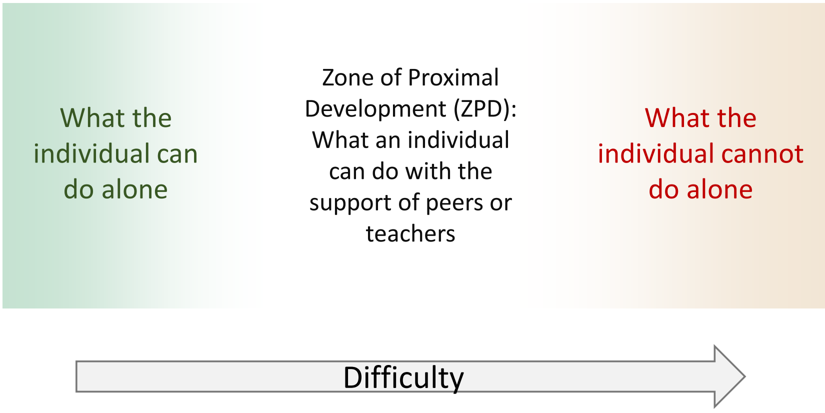

Figure 3.2 – Illustration of Vygotsky's Zone of Proximal Development (ZPD).

When ER activities are implemented in classrooms, students in most cases, collaborate in small groups to solve the tasks. The social aspects of the activities, represent another facet of the situated learning with ER. In this regard, we also have to consider another variant of the constructivist learning theory, which is the socio-constructivist theory. Socio-contructivism emphasizes the impact of social factors in learning, for instance by collaboration, discussion and negotiation between learners. The origins of socio-constructivism can be traced back to the works of Vygotsky, who argued that knowledge is co-constructed (Vygotsky, 1964, 1980). A key element of his theories is the Zone of Proximal Development (ZPD). It represents the range of tasks that are too difficult for an individual to master alone, but however, can be mastered if the individual is supported by a more-skilled peer or assisted by a teacher (Fig. 3.2). In the latter case, another concept closely related to Vygotsky's ZPD becomes important, namely the theory of scaffolding. First introduced in the context of children's language development by Wood et al. (1976), scaffolding describes the idea of providing students with the appropriate amount of support with an appropriate timing, so that it creates favorable learning experiences. More specifically, Bruner (1978) described scaffolding as "the steps taken to reduce the degrees of freedom in carrying out some task so that the child can concentrate on the difficult skill she is in the process of acquiring". As a matter of fact, many ER activities build on the concepts related to the ideas of socio-contructivism, such as Vygotsky's ZPD or scaffolding. We, therefore, argue that the theory of socio-constructivism should also be taken into consideration when designing ER tools and ER activities.

Both learning theories, constructionism and socio-constructivism, are fundamental pillars of the ER activities discussed in this thesis and thus need to be considered for their instructional alignment. Similar to the way that Biggs (2003) incorporated constructivism as a guiding principle in his alignment model, we here propose that the learning theories of constructionism and socio-constructivism should be considered as guiding principles at all steps of alignment for ERLS.

Using the ERLS alignment model as a guiding framework, educators may take a more situated perspective on ER activities, helping them to design and implement instructional activities, that are better aligned with the learning goals, the affordances of the used ER artifacts and the final assessment activities. On the other hand, developers may use the framework to make informed design decisions, helping them to better align their ER tools with classroom needs. Either way, it is important that both educators and developers, consider the entirety of the ERLS and take a situated learning perspective, when they design their tools and activities, respectively. In order to fully leverage the potential of ER activities, the designs should in particular embody learning theories underlying ER, such as constructionism and socio-constructivism.

Based on the devised ERLS framework, the next section will describe a backwards design approach, aimed at further supporting developers and educators in putting the main ideas of the alignment model into practice. It describes the different steps that should be taken to achieve well-aligned designs of ER tools and ER classroom activities.

3.4 Backwards design as an approach for alignment

As presented by Biggs (2003), a backwards design approach, i.e., the idea to start with the definition of the desired learning outcomes and then choose appropriate instructional activities and assessment methods, can support educators in the creation of well-aligned instruction. The approach has been adopted to the field of e-learning by Bower (2008), who extended the task design by an affordance analysis of the educational technology (see Section 2.3.2). Lauwers (2010) on the other hand, demonstrated that backwards design can also be helpful for developers, supporting them in the design of educational technologies. Here we will adopt a similar approach for the alignment of ERLS. As presented in the previous section, both developers and educators should consider the entirety of the ERLS framework, when they design their tools and activities, respectively. In the following we will therefore describe a procedure, incorporating this approach, that may be applicable by both developers and educators alike. While developers could follow the procedure for the design of ER tools, it might be equally useful for educators, who are designing and implementing ER classroom activities:

1. The design starts with a clear definition of the intended outcomes.
2. Based on the defined outcomes, developers design appropriate ER tools and educators design appropriate instructional activities. In either case, the design process should consider:
    (a) Which kind of tasks can lead to reaching the defined outcomes and how they do it
    (b) Which ER artifacts (i.e., robots, interfaces and playgrounds) can effectively support the activity and how they do it
    (c) Which affordances the ER artifacts provide and how they can effectively be leveraged to support the activity

(d) Which roles the students assume in the activities and how they interact with each other

(e) Which roles the educators assume in the activities and how they interact with the students

3. Finally, appropriate assessment methods, both formative and summative, should be identified to evaluate how the actual outcomes match with the intended outcomes.

While the above presented procedure may seem more intuitive for educators to apply, we argue that it may be considered equally important for developers designing new ER tools, especially the first two steps. In order to design well-aligned ER tools, it is necessary that developers are aware of the outcomes educators want to achieve using the tools, as well as the instructional activities foreseen to reach them and eventually even the assessment methods to evaluate them. ER tools, which are inherently aligned with these instructional components by design, may facilitate the final use by educators and students, and could therefore be educationally more meaningful.

Nevertheless, it has to be acknowledged, that reflecting on learning outcomes, appropriate instructional activities and assessment methods may, in most cases, not be part of the typical duties of technology developers and they may therefore, represent a particular challenge. Indeed, Antonenko et al. (2017) have pointed out, that a major limitation of their alignment framework for educational technologies is that it might not be usable by "technology designers and developers with little or no background in psychology and education". On the other hand, alignment has to be also addressed by educators who finally use the devised ER tools to design and implement classroom activities. While most educators should be familiar with the ideas of traditional instructional alignment (i.e., the alignment of intended outcomes, instructional and assessment activities), the integration of ER tools may pose new challenges. ER is still a novel field to many educators and only few have already developed the technological pedagogical knowledge (TPK) required to understand the affordances of ER artifacts and hence know how they can effectively be leveraged for classroom activities.

While it is possible to overcome these issues by applying participatory and design-based research methods as proposed by Lauwers (2010), such approaches are often time-consuming and not always feasible. As a matter of fact, not many developers have the possibilities to involve teachers and students in co-design projects. Moreover, having access to classrooms for testing can at times be difficult due to strict regulations. Other groups, such as the Thymio project (Mondada et al., 2017), have followed an open-source approach to actively involve educational researchers, engineers, teachers and designers in the development process of the tools and activities. Though this approach has shown to be effective for the development of an educationally valuable ER platform, it is rather suitable for projects in more advanced stages, as successfully building a community around a product requires significant amounts of effort and time.

To support educators and developers in addressing the above-mentioned challenges in a time- and cost-efficient way, the next chapter will introduce a methodology to further facilitate the backwards design procedure, particularly easing the second step of the proposed approach. To this end, we will present the development of heuristics for the design and evaluation of ERLS. While incorporating the main ideas presented in this chapter, the heuristics may represent a more operational and straight-forward complement to the alignment model, that can be applied in an easy, intuitive and time-efficient manner, by both educators and developers. The methodology might especially be helpful for developers with little background in education as well as for educators, who are not very experienced with ER.

4 Heuristics for the design and evaluation of ERLS

Disclaimer
The content of this chapter has been adapted from the following works - with permission of all co-authors and publishers:

- Giang, C., Piatti, A., & Mondada, F. (2019). Heuristics for the development and evaluation of educational robotics systems. *IEEE Transactions on Education, 62*(4), 278–287. https://doi.org/10.1109/TE.2019.2912351

  As the main author of this publication, my contribution to this work involved: conceptualization, methodology, formal analysis, investigation, data curation, visualization and writing - original draft preparation.

- Giang, C., Piatti, A., & Mondada, F. (2020). Aligning the design of educational robotics tools with classroom needs. *Manuscript submitted for publication.*

  As the main author of this publication, my contribution to this work involved: conceptualization, methodology, formal analysis, investigation, data curation, visualization and writing - original draft preparation.

4.1 Summary

In this chapter, we will introduce the heuristics for the design and evaluation of educational robotics learning systems (HERLS). The main objective of developing the heuristics was to provide an operational tool that could complement the ERLS alignment framework and help developers and educators put the main ideas of the framework into practice. The HERLS heuristics were devised in multiple iterations. First, a preliminary set of heuristics was compiled based on existing literature and presented to primary and lower secondary school teachers to explore the use of heuristic evaluation in the context of ERLS. In a second step, focus groups with teachers, developers and educational researchers were organized to revise

and validate the heuristics. The revised heuristics were then presented to two groups of developers, to assess their usability and utility in authentic development settings. This chapter will finally conclude with a discussion on the results obtained from these studies.

The HERLS heuristics are aimed at supporting both developers designing ER tools as well as educators designing and implementing ER classroom activities. In this and the following chapter, we will first focus our discussions on the former case. An illustration of how the HERLS heuristics can also be applied to guide the design of ER activities will then be presented in Chapter 6.

4.2 Heuristic evaluation

Heuristic Evaluation (HE), introduced by Nielsen and Molich (1990), is a time- and cost-efficient methodology to analyze the usability of digital user interfaces, that has been shown to be particularly useful in early design stages. HE is usually performed by experts in the domain (but sometimes also by non-experts), who evaluate a system by checking its compliance with a given set of design principles (i.e., the heuristics). The heuristics describe the characteristics of an ideal system and are often formulated as rule of thumbs such as *"minimize user memory load"* or *"provide good error messages"*. By performing HE, the evaluators identify usability issues, i.e., properties of the system that do not comply with the heuristics. Representing the weak spots of the system, the usability issues can then be fixed by the developers in the next iteration of the development cycle.

While the works of Nielsen and Molich mainly focused on the usability of digital user interfaces, their approach has been adapted for different target domains. For instance, previous works have presented heuristics adapted to games (Desurvire et al., 2004; Köffel & Haller, 2008; Sweetser & Wyeth, 2005).

In their study, Desurvire et al. (2004) have developed heuristics to evaluate the playability of video, computer, and board games and tested their efficacy against traditional user study methods. The authors concluded that, HE can be "very useful for creating a highly usable and playable game design, particularly in the preliminary design phase before expensive prototypes".

Also aiming at the evaluation of games, Sweetser and Wyeth (2005) developed the GameFlow model. The authors presented eight heuristics to evaluate enjoyment in games based on the concept of flow. Experts applied the heuristics to evaluate two real-time strategy video games and the results showed that the assessment based on heuristic evaluation appeared to be coherent with evaluations from professional reviews.

Köffel and Haller (2008) adapted the approach to develop a set of ten heuristics for augmented tabletop games. The heuristics incorporated aspects related to the virtual interface, the game play, the story line as well as the particularities of tabletop games. The authors revised the heuristics in multiple iterations and suggested that the final version may provide support for the evaluation of tabletop games.

Clarkson and Arkin (2007) adapted HE to the field of human-robot interaction (HRI) by developing a list of eight HRI heuristics. In a study with graduate students, the heuristics were tested with a robot that was teleoperated using a joystick controller. Based on their results, the authors suggested that HE can be considered useful in "guiding the formative stages of system development via early, iterative applications".

Likewise, several attempts have been made to leverage HE as a method to guide the development of educational technology. As presented in Section 2.3.2, the work of P. Kirschner et al.

(2004) has associated usability with technologicial affordances, while utility was related to educational affordances. While HE is traditionally concerned with a system's usability, several works have shown that it is also possible to integrate heuristics related to their utility. Previous works have presented heuristics specifically compiled for the domains of educational games (Barbosa et al., 2015; Jerzak & Rebelo, 2014; Malone, 1981; Mohamed & Jaafar, 2010) as well as novice programming systems (Kölling & McKay, 2016; Pane & Myers, 1996). These sets aim at covering more aspects than pure usability by integrating heuristics related to pedagogy and learning. Indeed, tailoring heuristics to a specific target domain appeared to be more effective: in their study, Kölling and McKay (2016) asked half of the participants to perform a HE for a novice programming system using Nielsen and Molich's set, while the other half used their tailored heuristics. The study showed that the issues identified using the latter appeared to be more diverse and included more items that were not identified with Nielsen and Molich's set than vice versa.

The findings from the existing literature suggest that HE may also be a suitable approach to guide the design and evaluation of ERLS, especially in the early stages of the development. In fact, the approach might be interesting for both developers who design ER tools, as well as educators designing and implementing ER activities. In the following, we will begin with a focus on the former case. As a matter of fact, ER tools incorporate many elements that are also found in video and tabletop games as well as in other types of educational technology. Hereafter, we will therefore present the iterative development of design heuristics specifically compiled for ERLS. While the first preliminary set was completely based on existing literature, the revision of the heuristics involved the input given by teachers, developers and educational researchers. As emphasized by Kölling and McKay (2016), the outcomes of HE heavily rely on the evaluators' findings, which in turn are strongly driven by the heuristics provided. It is, therefore, essential that any heuristics are tested for validity prior to the use in HE. To this end, focus groups were performed to revise the devised heuristics and to demonstrate their utility in authentic development settings.

4.3 Exploring heuristic evaluation for ER tools

In a first study, the suitability of HE as a method to assess the properties of ER tools was explored with twelve ER-experienced teachers. Therefore, an initial set of heuristics was compiled, consisting of fourteen categories adapted from existing literature on educational robotics, learning games, as well as augmented and classical tabletop games (Table 4.1). Moreover, a particular emphasis was put on the integration of the ideas presented in the ERLS framework. Indeed, ERLS incorporate many elements that are essential to classical tabletop games and digital video games, such as interaction, enjoyment and challenge. Many of these elements have been also considered core components for the design of learning games. These commonalities suggest that ERLS may be considered a kind of *educational augmented tabletop game* - a combined entity of tangible tabletop games and digital learning games. The initial set of heuristics therefore comprised categories adapted from both game- and education-related literature. Finally, to emphasize the idea that developers and educators should always consider the entirety of the ERLS, each heuristic was formulated with respect to the system, instead of focusing on each specific component of the ERLS.

4.3.1 Study design

The first set of heuristics was evaluated through experiments with twelve compulsory school teachers, who participated in a testing session, in which they tested five different ER tools. During the testing session, they were asked to identify usability and utility issues for each system. These issues were then mapped to the heuristics in order to illustrate the applicability, completeness and orthogonality of the heuristics. Furthermore, based on the issues determined by the teachers, the systems were ranked regarding their suitability for classroom use. The results were then compared to two rankings obtained from a questionnaire distributed at the end of the testing session: one ranking was based on the intuitive choices of the teachers, while the other concerned system characteristics and technical features (e.g., type of sensors/actuators, connection method etc.). The latter was based on a list of characteristics that was determined using the teachers' answers given in the questionnaire. Finally, the questionnaire also provided information about the acceptance of the heuristics by the teachers and their satisfaction with the testing procedure.

Participants

To validate the coherence of the selected heuristics with the expectations of compulsory school teachers, a heterogeneous group of twelve teachers (different gender, age, school level and professional background) was invited to participate in a user study (Table 4.2). At the time of the study, all teachers were in service and enrolled in a two-year training program for a Certificate of Advanced Studies (CAS) in Educational Robotics. This CAS is the first of its kind in Switzerland, and participants are considered to be so-called early adopters, taking a pioneering role among their peers. The course is lectured by an interdisciplinary team of engineers, educators and educational researchers knowledgeable in the field of ER.

Assessed ER tools

Five ER tools were assessed by the participating teachers during the testing session. All tools were promoted for educational purposes by their manufacturers and none of them were known to the teachers before the study. Aiming at validating the heuristics for a wide range of ER tools, the selection was based on the characteristics presented in Table 4.3. The goal was to include a selection of systems comprising a great variety of the shown characteristics, in order to achieve a diverse representation of ER tools. It is important to note that most of the information given in Table 4.3 refer to one possible configuration of the respective system. For instance, some systems allow the use of various programming interfaces and

Table 4.1 – First set of heuristics for ERLS presented in Giang et al. (2019)

No.	Description	Based on	Game	Edu
1	*Cognitive workload*: The system allows the user to maintain their sense of cognitive flow. Cognitive workload which is not related to the learning activities should be minimized.	Boller and Kapp, 2017; Köffel and Haller, 2008; Lee et al., 2011	X	X
2	*Challenge*: The system presents appropriate challenges tailored to the user. It should be "easy to learn, but hard to master". The user's fatigue is minimized by varying activities and pacing during the learning activities.	Barbosa et al., 2015; Boller and Kapp, 2017; Köffel and Haller, 2008	X	X
3	*Adaptability*: The system should be adaptable to the needs of the user. The system should be usable by all users of the target group regardless of their prior knowledge.	Barbosa et al., 2015; Köffel and Haller, 2008	X	X
4	*Interaction*: The interaction method should satisfy the expectations of the user and follow the logic of the learning activities. The user interfaces should be compliant with industry standards and be usable in a very natural, easy and understandable way.	Boller and Kapp, 2017; Köffel and Haller, 2008	X	
5	*Level of automation*: The user should be able to execute all actions relevant to the learning activities by him/herself. All actions that are perceived as boring and rather unimportant to the learning activities should be performed by the system.	Köffel and Haller, 2008	X	
6	*Collaboration and communication*: The entirety of the system should support interpersonal communication, collaboration and, if appropriate, competitiveness between users.	Boller and Kapp, 2017; Köffel and Haller, 2008	X	X
7	*Feedback*: The system should provide visual, acoustic or haptic feedback to help the user understand their performed actions and the resulting consequences.	Benitti, 2012; Boller and Kapp, 2017; Köffel and Haller, 2008	X	X
8	*Comfort of the physical setup*: The physical setup should be fast and easy to assemble, comfortable to use and not require the user to take an awkward position.	Köffel and Haller, 2008	X	
9	*Enjoyment and aesthetics*: The user should find the activities fun. The entirety of the system should be inviting and aesthetically appealing. It should quickly grab the user's attention and facilitate the user's concentration and immersion in the activities.	Barbosa et al., 2015; Boller and Kapp, 2017; Sweetser and Wyeth, 2005	X	
10	*Transparency*: The system should provide a rich and open environment, allowing the inspection of all underlying mechanisms.	Benitti, 2012; Lee et al., 2011		X
11	*Active learning*: The system encourages exploration, problem solving and enquiry. The user should feel safe in the knowledge that they can experiment without breaking the system.	Barbosa et al., 2015; Kölling and McKay, 2016		X
12	*Relevance*: The learning activities should be personally relevant to the user and allow him/her to relate the activities to the learning goals.	Barbosa et al., 2015; Benitti, 2012; Boller and Kapp, 2017		X
13	*Supports reflection*: The system should provide opportunities for reflection and debriefing on learning and highlight the process of learning to the user.	Barbosa et al., 2015; Boller and Kapp, 2017		X
14	*Computational thinking*: The entirety of the system should support the development of computational thinking competences.	Atmatzidou and Demetriadis, 2016; Mannila et al., 2014		X

Table 4.2 – Teachers participating in the study presented in Giang et al. (2019)

Teacher	Gender	Age	School level	Background
T1	Male	30-50	Primary	Pedagogy
T2	Male	30-50	Primary	Psychology
T3	Male	> 50	Primary	Pedagogy
T4	Female	30-50	Primary	Pedagogy
T5	Female	> 50	Primary	Psychology
T6	Female	> 50	Primary	Pedagogy
T7	Female	> 50	Primary	Pedagogy
T8	Male	< 30	Lower secondary	Electrical engineering
T9	Female	< 30	Lower secondary	Mathematics
T10	Male	30-50	Lower secondary	Aeronautical engineering
T11	Male	< 30	Lower secondary	Mathematics
T12	Male	30-50	Lower secondary	Electrical engineering

user devices or can be extended with additional sensors and/or actuators. For this study, all systems were presented using their standard available components. Moreover, the interaction methods and proposed tasks were chosen by the experimenters, while putting emphasis on the particularities of each system. Since the aim of this study was the evaluation of the suitability of HE for ER tools rather than the assessment of the selected ER tools, not presenting all possible configurations to the teachers is justifiably not a major issue.

Study protocol

At the beginning of the session, all participants were introduced to the heuristics, ensuring that there was no lack of clarity about the meaning of each heuristic. For the testing session, the teachers formed five groups of 2-3 people, a recommended group size for ER activities (Benitti, 2012). Each group had 30 minutes to test a system by working through a worksheet prepared by the experimenters. Each test was introduced by a short video made by the manufacturer highlighting the features and functionalities of the presented ER tool. This was followed by a user tutorial, showing the group how to get started with the system. Finally, the group was given time to further explore the system by either following proposed activity suggestions or by following their own ideas and interests. On a rotational basis, each group tested all five systems and was asked to identify as many usability issues as possible for each system while testing. A schematic overview of the study procedure is depicted in Fig. 4.1. A printout with the set of heuristics was available on all tables. However, the teachers were not obliged to use them, and they could identify issues not related to the heuristics. Moreover, the teachers were instructed to put emphasis on identifying usability issues related to the possible use of the ER tools in classrooms, which also includes difficulties that their students might encounter when working with the presented ER tool. Throughout the tests, each group was followed by one experimenter, who supervised the heuristic evaluation by taking the role of an

Table 4.3 – ER tools assessed in the study presented in Giang et al. (2019)

Characteristic	**Anki Cozmo**	**Calliope mini**	**Lego WeDo 2.0**	**Makeblock mBot**	**Ozobot Evo**
System type	Wheeled / social robot	Electronic kit	Construction kit	Wheeled robot	Wheeled / social robot
Programming	Graphical	Graphical	Graphical	Graphical	Tangible
User device	Tablet	PC	Tablet	Tablet	Pen and paper
Assembly	No	No	Yes	Yes	No
Extendable	No	Yes	No	Yes	No
Sensors	Camera with AI, ground sensors	Touch, buttons, light, temperature, gyroscope, compass, microphone	Distance, tilt	Distance, ground, light, button	Distance, RGB, ground
Actuators	Motor, loudspeaker, color LED, screen, mechanical arm	Loudspeaker, color LED, 5x5 LED matrix	Motor	Motor, loudspeaker, LEDs	Motor, loudspeaker, LEDs

"observer" (Nielsen, 1995a). The main function of the observers was to record the usability issues determined by the teachers, using written reports for later analysis. Furthermore, they provided technical support in order to facilitate the testing procedure.

Weighting and mapping of usability issues

The written reports obtained from the observers provided a summary of all usability issues for each ER tool identified by each group. However, since some issues may have a stronger impact on the user experience than others, it was not sufficient to only consider the quantity of the identified issues for each system. Instead, the teachers were also asked to assign a weight to each issue, in order to obtain a meaningful weighting among the issues. The weights were based on the Nielsen severity scale (Köffel & Haller, 2008; Nielsen, 1995b) for usability issues:

- 0 - Not a usability problem at all.
- 1 - Cosmetic problem only: It does not have a profound impact on the activity.
- 2 - Minor problem: It has a slight impact on the activity and influences the experience a bit.
- 3 - Major problem: This problem has a severe impact on the activity and negatively influences the user experience.
- 4 - Usability catastrophe: This problem has to be fixed in order to allow for a decent user experience.

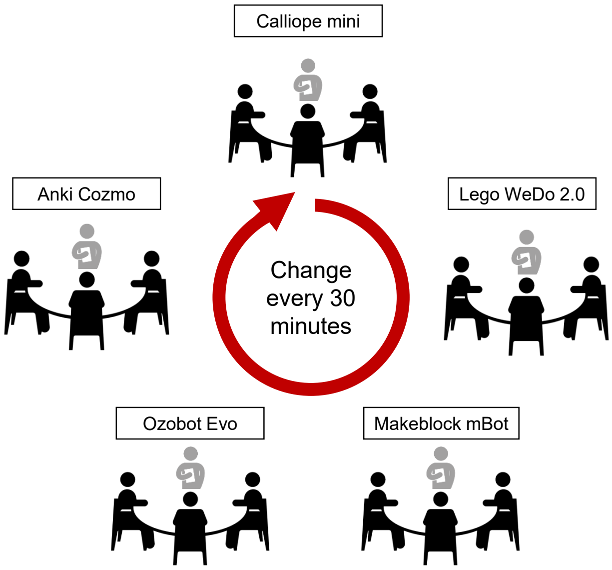

Figure 4.1 – Schematic overview of the procedure for the first study. Each group of 2-3 teachers (black) was followed by one experimenter (gray). On a rotational basis, each group tested each ER tool.

Assuming that different groups would identify different usability issues (Nielsen, 1995a), all the identified issues were merged into one list for each system by the experimenters at the end of the study. Subsequently, each issue was mapped to one of the heuristics. The information about the aggregate of all issues identified for a system and the heuristics they were mapped to, was hence only available to the experimenters and not to the teachers. The final ranking among the systems was determined based on the number and the severity of the issues identified for each system. The systems were ranked by always giving a higher importance to more severe issues (e.g., one usability catastrophe is always worse than many minor problems). If two systems had the same number of issues for one severity class, the next lower class was considered.

Questionnaires

At the end of the session, all participants were asked to complete a questionnaire comprising four subsections: first, they were asked to provide an intuitive ranking of the five systems based on the following question: “For the use in my class, I would choose the systems in the following order”. This information was used by the experimenters to determine a ranking based on the teachers’ intuitive choices. In the next subsection, the teachers were asked to assess the relevance of each heuristic using a 4-point Likert scale (strongly agree / agree / disagree / strongly disagree). Moreover, they were asked to propose amendments to the heuristics, in case they thought that some criteria were missing. The answers were used to analyze the

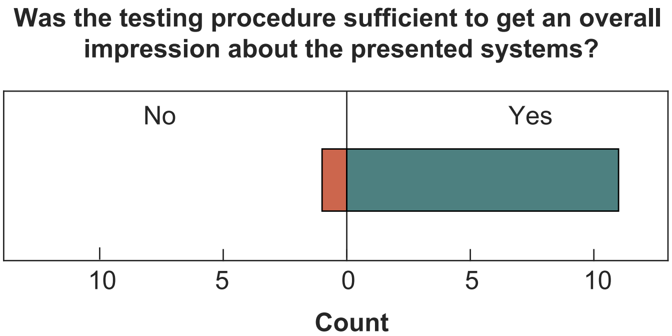

Figure 4.2 – Satisfaction with the testing procedure as reported from the questionnaire.

teachers' acceptance of the devised heuristics. Subsequently, the teachers were asked to rate the importance of different system characteristics of ER tools (e.g., type of sensors/actuators, connection method etc.) using the same 4-point Likert scale. Additionally, they were given the possibility to indicate features and components that were not listed. This information was used by the experimenters to determine a ranking based on system characteristics. Finally, in the last subsection of the questionnaire, the participants were asked to provide their opinion about the testing procedure.

4.3.2 Results

Since most of the results presented in this subsection rely on the assumption that all participants were able to form an opinion about the presented systems, it was important to ensure that the evaluation session allowed them to explore the systems sufficiently well. As reported in the last subsection of the questionnaire, it appears that the testing procedure indeed allowed the teachers to adequately discover the systems within the 30 minutes of testing (Fig. 4.2). A crucial point might have been the prepared worksheets, which facilitated the systematic exploration of the systems. Furthermore, it can be assumed that also the presence of the observers might have contributed to the efficacy of the testing procedure, since technical issues were resolved quickly, allowing the participants to remain focused on the main tasks.

Acceptance of heuristics

As a first means to examine the validity of the heuristics, the teachers were asked to assess the relevance of each heuristic using a 4-point Likert scale. Demonstrating a general acceptance of the devised heuristics from a heterogenous groups of compulsory school teachers would provide first evidence for their validity, since ultimately, teachers are the ones who select, adapt and create learning activities involving ER tools. The results illustrated that a large majority of the teachers agreed on the validity of the heuristics as a guiding tool for the evaluation of ER tools (Fig. 4.3). For some of the heuristics (e.g., *Level of automation* and *Comfort of physical setup*) the acceptance was lower compared to others (e.g., *Active learning*

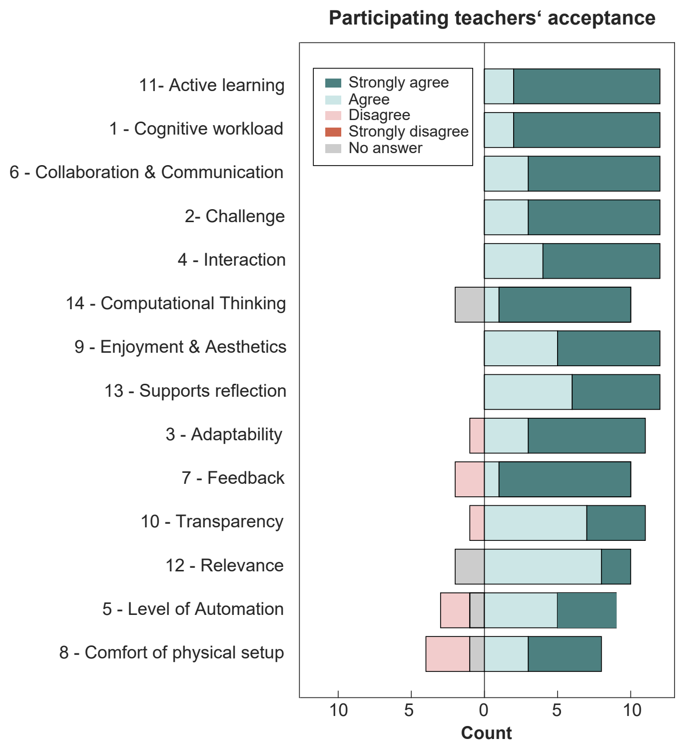

Figure 4.3 – Acceptance of each heuristic by the teachers participating in the first study.

and *Cognitive workload*). However, approval was dominating rejection for all the devised heuristics, demonstrating a general acceptance by the participating teachers. Interestingly, also heuristics based only on game-related literature (e.g., *Interaction* and *Enjoyment and aesthetics*) found wide acceptance, supporting the proposed approach to consider ERLS as a kind of "educational augmented tabletop game". Nevertheless, it should be mentioned that the two heuristics that found the least acceptance (i.e., *Level of automation* and *Comfort of physical setup*) were also based only on game-related literature. For some of the heuristics (e.g., *Computational thinking* and *Relevance*) a few participants did not respond, indicating that these heuristics may need a more precise description to prevent a lack of clarity. Finally, the teachers were also asked to propose amendments to the heuristics, in case they considered that some criteria were missing. However, none of the participating teachers suggested any amendments, indicating that this first set of heuristics comprised all the criteria the teachers considered to be important for the use of ER tools in classrooms.

Ranking based on usability issues

A total of 63 usability issues (1 usability catastrophe, 26 major, 22 minor and 14 cosmetic problems) were identified by the participating teachers for all systems (Fig. 4.4, bottom right panel). The results illustrated that almost half (31) of the identified usability issues could be associated to three heuristics: *Interaction* (12), *Adaptability* (11) and *Comfort of the* *physical setup* (8). This is not surprising, since usability issues related to these heuristics are often easily noticeable. Some of the usability issues associated to these heuristics were for example: “Programming interface not intuitive” (mapped to *Interaction*), “Limited number of actuators constrains possibilities” (*Adaptability*) or “Setting up the system takes too much time” (*Comfort of the physical setup*). Usability issues were found for all heuristics except for one (i.e., *Supports reflection*) and all the identified issues could be clearly mapped to one of the fourteen heuristics by the experimenters. Moreover, there were no issues identified by the teachers which could not be associated to a heuristic. Nevertheless, it seems that usability issues related to some of the heuristics would need more extensive testing for identification, since they could be less evident to discover. Extending the testing time could therefore reveal more usability issues for heuristics where none or only few issues have been identified (e.g. *Supports reflection*, *Level of automation* or *Feedback*), and hence provide a more exhaustive evaluation of the presented ER tool. The individual analysis for each of the five systems presented during the evaluation session revealed considerable differences (Fig. 4.4, first five panels). While most usability issues found for Makeblock mBot did not have a severe impact on the user experience (0 usability catastrophes (Ca) / 1 major problem (Ma) / 7 minor problems (Mi) / 3 cosmetic problems (Co)), the issues found for Anki Cozmo (Ca-0 / Ma-5 / Mi-0 / Co-6), Ozobot Evo (Ca-0 / Ma-5 / Mi-7 / Co-3) and Lego WeDo 2.0 (Ca-0 / Ma-8 / Mi-5 / Co-1) had a larger influence. The only usability catastrophe identified during the evaluation study was found for Calliope mini (Ca-1 / Ma-7 / Mi-3 / Co-1). The final ranking based on the identified usability issues, is presented in the second column of Table 4.5. Although the usability catastrophe found for Calliope mini (i.e., “Only German user language” (*Interaction*)) is strongly related to the non-German-speaking participants of this study, it seems like user language is a non-negligible factor for the use of ER tools in classrooms. Albeit some elements of Calliope mini were available in English, most teachers reported that they preferred a translation to their mother tongue (Italian). Under the present circumstances, the teachers believed they were not able to use this ER tools in class and hence, decided to assign the most severe weight (i.e., *usability catastrophe*) to this usability issue. As hypothesized, only few overlaps were found for the issues identified by the different groups. As typical for heuristic evaluations involving a limited number of evaluators (Nielsen, 1995a), each group identified different usability issues for each system. Nevertheless, it can be assumed that the aggregate of the usability issues identified by the five groups is a reasonable representation of the systems’ limitations, providing an extensive list of weaknesses and limitations for each system. As shown by Nielsen and Landauer (1993), five evaluators are usually sufficient to uncover almost 75% of the known usability issues of a system.

Ranking based on intuitive choices

By means of the questionnaire distributed at the end of the session, each teacher was asked to provide a personal intuitive ranking of the five presented systems regarding their usability for classroom teaching (Fig. 4.5). In order to quantify an overall result, a scoring system was introduced: a system was assigned 4 points every time it was selected as a first choice (3 for

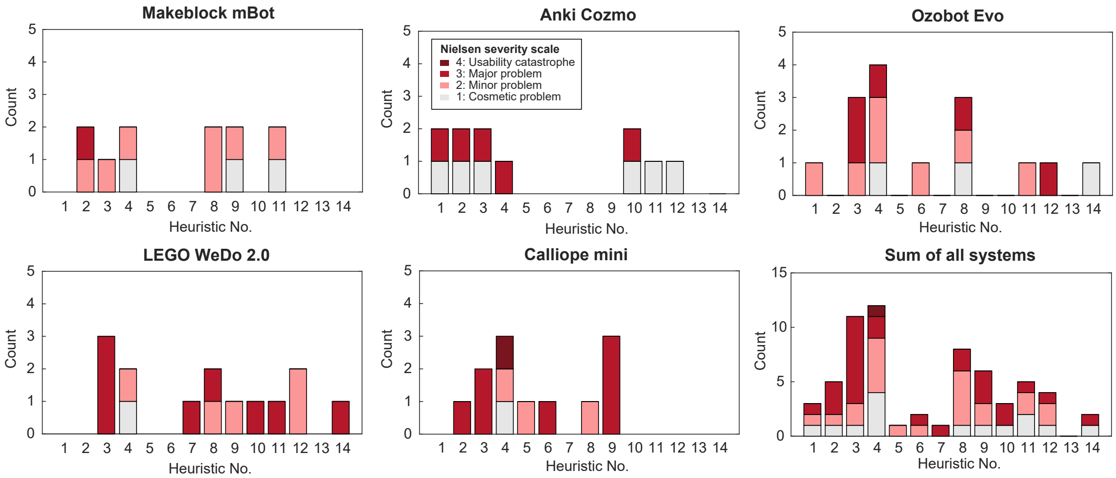

Figure 4.4 – Usability issues identified by the participating teachers during the evaluation session for each ER tool (first five panels) and for all systems together (last panel). The bars indicate the number and severity of the issues associated to each heuristic.

second, 2 for third, 1 for fourth and 0 for fifth choice). The final ranking (Table 4.5, third column) was established based on the ratio between the points obtained by each system and the maximum achievable score (48 points). The results yielded a consistent match with the ranking based on the identified usability issues. The dominant lead of Makeblock mBot in the personal ranking (85% of the maximal score) is equally reflected by the low number of severe usability issues identified by the teachers (Ca-0 and Ma-1). It is followed by Anki Cozmo (63%) and Ozobot Evo (58%), which similarly convinced the teachers. Coherently, an equal number of severe issues was found for both systems (Ca-0 and Ma-5 for both). The higher number of low severity issues for Ozobot Evo account for its lower positioning, which is in accordance with the results based on the intuitive choices. Lego WeDo 2.0 (33%) and Calliope mini (10%) ranked on the last two places of the personal ranking. A similar result was obtained from the ranking based on the usability issues, reflected by the high number of severe issues identified for both systems (Ca-0 and Ca-1, Ma-8 and Ma-7, respectively). The low rating for Calliope mini might again be related to the language barrier, caused by the German user language. As a result, most of the teachers could not explore the system in the way a German-speaker could, which led to a limited user experience, and thus, to the low positioning of this ER tool. In summary, the consistency between both rankings can be interpreted as another indicator for the usefulness of the heuristics as a guiding tool for the assessment of ER tools. The fact that the intuitive choices of the teachers matched with the ranking based on the usability issues, which could all be successfully mapped to the heuristics, indicates that there are no criteria outside the ones listed, that the teachers considered important enough to impact their opinion about the presented ER tools. It could be argued that the intuitive choices of the teachers could have been biased by the usability issues they identified before, and a matching between both rankings would therefore not surprise. However, as mentioned in the Methods subsection, each group identified different issues, and thus, none of the teachers was aware of

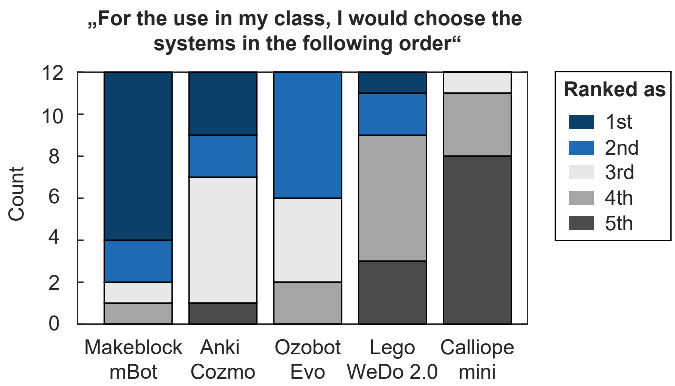

Figure 4.5 – Intuitive ranking of the ER tools by the participating teachers. Each bar indicates the number of teachers who ranked the corresponding system as their first, second, third, fourth or fifth choice for the use in class.

the aggregate of all issues when they indicated their intuitive choices.

Ranking based on system characteristics

Aiming at evaluating the effectiveness of the heuristics compared to more conventional assessment methods, a third ranking was established. For this purpose, the presented systems were rated by different characteristics and features, determined by means of the answers given by the teachers in the third subsection of the questionnaire. From a list of different system characteristics, the teachers had to indicate which ones they considered relevant using a 4-point Likert scale. For the ranking, the five characteristics with the highest values of approval (more than 83% of the teachers agreed or strongly agreed on their relevance) were included. Additionally, the teachers were asked to indicate the sensors and actuators they considered as most relevant for their teaching. For the ranking, the four most commonly mentioned sensors and actuators were included. The complete list with the most preferred system characteristics, sensors and actuators was then used by the experimenters to evaluate the presented ER tools: for each item on the list that a system complied with, it was given 1 point (Table 4.4). The final ranking (Table 4.5, fourth column) was then established based on the ratio between the points for each system and the maximum achievable score (13 points). The results illustrate that the ranking obtained following this approach did not provide a coherent match with the intuitive choices of the teachers. While the intuitive ranking determined Makeblock mBot as the teachers' clear preference, this approach ranks it in the second place (69% of the maximum score). Instead, Ozobot Evo (77%) was determined as the leader of the ranking, which has only been third based on the teachers' intuitive choice. Another remarkable difference is given by the good rating of Calliope mini (62%), which came in last far behind in the intuitive ranking. It is followed by Anki Cozmo (54%) and Lego WeDo 2.0 (38%), which ranked on the two last places following this approach. The obtained results underline the presumed weakness of conventional evaluation approaches for ER tools. Indeed, the isolated evaluation of system characteristics and components of ER tools appears to be insufficient to appropriately repre-

Table 4.4 – Participating teachers' preferred system characteristics

Characteristic	Anki Cozmo	Calliope mini	Lego WeDo 2.0	Makeblock mBot	Ozobot Evo
Wireless connection	X	X	X	X	X
Didactic material available		X	X		X
System extendable		X		X	
Usable on a desk	X	X	X		X
Preprogrammed	X	X		X	X
Distance sensor			X	X	X
Input button		X		X	
RGB ground sensor					X
B/W ground sensor				X	X
Motor	X		X	X	X
LED	X	X		X	X
Lifting arm	X				
Loudspeaker	X	X		X	X

Table 4.5 – ER tool rankings for three different approaches

Rank	By usability issues	By intuitive choice	By system characteristics
1	Makeblock mBot (Ca-0/Ma-1/Mi-7/Co-3)	Makeblock mBot (85%)	Ozobot Evo (77%)
2	Anki Cozmo (Ca-0/Ma-5/Mi-0/Co-6)	Anki Cozmo (63%)	Makeblock mBot (69%)
3	Ozobot Evo (Ca-0/Ma-5/Mi-7/Co-3)	Ozobot Evo (58%)	Calliope mini (62%)
4	Lego WeDo 2.0 (Ca-0/Ma-8/Mi-5/Co-1)	Lego WeDo 2.0 (33%)	Anki Cozmo (54%)
5	Calliope mini (Ca-1/Ma-7/Mi-3/Co-1)	Makeblock mBot (10%)	Lego WeDo 2.0 (38%)

sent the needs of compulsory school teachers. This is particularly interesting, since for this study, the teachers were given the possibility to specify the characteristics and components they considered as relevant. However, formerly such criteria have been selected and applied for the assessment of ER without demonstrating any evidence about their relevance to the users. Moreover, such approaches mostly focused on the properties of the robot. The findings of this study illustrate that HE may represent a more holistic approach, that can better reflect what is expected from ER tools when they are used for classroom activities.

4.3.3 Discussion

This study introduced the use of heuristic evaluation as a tool to guide the assessment of ER tools targeted to classroom use. Existing heuristics coming from different fields were combined and adapted to devise an initial set of fourteen heuristics, specifically aimed at meeting the needs and expectations of ER tools in formal education settings. A group of twelve compulsory school teachers participated in this study, to validate the suitability of HE for the assessment of ER tools. The findings illustrated that the heuristics embodied a good representation of the teachers' needs regarding the use of ER tools in classrooms, matching their personal choices based on intuition. The proposed heuristics could therefore guide researchers and engineers in the development of ER tools, providing design principles which are coherent with the needs and expectations of compulsory school teachers. In this context, the heuristics may not only be useful for the development of the robot, but also for the design of the programming/interaction interfaces. However, due to the small sample size, the results of this study should rather be interpreted as a proof of concept for the proposed set of heuristics. Therefore, studies with larger sample sizes should be considered, in order to draw more substantial conclusions. Nevertheless, it can also be argued that previous studies on heuristics for digital and tabletop games involved similar sample sizes for the validation (Desurvire et al., 2004; Köffel & Haller, 2008) and yet yielded recognized results. Another limitation of this study concerns the mapping of the identified issues to the heuristics, performed by the researchers and hence involving an implicit risk of bias. The next step of this research will therefore involve the application of the heuristics by independent evaluators, in order to consolidate their applicability. However, before moving on to such confirmatory studies, it is necessary, to ensure the validity of the heuristics through revision by external evaluators knowledgeable in the field. Indeed, this first set of heuristics was developed based purely on existing literature and the classroom experiences of the experimenters. Involving the feedback of teachers, developers and educational experts could provide valuable input to refine the heuristics and allow a better description of the heuristics. In this context, qualitative research methods, such as focus groups could be considered. As a matter of fact, focus groups have been suggested to be a useful method for the creation of design science research artifacts (Tremblay et al., 2010). The next section, will therefore present the organization of focus groups with teachers, developers and educational researchers to revise the heuristics and to demonstrate their utility in authentic development settings.

4.4 Focus groups to revise the heuristics and validate their utility

The use of focus groups is a well-established qualitative research method that has been applied in a wide variety of different fields, including political science, public health, marketing and education (Morgan, 1996). In focus groups, data is collected from moderated discussions of typically 4-12 participants who debate about a given topic. The discussions are usually guided by a moderator, who, depending on the intended degree of involvement, may direct the participants towards the matters of interest by following a predetermined questioning route. Compared to other methods, one of the main strengths of focus groups is that the findings can go beyond simple explorations of ideas. As suggested by Morgan, the interactivity of the setting may allow to provide insight into more complex behaviors and motivations. In the context of design science research, Tremblay et al. (2010) have suggested that focus groups can be used to improve design artifacts as well as to demonstrate their utility in a field setting. They proposed that incremental improvements can result from exploratory focus groups (EFGs), while evidence about utility can be achieved through confirmatory focus groups (CFGs). In both cases, participants of the focus groups are asked to use the design artifact prior to the discussions. However, since EFGs and CFGs have different objectives, they may follow different experimental protocols. With respect to these notions, here we describe the organization of EFGs to devise a new set of design heuristics for ER tools, and CFGs to demonstrate their utility in authentic development settings.

4.4.1 Study design

Participants

Seven focus groups were conducted with in total 53 participants (29 males, 24 females). Five EFGs were performed with 45 participants to iteratively revise the design heuristics. Participants were recruited from varying backgrounds to allow for different perspectives on the use of ER in classrooms. Finally, two CFGs were conducted with ER developers to demonstrate the utility of the heuristics in authentic development scenarios. A summary of the participants' details is presented in Table 4.6.

ER tools presented in focus groups

Participants in EFG1-4 were asked to select two systems among four commercially available ER tools to perform the HE: Cozmo, mBot, Lego WeDo, and Root. The interaction interfaces (all tablet-based) as well as the learning activities were selected by the experimenters with the aim of highlighting the distinguishing features of each robot, such as a camera and AI (Cozmo), electrical and mechanical assembly (mBot), brick-based construction (Lego WeDo) and the compatibility with whiteboards (Root). Based on the EduRobot taxonomy (Catlin et al., 2019), mBot and Lego WeDo can be classified as "build bots," while Cozmo and Root represent the class of "use bots." All systems were presented with their standard components (i.e., without any extra purchasable accessories). Participants in EFG5 and CFG1 worked with Thymio PaPL (Mehrotra et al., 2020), the successor of the Thymio TPL project (Mussati et al., 2019). This system consists of the Thymio educational robot (Riedo et al., 2013), a tangible programming interface based on physical paper tiles and an external webcam, and a set of playgrounds and programming tasks designed by the developers. CFG2 were the developers of the Azoresbot (Cascalho et al., 2019), a modular do-it-yourself educational robot based on the Arduino platform. The main objective of this project is to provide a low-cost educational robot that students and teachers can use for robotic competitions. The system used in CFG2

Table 4.6 – Summary of participants' details for all conducted focus groups.

Focus group	Gender (m/f)	Age (mean±sd)	Work experience (mean±sd)	Description
EFG1	3m/1f	35.8±3.9y	10.5±3.6y	Engineers from a non-profit association involved in the development of the educational robot Thymio. All participants had extensive experience with ER tools (e.g. BlueBot, Lego WeDo, Lego EV3, mBot, Ozobot, Edison, Sphero).
EFG2	1m/8f	39.4±6.3y	17.6±7.9y	Primary and lower secondary school teachers participating in a compulsory training program of a cantonal digital education initiative. Most teachers have worked with different ER tools (Thymio, BlueBot, WeDo, mBot, Lego NXT, Lego EV3) for at least three months.
EFG3	7m/5f	33.4±6.0y	6.1±5.4y	Ten primary and lower secondary school teachers enrolled in a certificate of advanced studies (CAS) in educational robotics and two lecturers of the class. Both lecturers had extensive experience with ER tools (especially with Lego EV3). Two teachers had more than one year of experience with the Lego EV3 robot, the rest had worked at most three times with different ER tools (Thymio, Lego EV3, Arduino, MIND, mBot).
EFG4	6m/4f	40.3±4.5y	16.1±7.3y	Primary and lower secondary school teachers representing local contact persons at schools participating in a cantonal digital education initiative. Two teachers had more than two years of experience with ER tools, one mainly with the Lego EV3 robot, the other with Thymio and BlueBot. The rest of the teachers have worked up to five times with different ER tools (Thymio, BlueBot, Lego EV3). One teacher indicated that they had no previous experience with any ER tool.
EFG5	4m/6f	39.3±5.9y	15.4±6.1y	A group of experts in learning sciences and educational technology. With one exception, all participants had a PhD degree in one of the following fields: computer science, educational technology, robotics, psychology, sociology, human-computer interaction, teacher education and media psychology. Only three participants had experience with ER tools (Thymio, Cellulo). However, all participants had prior experience with educational technology in general (e.g. e-learning platforms, MOOCs, social robots).
CFG1	4m	23.3±1.3y	2.2±1.4y	A group of young engineers developing Thymio PaPL. All participants had a BSc degree in Robotics or Mechanical Engineering, and they all had at least three months of experience with other ER tools (Lego EV3, ePuck) before joining the PaPL project. At the time of the study, all participants have worked for at least four months in the project.
CFG1	4m	54.3±2.3y	27.0±5.4y	A group of three experienced researchers and one teacher developing the Azoresbot. The teacher had a MSc degree in Biology, while the three researchers had degrees in Mathematics and Informatics (MSc), Computer Science (PhD) and Systems Engineering (PhD). Except for one researcher, all participants had prior experience with different ER tools (Lego EV3, Picaxe, Arduino).

Table 4.7 – Questions of the surveys presented to EFG2-4 and CFG1-2

Item	In survey presented to EFG2-4 and CFG1-2
Q1	I think it is important to consider alignment as presented in the ERLS framework when assessing the design of ER tools.
Q2	I think that heuristic evaluation is an appropriate method to assess the design of ER tools.
Q3	I think that the overall length of the heuristics is appropriate.
Q4	I think that the description of the heuristics is sufficiently detailed and precise.
Q5	I think that the heuristics are easy to use.
Item	*Questions additionally presented to CFG1-2*
Q6	I would naturally (i.e., without instruction) consider alignment as presented in the ERLS framework when designing ER tools.
Q7	I think that the heuristic evaluation helps to uncover a larger amount of issues.
Q8	I think that the heuristic evaluation helps to uncover issues that cannot be found only by intuition.
Q9	I think that the organization of the heuristics is clear and well-structured.
Q10	I think that the heuristics cover all relevant aspects for the design of ER tools.

consisted of an Azoresbot, the Arduino IDE, and the playgrounds and programming tasks that students face in robotic competitions in the Portuguese Autonomous Region of the Azores.

Study protocol

For all focus groups, participants were first divided into subgroups of three to six people. Through a 15-minute presentation by the moderator (one of the experimenters), the groups in EFG1-4 were then introduced to the ERLS framework and the initial set of heuristics presented in Giang et al. (2019). Each group then chose two of the four ER tools at disposal and performed a HE for both systems for around 30 minutes each, identifying issues for each system guided by the original heuristics. Following EFG1, it was decided to ask for the participants' opinion on the ERLS framework and the HE by a survey with 4-point Likert scale questions (Table 4.7). The session concluded with a plenary discussion lasting around 30 minutes that was always moderated by the same experimenter. Based on the findings from EFG1-4, the heuristics were then revised, resulting in a new set of heuristics (see Table 4.8) that were then presented to EFG5 and CFG1-2.

EFG5, conducted with a group of ten educational experts, served as a pilot for the CFGs and, therefore, shared the same experimental protocol. Since the goal of the CFGs was to validate the utility of the heuristics in an authentic setting, a manipulation, as suggested by Tremblay et al. (2010) was applied. Therefore, after a 10-minute introduction to the ERLS framework, groups were first given 30 minutes to identify issues based only on their intuition. After the completion of this first evaluation, the participants were then introduced to the revised heuristics by a 5-minute presentation. The groups were then given another 30 minutes to do another assessment of the same ER tool – this time guided by the revised heuristics. If during the HE participants identified issues similar to the ones from the intuitive evaluations, they were asked to list them, nonetheless. Issues for which participants assigned multiple heuristics and for which no consensus was found, were assigned to the "Not clear" (NC) category by the experimenters. Following the HE, participants were also asked to perform a post-hoc assignment of the issues identified based on intuition to the revised heuristics. The focus groups concluded with the surveys and the plenary discussions. The surveys, in these cases, included some additional questions about the utility of the ERLS framework and the revised heuristics as perceived by the developers (Table 4.7). However, due to time constraints, the survey was only completed by CFG1-2.

4.4.2 Results

Revising the heuristics with EFG1-4

In EFG1-4, a total of 14 HEs were performed by all participants. Overall, all subgroups together identified 141 issues, with an average of 10.1±3.8 issues (mean±std) identified per HE. The distribution of the issues across the 14 heuristics was very diverse (Fig. 4.6). While a large proportion of the issues were related to the Pre-4 heuristic (*Interaction*, 34 issues), none of the 14 HEs identified issues related to *Challenge* (Pre-2), *Level of Automation* (Pre-5) and *Enjoyment and Aesthetics* (Pre-9). Moreover, for 22 issues, participants were not able to unambiguously assign them to a heuristic and therefore indicated two or more possible categories. These issues were therefore classified by the experimenters as *Not clear* (NC). No cases were reported in which issues could not be related to any heuristic.

The results of the surveys completed by EFG2-4 (Fig. 4.7) illustrated that with only one exception, all participants approved the importance of the ERLS framework (Q1). Moreover, all participants believed that HE would be an appropriate method to assess the design of ER tools (Q2). However, they also stated that improvements could be done in terms of the ease of use of the original set of heuristics (Q5). Specifically, some participants were not satisfied with their description (Q4), and more than half of the participants indicated that the length should be revised (Q3).

The plenary discussions allowed to further elaborate these perceptions. Indeed, most participants approved that the use of the heuristics in combination with the ERLS framework, was helpful in guiding the evaluation. However, with respect to the original heuristics presented in Giang et al. (2019), participants suggested a rearrangement of the categories to reduce conceptual overlaps between heuristics, and thus to improve orthogonality. Moreover, many participants advocated for a fewer number of categories, but with a more detailed description. They argued that this would make the heuristics clearer and thus improve their usability. Finally, participants also suggested to explicitly amend heuristics related to ER artifacts and the ERLS framework. Based on these findings, the heuristics were substantially revised (Table

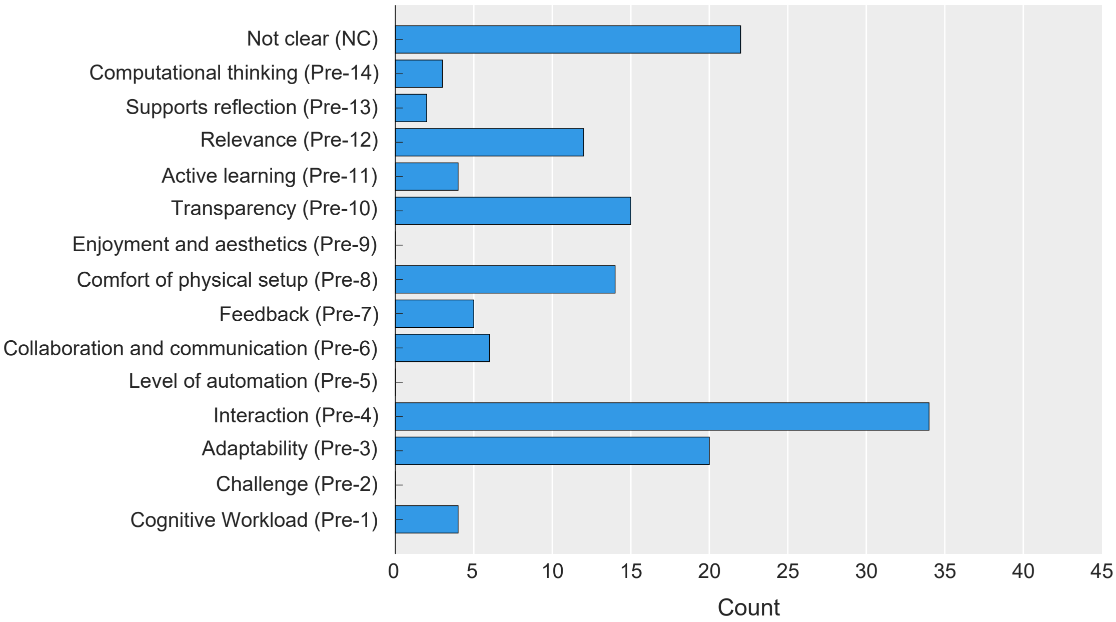

Figure 4.6 – Summary of the issues identified by EFG1-4 associated to the first set of heuristics.

4.8) before being presented to EFG5 and CFG1-2.

While some of the existing heuristics were kept unchanged in the revised set, some were strongly modified. Specifically, heuristics for which only few issues were found during the HEs of EFG1-4, were either removed or merged to create new categories. Furthermore, more detailed descriptions and a different structure were implemented for each category to prevent a lack of clarity and to facilitate the use of the heuristics. Following the development of the revised heuristics, all 141 heuristics identified during EFG1-4 were mapped to the revised heuristics by the experimenters (Fig. 4.8). Though also in the revised heuristics, a dominance of the Rev-4 heuristic (*Facilitating User Interaction*, 42 issues) was observed, the distribution

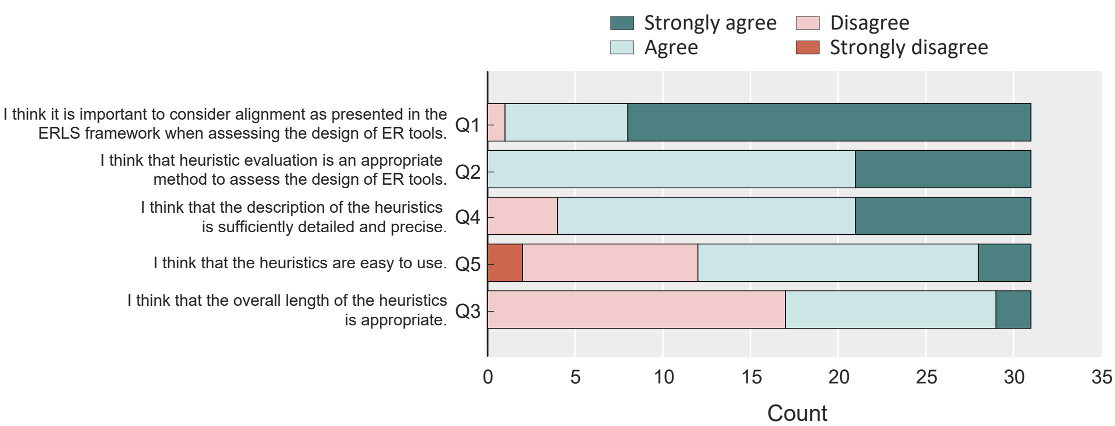

Figure 4.7 – Results of the surveys completed by the exploratory focus groups EFG2-4.

Table 4.8 – The revised set of design heuristics for educational robotics learning systems (HERLS)

Heuristic	Description
HERLS-1	*Educationally relevant* - The ERLS should embody constructionist learning approaches, involving the use of tangible objects and fostering active learning. - The ERLS should implement activities complying with desired learning goals, such as the ones listed in curricula and relevant competence models. - The ERLS should embody constructive alignment of learning goals, instruction, ER artifacts and assessment activities.
HERLS-2	*Inviting and engaging* - The ERLS should quickly grab the learners' attention and facilitate cognitive activation and immersion in the activities. - Learners should feel comfortable exploring and using the ER artifacts, without fear of breaking them. - The ER artifacts should be aesthetically appealing.
HERLS-3	*Supporting reflection* - The ERLS should encourage learners to reflect about their performed actions. - The ERLS should highlight the learning process. - The ER artifacts should provide a transparent environment allowing the inspection of underlying mechanisms.
HERLS-4	*Facilitating user interaction* - The ERLS should be usable in an easy and intuitive way, satisfying the learners' expectations. - The ER artifacts should provide feedback to facilitate the learners' interaction with the system. - The interaction with the ER artifacts should follow the logic of the instructional activities.
HERLS-5	*Versatile* - Learners should be able to use the ER artifacts regardless of their prior knowledge. - The ER artifacts should provide a rich environment allowing to implement a large variety of different learning activities. - The ERLS should allow for personalization of the learning activities.
HERLS-6	*Promoting collaboration and communication* - The ERLS should promote small group learning (i.e., groups of 2-3). - The ERLS should provide equal opportunities for participation to all members of a learning group. - The ERLS should foster collaboration and communication within a learning group and, if appropriate, between learning groups.
HERLS-7	*Compatible with classrooms* - The ER artifacts should be easy and fast to set up, disassemble, and maintain. - The ER artifacts should support teachers in classroom orchestration. - The ER artifacts should comply with the highest quality and safety standards. - The ER artifacts should be accessible at a reasonable cost.

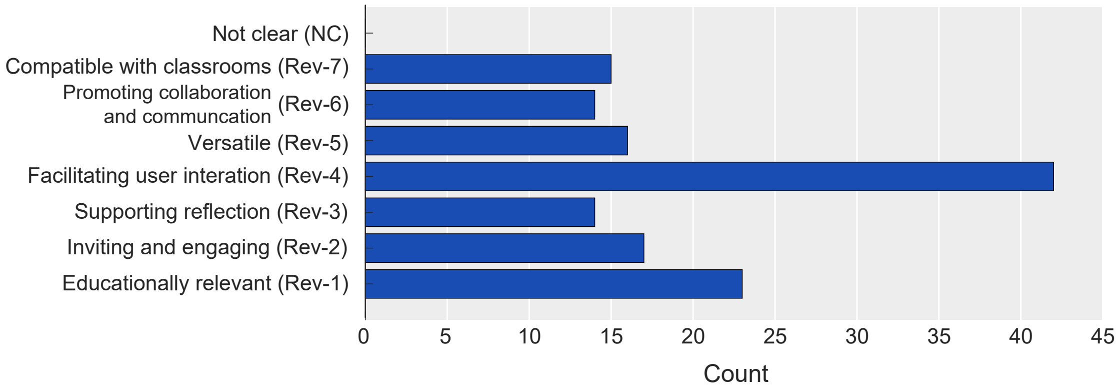

Figure 4.8 – Summary of the issues identified by EFG1-4 mapped to the revised heuristics.

across the remaining categories appeared to be more balanced than in the previous set. Particularly, in this new set, there were no categories with less than 14 issues, while for the previous set this was the case for ten categories. Moreover, the experimenters did not encounter a single case where an issue could not unambiguously be assigned, indicating a higher orthogonality of the revised heuristics. It was therefore decided that the revised set of heuristics was satisfactory and ready to be tested in confirmatory focus groups involving authentic development settings.

Demonstrating utility with EFG5 and CFG1-2

The participants from EFG5 identified 13 issues by intuition and 14 issues based on the HE with the revised heuristics for Thymio PaPL (Fig. 4.9). For the same system, CFG1 identified 16 issues by intuition and 14 issues by HE. For the Azoresbot, CFG2 identified 9 issues by intuition and 17 items by HE. Based on intuition, all groups identified at least two issues that were classified as NC (2 for EFG5 and CFG1, 5 for CFG2). Excluding the problems in the NC category, most of the issues identified by intuition (10 for EFG5, 11 for CFG1 and 3 for CFG2) were assigned to HERLS-4 (Facilitating User Interaction). This result is not surprising, since it can be assumed that issues related to user interaction are more obvious. Remarkably, issues related to any of the other six categories of the heuristics appeared to be more challenging to determine (only 1 issue for EFG5 and CFG2, 3 for CFG1). The distribution of the issues drastically changed when groups were performing the HE. With the heuristics, EFG5 and CFG1 found problems for all categories except for one (HERLS-5 and HERLS-6, respectively). CFG2 did not only found almost twice as many issues as compared to the intuitive evaluation. Based on HE, the group found at least one issue related to each category of the heuristics. With the heuristics, all groups identified problems that were clearly assignable, as opposed to the intuitive evaluation, where some issues had to be assigned to NC.

The surveys completed by CFG1-2 (Fig. 4.10), showed that participants agreed on the heuristics helping them to uncover larger amount of issues (Q7), and especially those that they would not have found based only on their intuition (Q8). Overall, the developers expressed high

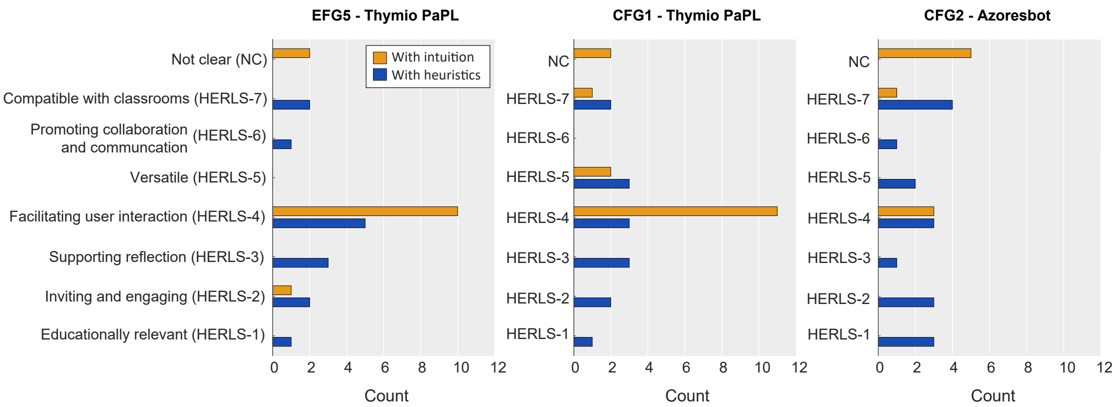

Figure 4.9 – Issues identified by EFG5, CFG1 and CFG2 based on intuitive evaluation (yellow) and heuristic evaluation using the revised heuristics (blue).

appreciation for the heuristics in combination with the ERLS framework: seven of the ten statements presented in the survey received entirely positive responses from the developers, supporting the utility of the heuristics as a guiding tool for the design of ER tools. However, one participant still found that description of the heuristics could be improved (Q4). Moreover, two participants did not indicate whether they believed that the heuristics cover all relevant aspects for the development of ER tools (Q10). Upon request, they stated not having chosen an answer because they felt they would lack the background to appropriately address this question. Finally, 6 out of the 8 developers acknowledged that they would not have considered alignment as presented in the ERLS framework without having been instructed to do so (Q6). This is particularly interesting in view of their responses to statement Q1, where all participants agreed on the importance of the ERLS framework. The plenary discussions with CFG1-2 provided further evidence that the proposed heuristics could be considered a useful tool to guide the design of ER tools. Participants particularly appreciated the conceptualization by the ERLS framework and affirmed that it helped them to approach the evaluation in a more systematic way. Interestingly, all participants stated that considering alignment of robot and interface was already intuitive to them before the focus groups. However, also thinking about playgrounds, learning goals, as well as instructional and assessment activities was natural to only two developers in the CFGs, of which one was a teacher. Most of the developers stated that they would first develop robot and interface and then think about the teaching and learning activities. As the main reason for this approach, participants indicated limited experience in the pedagogical domain. In this context, the heuristics provided them guidance throughout the evaluation process and helped especially those with no pedagogical background to give more attention to the alignment of the ER tool design with classroom needs. Even those coming from a pedagogical background, namely the teacher in CFG2, acknowledged that without the heuristics, they would have missed to identify some important issues.

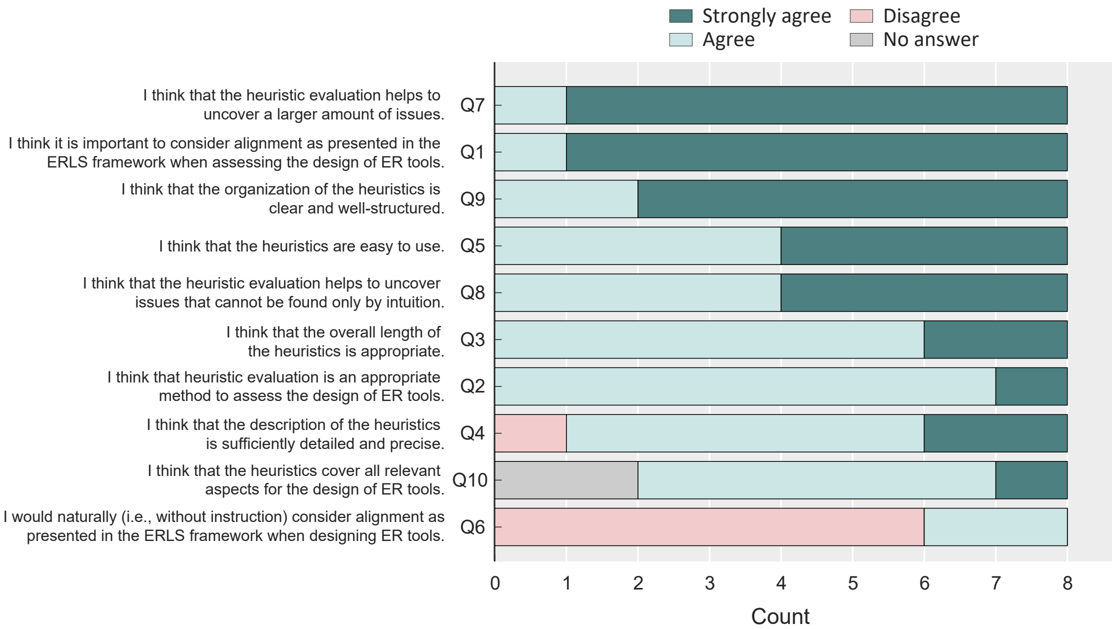

Figure 4.10 – Results of the surveys completed by confirmatory focus groups CFG1-2.

4.4.3 Discussion

Naturally, heuristics cannot fully replace teachers' know-how and years of experience. Nevertheless, the findings of this study illustrate that the devised heuristics appeared to be helpful in raising awareness for pedagogical issues, that even developers with no classroom experience were able to grasp. Especially when put in contrast with an evaluation based only on intuition, participants praised the HE as more comprehensive and purposeful. The confirmatory focus groups with two different groups of developers provided insights into the benefits of the approach in authentic development scenarios of ER tools. The heuristic evaluation in combination with the ERLS framework was highly appreciated by the developers and uncovered a broader variety of issues that also appeared to be better defined. A precise definition of issues could provide important support to developers and make design iterations more targeted and effective. Nevertheless, it should be acknowledged that the proposed approach also comes with the limitations that are inherent to research studies based on focus groups. Indeed, generalizability can be considered limited, due to the restricted number of focus groups that can be organized. Furthermore, the selected participants represent only a small sample of the target group and their contributions naturally involve a personal bias. The present study attempted to alleviate the effects of these limitations by a rigorous planning and by including a comparatively large number of participants, all with prior experience in ER. Quantitative data was integrated by means of the surveys as well as the analysis of the identified issues, to support the presentation of the findings. Though the revised set of heuristics is not claimed to be exhaustive, the results of the present study provide some promising insights about its utility in the context of authentic development scenarios for ER tools. Nevertheless, further studies with other groups of developers could be performed to further improve the heuristics and to provide more evidence supporting the proposed methodology. To take full advantage of the methodology, developers could consider performing the HE in the presence of teachers, as showcased by CFG2. The combination of HE and teachers' experience could provide more in-depth evaluations and help to uncover more issues relevant for the alignment of ER tools with classroom needs. Furthermore, future work could study the utility of the ERLS framework from the educators' viewpoint. In this context, the framework may also prove useful for the design of instructional and assessment activities.

4.5 Conclusion

In this chapter, we have introduced the use of heuristic evaluation as a method to complement the ERLS framework as an operational method to guide the development of ER tools and ER activities. Design heuristics were iteratively devised and evaluated in studies with teachers, developers and educational researchers. The final set of heuristics was presented to two groups of developers, who applied them to evaluate their designs in the context of authentic ER tool development settings. The results showed that the combination of the ERLS framework and the HERLS heuristics helped the developers to uncover pedagogical issues in their designs that they would not have identified based only on their intuition. The developers appreciated the proposed approach as a way to make design iterations more efficient and purposeful. Especially the fact that the heuristics helped to draw attention to education-related issues could therefore contribute to a better alignment of ER tools with classroom activities. Considering these issues already in the early design stages would help to make ER tools inherently better aligned with classroom activities and thus educationally more meaningful. To further illustrate how exactly the heuristics can guide the development of ER tools, the next chapter will outline two development projects: PaPL and CreroBot. Subsequently, we will present examples for the design of ER activities that have been guided by the proposed framework.

5 Alignment of the design of ER tools

Disclaimer

The content of this chapter has been adapted from the following works - with permission of all co-authors and publishers:

- Mussati, A., Giang, C., Piatti, A., & Mondada, F. (2019). A tangible programming language for the educational robot thymio, In *2019 10th international conference on information, intelligence, systems and applications (iisa)*. IEEE. https://doi.org/10.1109/IISA.2019.8900743

  As the main supervisor of the student project that resulted in this work, my contribution to this work involved: conceptualization, methodology, investigation, supervision and writing - original draft preparation.

- Giang, C. (2019). An open development environment for paper-based programming languages. *Research proposal for small projects approved by the Hasler Stiftung.*

  As the main applicant of this research proposal for small projects submitted to the Hasler Stiftung, my contribution to this work involved: writing - original draft preparation.

- Mehrotra, A., Giang, C., Duruz, N., Dedelley, J., Mussati, A., Skweres, M., & Mondada, F. (2020). Introducing a paper-based programming language for computing education in classrooms, In *Proceedings of the 2020 acm conference on innovation and technology in computer science education*. https://doi.org/10.1145/3341525.3387402

  As the main supervisor of the student project that resulted in this work, my contribution to this work involved: funding acquisition, conceptualization, methodology, investigation, supervision and writing - original draft preparation.

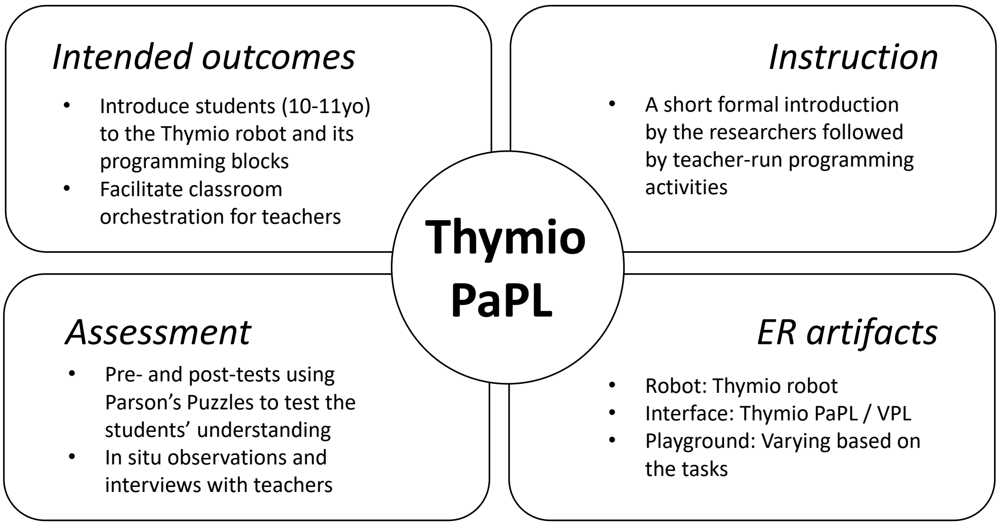

Figure 5.1 – ERLS implemented in the teacher-run classroom study conducted with the PaPL platform.

5.1 Summary

In this chapter we will outline how the ERLS framework and the HERLS heuristics have been applied in two authentic development projects of ER tools. The chapter will start with an introduction to tangible programming languages (TPLs), presenting previous works in this field and providing a motivation for the development of the Thymio PaPL, the first example discussed in this chapter.

The Thymio PaPL project studied the implementation of a tangible programming interface for the educational robot Thymio based on accessible materials such as paper and cardboard. Launched as a semester project for master students at EPFL, the development progressed rapidly, eventually leading to a successful research grant application funded by the Hasler Stiftung. In the following, we will first describe the development of a prototype of the Thymio PaPL platform, for which instructional alignment was not yet deliberately considered. We will then illustrate how the ERLS framework and the HERLS heuristics have been used to devise a classroom version of the platform. This version was finally tested in a teacher-run classroom study involving two teachers and their classes. The objective of this study was to assess whether the use of the PaPL platform can facilitate the introduction of basic Thymio programming activities in classrooms with respect to the screen-based alternative Thymio VPL. Therefore pre- and post-tests were performed using Parson's Puzzle quizzes presented to the students who worked with both interfaces. Moreover, we performed in situ observations and interviews with the teachers to explore how they perceived the PaPL platform as a way to facilitate classroom orchestration. A summary of the ERLS implemented in this classroom study is depicted in Fig. 5.1.

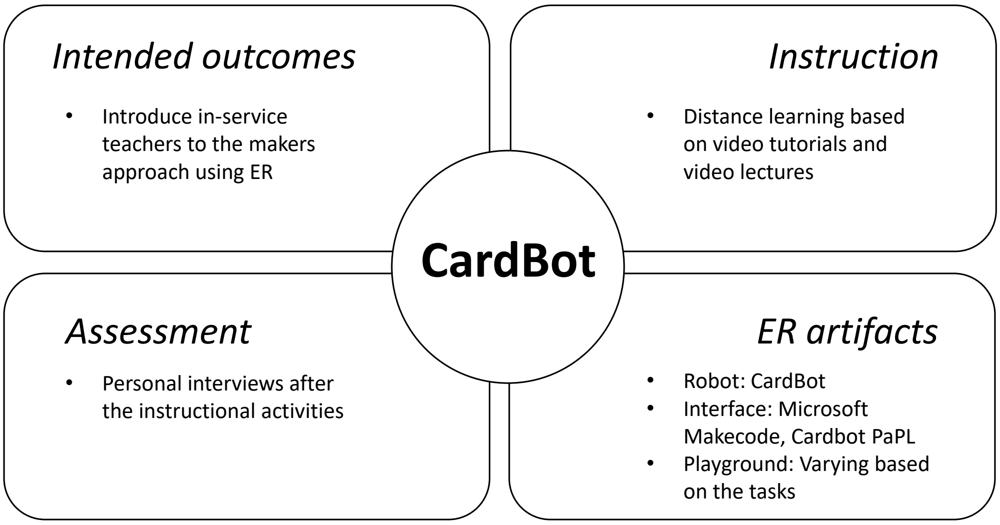

Figure 5.2 – ERLS implemented in the distance learning study with CreroBot.

The idea of the CreroBot project emerged from the Covid-19 pandemic situation that befell the planet in the winter of 2020. Restrictions on personal meetings made it impossible for educational institutions to continue business as usual and therefore, many distance learning solutions had to be devised. The CreroBot project represents such a solution. On the one hand, it allowed master students at EPFL to continue their semester projects on the development of an educational robot. On the other hand, it provided a way for in-service teachers participating in a continuous professional development program about ER to complete their coursework. In order to adapt to the special circumstances, the approach of using accessible materials from the PaPL project was expanded to the creation of a do-it-yourself educational robot based on cardboard and low-cost electronic components. In addition, the pandemic situation emerged as a possibility to explore the use of the ERLS framework in a new context, namely that of distance learning. Using it as a guiding tool, we designed CreroBot and a distance learning unit for the teachers based on video tutorials and lectures. To align the learning outcomes and the instruction with the use of CreroBot, the unit was framed in the context of making, a learning approach which has often been associated with ER activities. The teachers' experience with the system was finally evaluated in individual interviews organized at the end of the unit. A summary of the ERLS implemented in this distance learning study is depicted in Fig. 5.2.

Figure 5.3 – Two examples of tangible programming languages (TPLs): the KIBO robot and its wooden programming blocks (left) and the Osmo Coding Kit (right).

5.2 Tangible progamming languages (TPLs)

In this section we will introduce the idea of tangible programming languages (TPLs) as an approach for computing education. We will first outline the motivation behind such tools and we will discuss previous efforts made to develop them. Finally, we will identify current limitations of TPLs, that will motivate the first development project presented in this chapter, namely Thymio PaPL.

5.2.1 Motivation

Recently, more and more countries have recognized the importance of integrating digital literacy, computational thinking and related concepts into their school curricula. These competencies have been considered essential skills for future generations and therefore several countries have started to teach them or are planning to integrate them into compulsory schooling (Bocconi et al., 2016; Bocconi et al., 2018). In this context, the introduction of programming and fundamental computer science concepts has been deemed a major objective. To introduce novices and especially young learners to these concepts, usually graphical programming languages are used, such as Scratch (Resnick et al., 2009), Blockly (Fraser, 2015) or Thymio VPL (Shin et al., 2014). These interfaces aim at facilitating the introduction to programming by providing a simplified, more structured and playful environment, that allows programs to be created by simple drag and drop actions. Although such interfaces may indeed facilitate the introduction of programming and computer science concepts, they also have certain limitations. In most classroom scenarios, children work in groups and share a computer to perform the programming tasks. However, computers usually provide single user input devices, often resulting in unbalanced opportunities for participation among group members. As shown in the late nineties (Inkpen et al., 1999), providing opportunities for multiple children to interact with a digital environment simultaneously, can positively impact their levels of engagement, activity and motivation. In this regard, the use of tangible programming languages (TPLs) appears to be an interesting approach to address this issue. TPLs usually consist of physical blocks that can be assembled in the real world to control physical objects (e.g. educational robots, Fig. 5.3 left) or virtual agents (e.g. digital avatars, Fig. 5.3 right).

Previous work has demonstrated that with regard to screen-based graphical programming languages, TPLs may improve collaboration among a group (Horn et al., 2008), increase situational interest (Horn et al., 2008; Melcer & Isbister, 2018; Sapounidis et al., 2015), and have a positive impact on learning (Melcer & Isbister, 2018; Strawhacker et al., 2013). By involving tangible objects, they may also provide opportunities to implement more playful ways of learning, allowing to draw interest especially from younger children. The physicality of TPLs facilitates the implementation of constructionist learning approaches, which is one of the fundamental learning theories underlying ER. Moreover, by bringing the programming activities back to the real world, TPLs may also be valuable to address the concerns of those teachers and parents, who are still reluctant about introducing screens into classrooms. Finally, TPLs may also allow teachers to better monitor classroom activities and may hence provide support for classroom orchestration (Dillenbourg, 2013).

5.2.2 Related work

The potential of enhancing learning experiences through tangible interfaces has motivated many scientific groups to explore this field of research. Therefore, several attempts have been made in the past to develop different TPL systems. For instance, Horn et al. presented TERN, a TPL that allowed museum visitors to program the movements of a mobile robot (Horn et al., 2008). It was implemented using wooden blocks that could be connected in a puzzle-like fashion to create the programs. The blocks were captured by a camera mounted on top of the setup, that allowed the system to send the commands to the robot. In their study, the authors compared TERN with a purely screen-based implementation of the interface and found that with TERN, participants were more likely to explore the setup and to participate in groups.

Another example is T_ProRob, a TPL presented by Sapounidis and Demetriadis (2013). In several studies, they used T_ProRob to control Lego NXT robots. The system consists of different cubes that can be assembled to program the robot. Here, the links between the blocks are established electronically, and all cubes are connected to a master box, that allows to transfer the program to the robot. Their studies showed that compared to a graphical implementation of the interface, T_ProRob appeared to be particularly attractive for girls and easier to use for younger children (Sapounidis & Demetriadis, 2013). Moreover, they showed that children using T_ProRob produced fewer errors and debugged more efficiently (Sapounidis et al., 2015). Furthermore, younger children appeared to need less time to accomplish the tasks and older children appeared to be more engaged and interested in the activities (Sapounidis et al., 2015).

Another TPL developed to control a mobile robot is CHERP, developed at Tufts University. CHERP consists of wooden cubes that can be connected to create programs captured by a standard webcam. In a comparative study (Strawhacker et al., 2013), students were asked to solve different kinds of programming tasks, using CHERP, a graphical interface, or a combination of both. The results showed that for most kind of tasks, children using CHERP reached higher scores than children in the other conditions.

A modified version of CHERP was later used to implement the TPL of KIBO (see Fig.5.3 left), an educational robot developed by the same research group and aimed at preschool children. In a study with preschoolers, the authors showed that by using KIBO and its TPL, even children younger than 5 years old were able to learn fundamental programming concepts such as sequencing or repeat loops (Elkin et al., 2016).

In addition to TPLs used for mobile robots, there have been also several examples of TPLs which were developed to control virtual objects. One example is the Strawbies platform (Hu et al., 2015), the precedor of the TPL which today is known as Osmo Coding (see Fig.5.3 right). Using a mirror attached to the webcam of a tablet, the developers were able to capture blocks that were placed on a desk. The combination of several of these blocks allows the user to control the movements of a digital avatar. The goal of the activity is to make the avatar find its way through a maze, that is displayed on the tablet. A similar approach was chosen by Melcer and Isbister (2018), who developed Bots & (Main)Frames, a virtual game, in which the user has to guide the avatar through a maze, by passing the correct sequence of motions. The commands could be given through a TPL or a graphical interface. The results of their study with undergraduate students showed that compared to the graphical interface, the use of the TPL had a greater positive impact on learning, situational interest, enjoyment and programming self-beliefs.

Also Tada and Tanaka (2015) adapted this approach to create a TPL called "Sheets". Using paper cards, user could control a graphical object displayed on a computer screen. Each card represents a different command and multiple cards can be lined up to create a desired behavior of the object. The cards were captured by an external webcam mounted above the placement area. Moreover, in this work, the use of paper allowed the user to interact with the system by directly drawing on the cards, providing them another intuitive mean of interaction.

5.2.3 Current limitations

Although previous works have demonstrated promising results regarding the use of TPLs for computer science education, until today they have not succeeded to prevail in great scale. In most classroom scenarios, graphical programming languages such as Scratch or Blockly are still the preferred choice, when it comes to the introduction of programming and computer science concepts. One reason is certainly the fact, that graphical interfaces are easier to setup and do not require any additional accessories to work. Instead, integrating TPLs can be more time consuming and costly, since many systems involve mechanical and/or electrical components, that have to be purchased, maintained and stored. Considering the existing expenses for computers, tablets and/or educational robots, the additional costs for introducing TPLs can represent a crucial argument against their use in schools. Often, TPLs also involve complex electronics, which do not make it possible for teachers to use them in their classrooms without a certain level of expertise. Moreover, most TPLs are single purpose systems, developed for one specific environment, limiting their flexibility to be reused for other activities. Although teachers have expressed appreciation for TPLs, it seems like the current drawbacks have hindered a more widespread use of such interfaces.

While the graphical programming language Blockly has been adapted many times for different purposes, there are only very few TPLs that provide a framework for customization. One promising approach towards the creation of a more flexible and versatile TPL was Google's Project Bloks (Blikstein et al., 2016). Intended as a physical counterpart for Blockly, the idea was to provide basic blocks that developers could modify to create their own TPLs. Each block includes some basic electronics which is powered and communicating through the connection to a central block. The central block can then communicate with digital applications or robots through a wireless connection. So far, three different applications have been demonstrated (Blikstein, 2019), however, the impact of the platform is still far from being comparable with the one of Blockly.

5.2.4 Goals of this project

To promote the use of TPLs in classrooms, this project was aimed at implementing a versatile environment for the development of paper-based programming languages (PaPLs). Based on computer vision algorithms, this platform provides a low-cost and customizable solution, which does not require schools to purchase any extra accessories. Instead, it aims at better exploiting existing technological resources such as computers and tablets with standard webcams. The use of paper as a principal mean of interaction allows teachers and students to fabricate and customize their own programming blocks. Moreover, it provides teachers with new possibilities to design, conduct and evaluate classroom activities. Past research has highlighted the affordances of paper as an interaction medium. Not only is paper intuitive to use, it also represents a generic medium that can carry any kind of information (Zufferey et al., 2009). Although many attempts have been made to create paper-based user interfaces for different purposes (e.g. augmented office applications (Arai et al., 1995), interactive presentations (Nelson et al., 1999) or architectural tools (Terry et al., 2007)), only very few have implemented PaPLs (Sabuncuoğlu & Sezgin, 2020; Tada & Tanaka, 2015). Combining the versatility of paper with constantly improving computer vision techniques can provide countless opportunities to create new PaPLs at very low-cost.

To this day, paper is still the predominant medium used in classrooms. As a matter of fact, many teachers already use printed versions of programming blocks to introduce their students to graphical programming languages. Adding functionality to these blocks through computer vision algorithms could open up new opportunities for teachers to implement classroom activities. In contrast to solutions involving complex electronics, PaPLs allow teachers and students to fabricate their own blocks, leading to strongly reduced costs and more possibilities for personal involvement. Furthermore, past research has identified paper interfaces as beneficial media to support teachers in classroom orchestration, since they allow to incorporate five important design principles: control, visibility, flexibility, physicality and minimalism (Dillenbourg, 2013). The use of PaPLs may therefore facilitate the teachers' supervision, allowing them to more efficiently monitor and adapt classroom activities than they can with traditional screen-based interfaces. These findings suggest that the use of PaPLs could thus be especially favorable for the alignment of ERLS in the context of classroom activities. In the following, we will describe the development of the PaPL platform, and outline the studies evaluating the implementation for the Thymio robot. In a second iteration of the Thymio PaPL project, the HERLS heuristics presented in the previous chapter were used as a tool to guide the development of the platform.

5.3 Introducing a paper-based programming language (PaPL)

This section describes the development of the first prototype for Thymio PaPL, a paper-based programming language for the educational robot Thymio. In the following, we briefly present the Thymio robot and its graphical programming interface Thymio VPL. Subsequently, the development of the first prototype of Thymio PaPL is described, followed by the results of a pilot user study conducted with 77 university students.

5.3.1 A first prototype for Thymio PaPL

The PaPL project started as an semester project for EPFL master students in spring 2019. The original idea was to explore the implementation of a tangible programming interface for the educational robot Thymio using accessible and low-cost materials. One of the predominant programming languages used to control the Thymio robot is the block-based, graphical programming language Thymio VPL (Shin et al., 2014). The language consists of two types of blocks, events and actions, that are combined using a drag-and-drop interface to create instructions for the robot. In the first step of the PaPL project, the objective was to develop a tangible version of Thymio VPL.

The Thymio robot and the VPL programming interface

The educational robot Thymio (Fig. 5.4 left) was designed to make robotics education accessible to a wide audience of potential users, regardless of age and prior knowledge (Riedo et al., 2013). The robot is equipped with several infrared sensors that can for instance detect obsta-

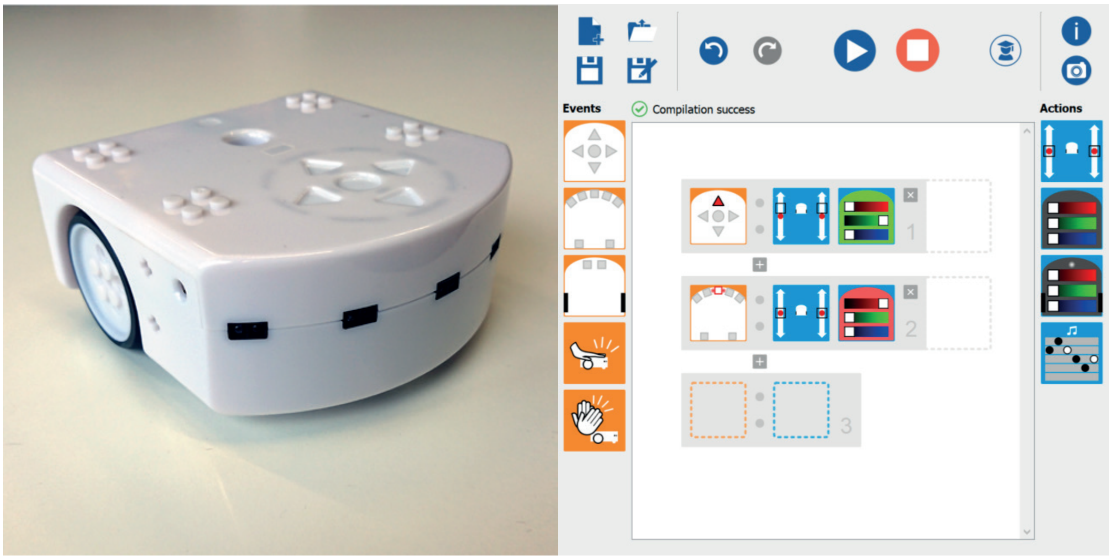

Figure 5.4 – Front view of the Thymio robot (left) and screenshot of the Thymio VPL programming interface (right).

cles surrounding the robot or the brightness of the surface the robot is placed on. Moreover, to perceive external stimuli, the robot has capacitive touch buttons, an integrated microphone, an accelerometer and an IR receiver for remote controls. To interact with its environment, the robot can be put into motion using its two-wheel differential drive, it can play sounds using its integrated loudspeakers or it can use its integrated LEDs to light up in different colors.

The robot comes with six pre-programmed modes, that facilitate the exploration of the robot's sensors and actuators. By connecting the robot with a computer using a USB-cable or a wireless USB-dongle, Thymio can also be programmed manually. The robot is compatible with different programming interfaces: graphical languages such as Scratch or Blockly can be used as well as textual languages such as Python or ASEBA (Rétornaz et al., 2013). Moreover, to help especially inexperienced users get started, Thymio VPL (Visual Programming Language) has been developed, another graphical programming interface, specifically tailored to the characteristics of the robot. The VPL language does not involve any text and relies purely on iconic representations of the robot's properties (see Fig. 5.4 right). The orange event blocks represent the different inputs the robot can receive through its sensors (e.g. infrared sensors, capacitive touch sensors, accelerometer, acoustic or remote control signals). The blue action blocks represent the different actions that the robot can perform (e.g. using the motors, playing sounds or changing the colors of the LEDs). Some blocks are parametric, allowing for instance to select specific infrared sensors of the robot or to set the speed of the motors. By combining event and action blocks, a specific behavior of the robot can be programmed. Merging multiple lines of event-action instructions allows the user to create more complex behaviors. For instance, the combination of blocks depicted in Fig. 5.4, makes the robot move forward and light up in green when the forward button is pressed, and stop and light up in red when the central front sensor detects an obstacle. In contrast to sequential programming

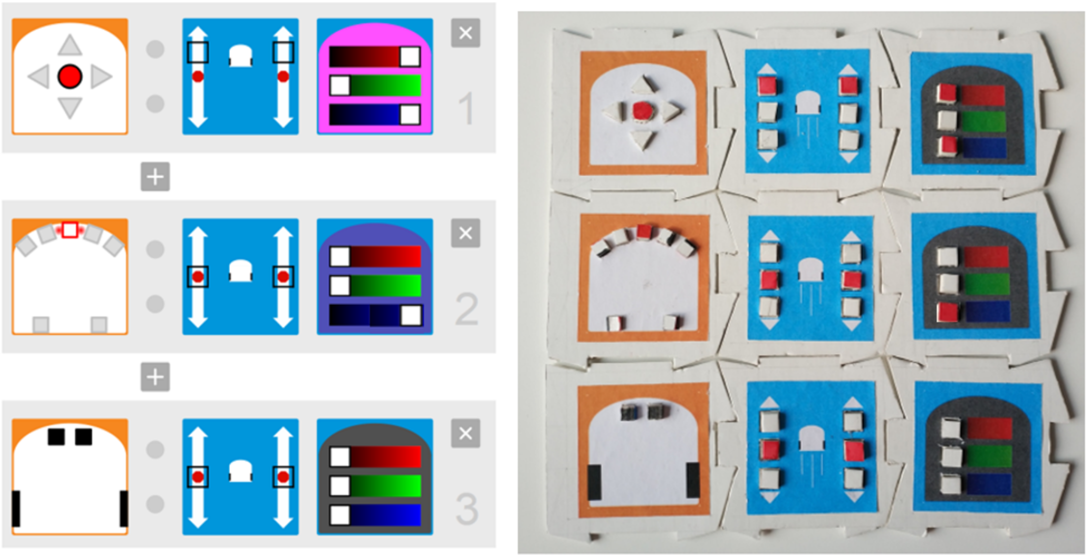

Figure 5.5 – Example of the same program implemented in Thymio VPL (left) and the first prototype of Thymio PaPL (right). When the central button is pressed, Thymio moves forward and lights up in purple. When the central front sensor detects an obstacle, Thymio stops and lights up in blue. When the bottom sensors detect void (e.g. at the edge of a table), Thymio stops and turns off the lights.

languages, the order of the lines in this event-based programming language does not matter. Actions are performed every time the corresponding event is being triggered.

Developing the first prototype of Thymio PaPL

The first prototype of Thymio PaPL (at that time still called Thymio TPL), consisted of programming blocks built from paper and cardboard and an external webcam attached to a laptop. In contrast to many other graphical programming languages, Thymio VPL purely consists of iconic representations, strongly facilitating the realization of a tangible version. Therefore, the design of the Thymio PaPL blocks was for the most part inherited from the existing VPL blocks, however, with minor modifications to ease the implementation. For instance, sliders in the VPL interfaces were replaced by fixed slots in which little cubes of different colors could be plugged in. On the one hand this reduced the granularity of the corresponding parameters, but on the other hand, it also allowed for a simplified construction of the blocks. The blocks were printed on paper and glued on cardboard pieces that could be assembled in a puzzle-like fashion (see Fig. 5.5 right).

An external webcam mounted on a physical support was used to capture the work space. The webcam was connected to a laptop on which a computer vision algorithm was running to detect which blocks were assembled together to form a program. The computer vision algorithm of the first prototype was implemented using C++ and the OpenCV library. Since

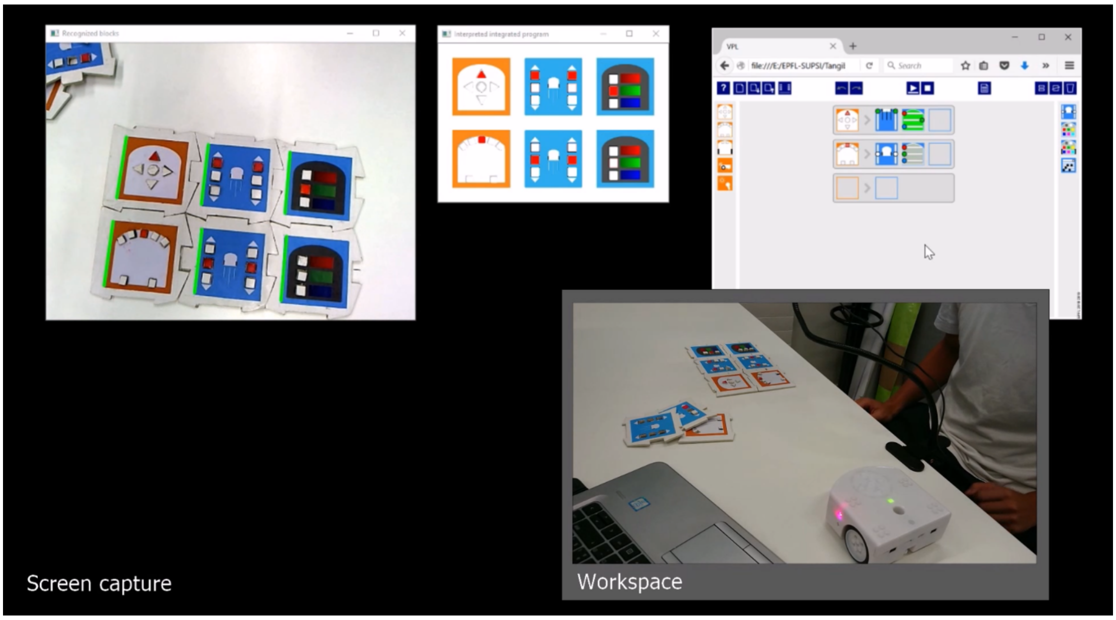

Figure 5.6 – Overview of the complete setup for the Thymio PaPL prototype.

most of the work related to the computer vision algorithms was done by the master students involved in the PaPL project, it will not be further discussed in this dissertation. Readers interested in the technical details of the computer vision algorithm are encouraged to look at the latest student report of the PaPL project (Guinchard, 2020).

The connection to the Thymio robot was established using the existing *Thymio Web Bridge* application and the protocols used for the newest, web-based version of Thymio VPL. Using this infrastructure, programs created with Thymio PaPL blocks could simply be sent to the robot by pressing the "arrow down" button on the laptop's keyboard. Fig. 5.6 gives an overview of the complete setup for the Thymio PaPL prototype. The bottom right depicts the work space with the Thymio robot, the PaPL programming blocks and the external webcam attached to a laptop. The top left depicts the live video feed from the webcam, where detected programming blocks have been highlighted with a green bar. The top middle depicts the blocks identified by the computer vision algorithm through template matching with the block library. The top right shows the web browser-based VPL3 interface, that has been used to transmit the commands to the robot. In the following, we will present how this this setup was used for a pilot study to perform a first evaluation of the usability and the potential of the Thymio PaPL prototype.

5.3.2 Study design

A pilot user study with 77 university students was organized to perform a first evaluation of the usability and the potential of the Thymio PaPL prototype. The study was conducted in the framework of demonstrations organized for first year master students at EPFL (master in Microengineering). In this context, the Thymio PaPL project was showcased as one of the projects realized at the Mobile Robotics Laboratory and the students were invited to test the platform. For this exploratory study, data was recorded based on qualitative observations of the experimenters. The study was organized in two afternoons of two consecutive weeks, with 39 students participating in the first afternoon (group A) and 38 students participating in the other (group B). Each demonstration afternoon was organized in four consecutive slots of thirty minutes. Therefore, during each slot around ten students were participating in the study. None of the students have worked with any of the interfaces before and each student gave their informed consent to participate in the study.

Study protocol group A

At the beginning of each slot, the student were welcomed by the experimenters and then split into two subgroups of 4-5 students. The session then continued with a five-minute speech by one of the experimenters, presenting the Thymio PaPL project and other activities of the research group. Subsequently, each group was given a set of Thymio programming tasks (e.g., "make Thymio follow a black line") and was assigned to one of the two work spaces in the room. One work space provided a laptop with the classical screen-based Thymio VPL interface (Fig. 5.7 left), while the other was prepared with the Thymio PaPL prototype (Fig. 5.7 right). The order of presentation for both interfaces was randomized: while one group started the activities with the VPL interface, the other started with the PaPL prototype. Each group worked in a separate work space on opposite sides of the room. After six minutes the groups switched platforms and were given a new set of tasks to perform during another six minutes. The main purpose of the tests with group A was to evaluate the usability and robustness of the devised Thymio PaPL prototype. Moreover, to perform a first evaluation on the potential of Thymio PaPL to foster collaboration, the experimenters took written notes during the activities. Based on direct observations, the number of students participating in the work (i.e., those who, during the six minutes work, were actively participating in solving the tasks at hand) was assessed. After each slot, the notes were discussed among both experimenters to reduce observational biases.

Study protocol group B

For the user tests with group B some changes were applied to the study protocol with respect to the tests with group A. While the overall procedure for the students remained the same, minor changes were applied to the introductory speech and the exercise sheets to allow for a better user experience. The most significant change in the study protocol was applied to the observational methods. For the sessions with group B, the objective was to gather more quantitative data regarding the potential of Thymio PaPL to foster team work and collaboration. Therefore, each group was videotaped while working with each interface. Each student gave their informed consent to participate in the study and to being video recorded during the study. The video recordings were then analyzed by two experimenters independently and the

Figure 5.7 – Examples of Thymio VPL (left) and Thymio PaPL (right) work spaces used in the pilot user studies.

results subsequently discussed among both experimenters to reduce observational biases. For these students, the timing and the type of the students' interactions with the platform and within the group were analyzed in detail. Therefore, four levels of interaction with the platforms were established. Every student was classified to be in one of these four states at any point during the six minutes:

- Not Interacting, Not Interested (NINI): e.g. checking the mobile phone, looking at other parts of the room.
- Not Interacting, Interested (NII): not working on the platform nor discussing with the others, but watching the platform and the work being performed on it.
- Non-Direct Interaction (NDI): discussing strategies, proposing solutions, telling other people to perform actions on the platform.
- Direct Interaction (DI): directly interacting with either the programming platform or with the robot to test the program.

Similarly, three states of collaboration among the members of a group were established. Every student was classified to be in one of these three states at any point during the six minutes, independently and in parallel with their classification for interaction:

- Not Collaborating (NC): either not working on the tasks or working independently of the others.
- Verbal Collaboration (VC): collaborating with others by discussing strategies, or by telling other people what action to perform.

Table 5.1 – Active participation of group A (n=38)

Platform	Actively participating	Not participating
Thymio VPL	13	25
Thymio TPL	25	13

- Physical Collaboration (PhC): directly working in parallel and with coordinated action on different parts of the system, e.g. one person programming while another tests the result on the robot.

If students were both verbally and physically collaborating, they were classified as PhC. The amount of time spent in each state of both dimensions was analyzed for each student independently. Finally, the results for both platforms were compared against each other. Moreover, qualitative data was collected by video analysis: whether one or a few persons would be the only ones directly interacting with the platform; how uniform the experience would be between participants, i.e., how similar the activity profile of each student would be; and whether the group would continue working on the tasks after being told that the time was over.

5.3.3 Results

The tests with group A showed that overall, Thymio PaPL was well received by the participants. The platform worked robustly and did not display any major technical issues. Most students acknowledged that Thymio PaPL provides an interesting alternative approach to coding, especially for novices. The physicality of the platform appeared to be more inviting and allowed multiple students to work in parallel. However, the students also argued that for more complex tasks, the use of the Thymio PaPL platform would become cumbersome. For these kind of programming tasks, the students would prefer the use of the screen-based Thymio VPL platform. Moreover, some of the PaPL blocks did not appear to be intuitive to the students and required some explanations by the experimenters. Table 5.1 summarizes active participation of the students for both platforms as observed by the experimenters. While around two thirds (25) of the students appeared to be rather passive when working with the graphical programming language VPL, this was true for only one third (13) of the students who worked with Thymio PaPL. For active participation instead, reverse conditions were found (13 students actively participating with VPL against 25 students with PaPL).

The tests with group B further emphasized the potential of Thymio PaPL to foster collaboration and team work. The analysis of the video recordings showed that on average, the students working with VPL spent the majority of their time (170 seconds, corresponding to 47% of the six minutes) in the Non-Direct Interaction (NDI) state, working with the platform only through others (Fig. 5.8). Almost a third of the time (30%) was spent in a Non-Interacting state (NINI: 5%; NII: 25%), while only about one quarter of the time (23%) was spent in direct interaction with either the VPL platform or the robot. In contrast, when working with PaPL, the students

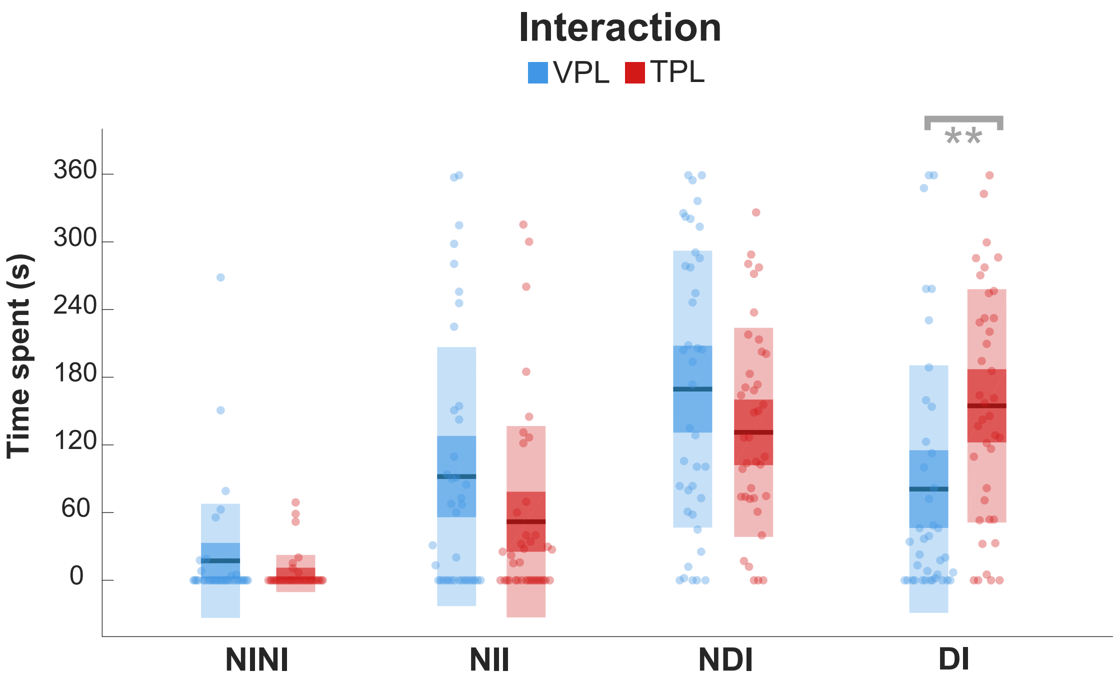

Figure 5.8 – Time spent in each interaction state for both interfaces in the tests with group B (n=39). Colored dots represent the time spent in each state by each individual student. Colored lines indicate median values for each state. Dark areas indicate the 95% confidence intervals, bright areas indicate one standard deviation. Results of a Wilcoxon signed-rank test between values are indicated by asterisks (**, $p < 0.01$). NINI = Not Interacting, Not Interested; NII = Not Interacting, Interested; NDI = Non-Direct Interaction; DI = Direct Interaction.

spent the majority of their time in Direct Interaction with the tiles or the robot (DI: 44%). A little less time was spent in Non-Direct Interaction (NDI: 39%) and only about a sixth of the time in a Non-Interacting state (NINI: 2%; NII: 15%). A Wilcoxon signed-rank test validated a statistically significant difference ($p < 0.01$) for direct interaction between both platforms.

As for collaboration, on average the students working with VPL spent most of their time (245 seconds, corresponding to 68% of the six minutes) in Verbal Collaboration (VC) with each other (Fig. 5.9), and more than a quarter of the time in Non-Collaboration (NC, 26%); mostly due to not participating in the activity at all (being in a Non-Interacting state), rarely because they were working on their own (e.g. silently checking the Thymio robot while the others were working on the program). Very little time (6%) was spent in Physical Collaboration (PhC); which mostly happened when one was testing the program on the Thymio robot while another corrected the code in VPL. With PaPL, the majority of the time (45%) was still spent in Verbal Collaboration (VC), however, the amount of time in this state was considerably lower than observed for VPL. Almost as much time (39%) was spent in Physical Collaboration (PhC), mostly with students preparing blocks or building instructions in parallel, by often implicitly, sometimes explicitly, splitting the assembling task between individuals or subgroups. Finally, only about a sixth of the time (16%) was spent in Non-Collaboration (NC), almost exclusively

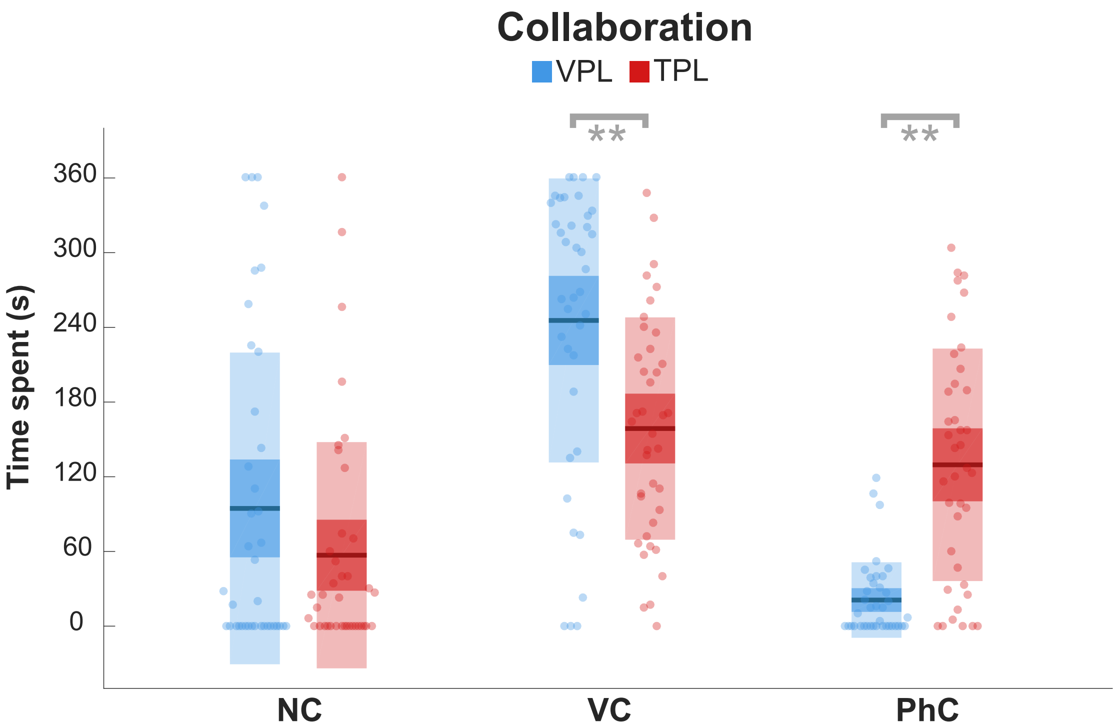

Figure 5.9 – Time spent in each collaboration state for both interfaces in the tests with group B (n=39). Colored dots represent the time spent in each state by each individual student. Colored lines indicate median values for each state. Dark areas indicate the 95% confidence intervals, bright areas indicate one standard deviation. Results of a Wilcoxon signed-rank test between values are indicated by asterisks (**, $p < 0.01$). NC = Not Collaborating; VC = Verbal Collaboration; PhC = Physical Collaboration.

from students who were in a Non-Interacting state. A Wilcoxon signed-rank test validated statistically significant differences in Verbal Collaboration ($p < 0.01$) and Direct Collaboration ($p < 0.01$) between both platforms, indicating a shift of observed collaboration states from one platform to the other.

From the video recordings made during the tests with group B, further qualitative analyses were performed: over the course of the six minutes of the VPL experiments, the number of interacting, interested, and collaborating students tended to peak in the first three minutes; while in the last three minutes those who were not interacting with the group tended to lose interest in the activity. On the opposite, interest appeared to be more constant during the PaPL experiments; some students who were not interacting with others continued manipulating tiles, even without major communication with the group. Moreover, those who at first were not collaborating in the activity tended to join in after the third or fourth minute. During the VPL experiments, most groups had only one to two students interacting with the platform; and sometimes another manipulating the Thymio robot; others would only help verbally. In contrast, with PaPL, most groups had either four or five students interacting with the tiles; and usually three to four of them would also manipulate the robot. Interestingly, even in strongly collaborating groups, when working with VPL, the main conversation would usually leave some people out. With PaPL instead, those who would be left out of the conversation would either join in through interaction with the tiles (e.g. picking up a tile and pointing out its features) or gather in secondary groups building an instruction, that could eventually be added to the program being built, giving them an opportunity to rejoin the main conversation. Six minutes after the start of the experiment, each group was told that the time was up, and they were asked to switch platforms or head to the debriefing, respectively. Of the eight groups working with the VPL platform, three continued their work; two for less than two minutes, one for another five minutes. Of the eight groups working with the PaPL platform, seven continued their work; five for less than two minutes; one for more than four minutes; another for more than eight.

5.3.4 Discussion

The findings of this exploratory study have illustrated the usability of the Thymio PaPL platform and its potential to foster active participation and collaboration within work groups. Thanks to the simple craft materials used for the fabrication of the PaPL blocks, they can be easily reproduced by teachers and students in a low-cost manner. Indeed, in the simplest implementation, the PaPL programming blocks can be purely paper-based, and the computer vision algorithm can be executed on a tablet or smartphone with an integrated camera. This provides schools, which are already using the Thymio robot, a possibility to acquire the platform without any additional expenses. With regard to accessibility, this represents a major advantage of the platform, since Thymio robots are already widespread in many schools in Switzerland and other European countries. In contrast to other platforms, Thymio PaPL does not have to be purchased as an additional accessory, hence considerably facilitating the access to a tangible interface. Another benefit of the devised platform is that it allows for a customization of the PaPL blocks. The block symbols can thus be easily adapted to fit the needs of a class (e.g. simplified, non-parametric symbols for younger students, or more complex symbols to introduce the notions of variables and functions to more advanced students). This provides teachers the opportunity to specifically adjust the presented blocks based on the level of their students. In terms of adaptability, this feature represents another advantage of the devised platform. Students may even be asked to design and fabricate their own hand-drawn symbols, allowing for a more personalized learning experience. Such approaches may not only be more intuitive to novices, but they may also better harness the pedagogical concept of constructionism underlying educational robotics activities. Although the results presented in this section appear to be promising, they should rather be interpreted as preliminary findings on the educational potential of the devised Thymio PaPL platform. It should be acknowledged that the results presented in this study are mostly based on qualitative observations of the experimenters. Although all results were discussed and double-checked by both experimenters independently to reduce any kind of observational bias, this condition certainly represents a limitation of this study. Moreover, it can be assumed that representing interaction and collaboration using a set of predefined states may not exhaustively cover the complexity of both concepts. Likewise, the comparatively short testing time does not allow for generalization to regular educational robotics activities, which usually take longer. In order to draw more substantial conclusions, further studies under more authentic conditions are needed. In particular, such studies have to involve participants from the actual target groups, i.e., pupils and their teachers. In the next section, we will therefore describe the next development iteration of the Thymio PaPL platform, which was finally tested in user studies with two classes of sixth graders and their teachers.

Table 5.2 – Usability issues identified in the heuristic evaluation of the Thymio PaPL prototype

ID	Description usability issue	Refers to heuristic
PaPL.1	The system is meant to be used for introducing basic programming and robotics concepts with Thymio. Considering this purpose, it is too complex.	HERLS-1
PaPL.2	The system lacks aesthetic appeal and is not very inviting.	HERLS-2
PaPL.3	There is no real user interface and too many different elements are involved (Thymio Web Bridge, Firefox browser, computer vision algorithm)	HERLS-2
PaPL.4	The current system does not provide any means for the user to reflect on the activities conducted.	HERLS-3
PaPL.5	The blocks displayed in the VPL3 window are distracting since they are not coherent with the PaPL blocks.	HERLS-4
PaPL.6	The system does not tell the user which part of the assembled program code is currently being executed.	HERLS-4
PaPL.7	Some of the icons (e.g. ground sensor or motor blocks) are not intuitive and difficult to understand without instruction.	HERLS-4
PaPL.8	The computer vision algorithm does not always detect the correct programming blocks.	HERLS-4
PaPL.9	The programming blocks may not comply with the highest safety standards (e.g. swallowing of parameter cubes by younger children).	HERLS-7
PaPL.10	Preparing the supporting cardboard structures is not easy and takes some time.	HERLS-7
PaPL.11	Setting up the physical support for the external webcam is very cumbersome.	HERLS-7

5.4 Aligning Thymio PaPL with classroom activities

In this section the second design iteration of the Thymio PaPL project is described. The project continued as a EPFL semester project in the fall semester 2019 with three students working on it as developers. The main objective of this design iteration was to improve and adapt the existing prototype for the use in classrooms. Therefore, the developers first performed a heuristic evaluation using HERLS to determine the shortcomings and limitations of the first prototype. These weak spots were then addressed in the development process, before the new version was tested in user studies with 32 sixth-graders and two teachers.

5.4.1 Heuristic evaluation of the Thymio PaPL prototype

The second design iteration of the Thymio PaPL platform started with a heuristic evaluation (HE) performed with the prototype from the previous semester project (Table 5.2). To this end, the three students that were going to work on the project performed the HE together. At the onset of the evaluation, the students were briefed about the learning goals intended to be achieved with the new platform. These were mainly to provide a smooth introduction to the functioning of Thymio's actuators and sensors and the event-based code structure. The new Thymio PaPL platform should provide a stepping stone to students before moving on to the screen-based VPL platform. Following this briefing, the developers were introduced to the ERLS framework and the HERLS heuristics. The students were then given 30 minutes to test the Thymio PaPL prototype. Therefore, they were asked to solve a set of Thymio programming tasks using the prototype. After the 30 minutes, the students performed the HE using HERLS. Each student identified usability issues independently. The testing session concluded with a plenary discussion, in which the students shared their identified usability issues and discussed the importance of each issue with their peers. Finally, a list with the most important usability issues was compiled (see Table 5.2), that was then used to guide the second design iteration of the project. Based on the results of the HE, the following design goals were defined for the next iteration of the Thymio PaPL platform:

- It should include a digital user interface that eases the use of the platform and that moreover, can provide user feedback (adressing PaPL.2, PaPL.3, PaPL.4 and PaPL.6)
- It should embed all important functionalities (computer vision, communication with Thymio, user interface) in one single application (PaPL.3 and PaPL.5)
- It should be completely paper-based to improve its compatibility with classrooms (PaPL.9 and PaPL.10)
- It should include a redesign of the programming blocks to make them more intuitive and easier to understand, especially for beginners. Moreover, a new design could also improve the robustness of the computer vision algorithm (PaPL.1, PaPL.7 and PaPL.8)
- It should include a new method to capture the assembled programming blocks. It should be easy to setup and at the same time facilitate strategic breaks to promote discussions and reflection (PaPL.4 and PaPL.11)

5.4.2 Developing a classroom version of Thymio PaPL

Based on these design goals a new version of Thymio PaPL was developed, that was expected to be better aligned with classroom activities. The new version was implemented in Python, allowing to integrate the computer vision algorithm, the user interface and the communication with the Thymio robot into one single application. The new user interface (Fig. 5.10) was aimed at making it usable by anyone regardless of their prior knowledge. Moreover, it provided feedback about the last event that was triggered in the Thymio robot, facilitating the debugging of the created program code.

The design of the programming blocks was substantially revised to make it more intuitive for beginners and to remove the dependency on the parametric cubes. The new blocks were

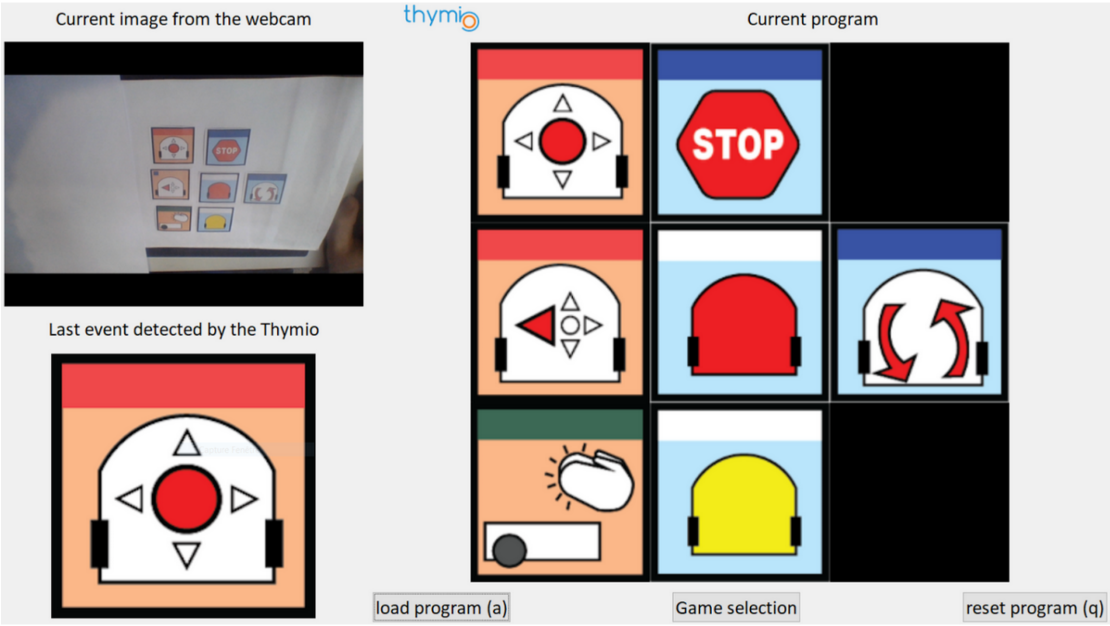

Figure 5.10 – Screenshot of the second version of the Thymio PaPL user interface.

entirely paper-based and were used by putting them on a placing grid printed on a DIN A4 paper sheet. The redesign of the blocks, involving new color schemes, also allowed to implement new computer vision methods to make the block detection more robust. For the camera setup, it was decided to remove the physical support. Instead, it was decided that the external webcam would have to be manually directed towards the programming blocks every time anew. The main idea was to deliberately decelerate the code creation process, in order to foster discussions and reflection within work groups. Moreover, to a certain extent, this setup already simulated the use of tablets for the computer vision algorithm, a feature that was in any case intended for future design iterations. An illustration of the new setup for the second version of the Thymio PaPL platform is depicted in Fig. 5.11. This new version was evaluated in a user study with 32 sixth-graders and 2 teachers.

5.4.3 Study design

To test the platform in an authentic context, the second user study focused on its use in a classroom environment. Thus, the objectives were to determine (1) the interaction of the target audience with Thymio PaPL and (2) the ability of the platform to facilitate the teacher's supervision of the programming activity.

Participants

The user study was conducted with 32 students and 2 teachers from two sixth-grade classes. At the time of the study, the students were 10-11 years old and parental consent was obtained

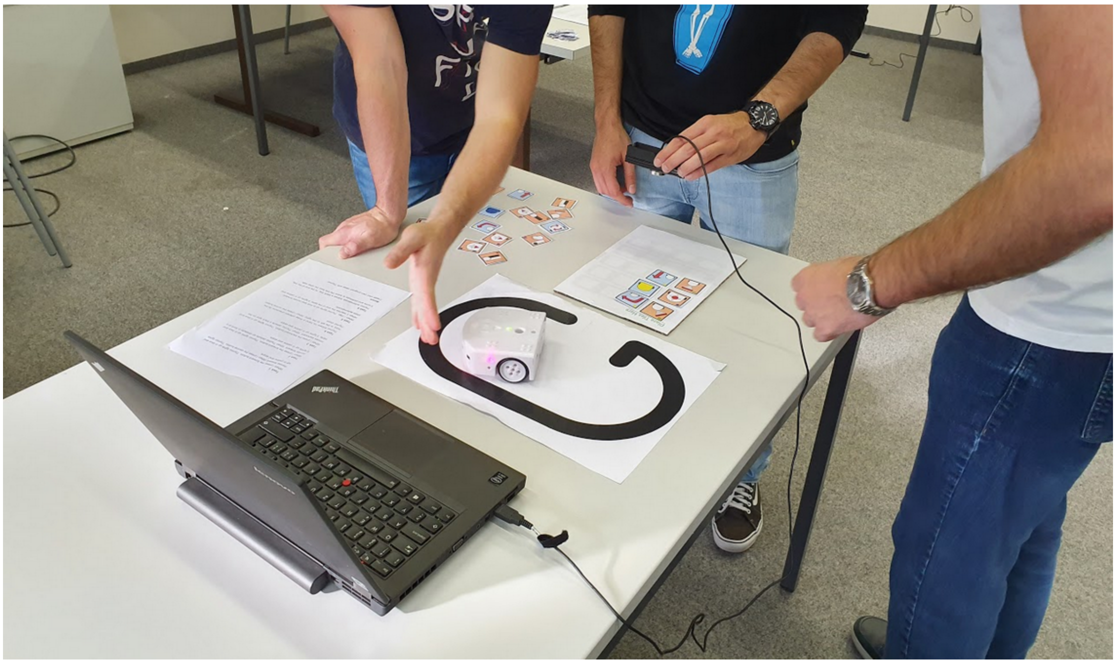

Figure 5.11 – Illustration of the second version of the Thymio PaPL platform.

prior to the study, allowing for the students' participation. Likewise, both teachers gave their informed consent to participate in the study. None of the students or teachers reported prior experience with the Thymio robot or with its graphical VPL programming interface.

Study protocol

The whole user study lasted around two hours and started with a plenary session, during which all students were given a five-minute introduction to the robot. Subsequently, the students were divided into groups of 2-3 (within their class) and given 25 minutes to explore the robot's pre-programmed modes under the supervision of one researcher. Meanwhile, 6 laptops with Thymio VPL were prepared in one room, while in another room, 6 Thymio PaPL setups were prepared. Another researcher briefed both teachers on the use of the Thymio PaPL and Thymio VPL interfaces. The teachers were also presented the solution sheets for the subsequent activities.

The exploration activity ended with a Parson's Puzzle (Parsons & Haden, 2006) containing five Thymio programming questions. Each question described a program specification, for which the students had to determine the correct combination of Thymio programming blocks. The goal of this activity was to familiarize the students with the main sensors and actuators of the Thymio robot, as well as introducing them to the event-based program structure. Therefore, the use of Parson's Puzzles appeared to be an appropriate assessment tool to attain alignment as described in the ERLS framework. The puzzles were attempted by each student individually. While the students of one class (i.e., class A) had to solve the questions using Thymio PaPL

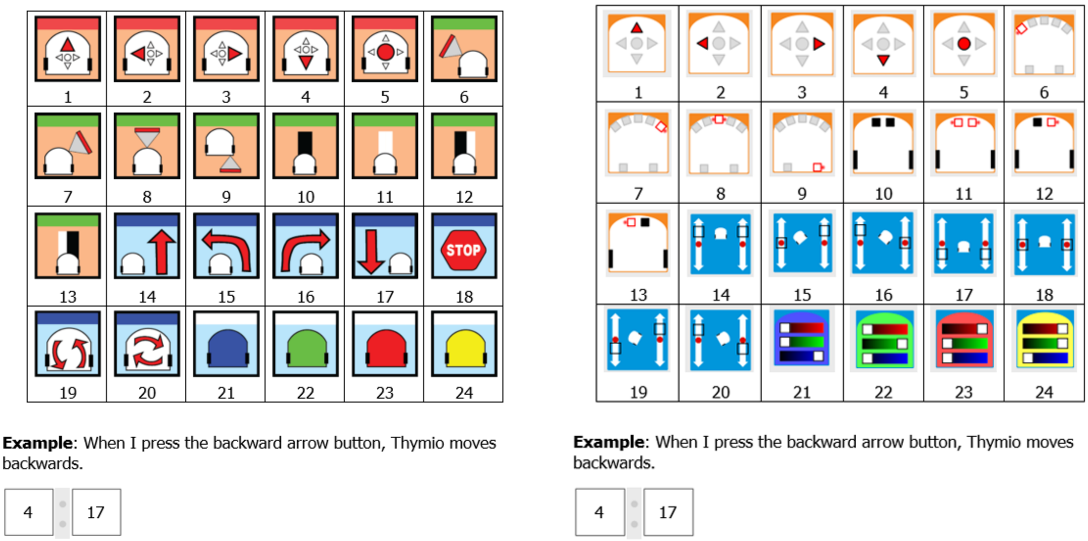

Figure 5.12 – Example of a Parson's puzzle question with the solution for Thymio PaPL (left) and Thymio VPL (right) programming blocks.

blocks (Fig. 5.12 left), the other class (i.e., class B) was presented the same questions with Thymio VPL blocks (Fig. 5.12 right). Subsequently, the teachers and their students were sent to the classrooms with the setup corresponding to their Parson's puzzle (i.e., class A to the Thymio PaPL setups and class B to the Thymio VPL setups). In the next five minutes, the students were given a short introduction to the respective platform by a researcher. This was followed by 20 minutes of programming activity, during which the students were asked to solve the same five tasks as presented in the Parson's Puzzle.

The whole activity was administered by the teachers, who were asked to conduct this as if it would be a regular class. In both classrooms, one researcher took the role of an observer and another, the role of technical support, whenever required. Following the programming activity, both classes were asked to complete the same Parson's Puzzle as before. Finally, in the last part of the experimental session, both classes switched setups, i.e. class A now worked with Thymio VPL and class B with Thymio PaPL. The students were given two new, more difficult, Thymio programming tasks. Again, both classes were introduced to the new platform and then given 20 minutes to solve the tasks.

Measures

The scores obtained in the Parson's Puzzles were used as a measure for the intuitiveness of the Thymio PaPL/VPL platforms. Each correct answer was awarded one point, leading to a maximum possible score of sixteen points. Parson's Puzzle scores were compared before (Pre) and after the programming activity (Post) to analyze improvements in both groups. The activity of the teachers was tracked by an observer using a software dedicated to the creation

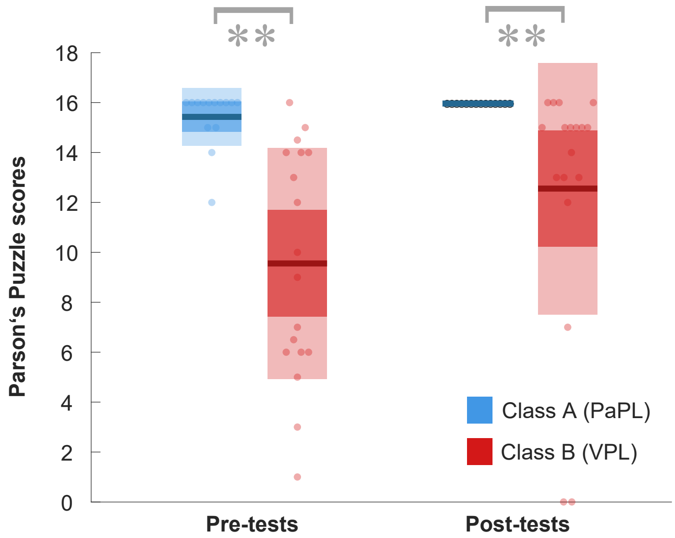

Figure 5.13 – Parson's Puzzle scores before (Pre) and after (Post) the programming activity for class A (PaPL first) and class B (VPL first). Colored dots represent the time spent in each state by each individual student. Colored lines indicate median values for each state. Dark areas indicate the 95% confidence intervals, bright areas indicate one standard deviation. Results of a Mann-Whitney U test between values are indicated by asterisks (**, $p < 0.01$).

of activity chronicles such as Actograph (SymAlgo Technologies, Paris, France). Specifically, observers monitored the interaction of the teacher with the groups during the whole duration. The teachers were not informed about these observations beforehand to prevent any change in behavior. Five days after the user study, both teachers met with the researchers for a 20-minute interview.

5.4.4 Results

Parson's Puzzle scores from students in both classes are shown in Fig. 5.13. Class A reached much higher scores (15.4±1.2 points) than Class B (9.6±4.6 points) before the programming activity. A Mann-Whitney U test validated a statistically significant difference ($p < 0.01$). After performing the tasks with the robot, the performance of both classes improved. While all students of Class A were able to achieve a perfect score (16 points), students in Class B still had varying results (12.6±5.0 points). Though the difference between both groups decreased, it was still statistically significant ($p < 0.01$).

The distribution of time spent with each group was evaluated for both teachers based on observational data (Fig. 5.14). Teacher A, starting with the Thymio PaPL platform, spent around half of their time attending to particular groups. While some groups needed more attention (e.g. group 2), others needed less (e.g. group 5). Throughout the 20 minutes of activity, the teacher followed all groups at least once. Moreover, the other half of the time, the teacher was not engaged with any specific group but monitoring the classroom in general.

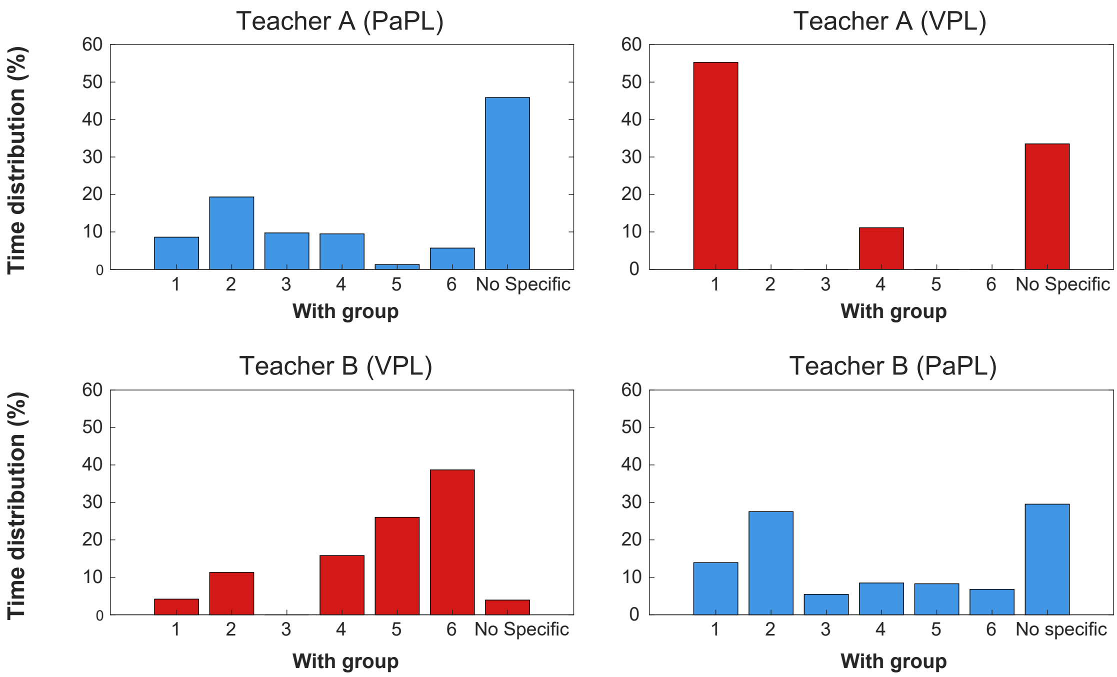

Figure 5.14 – Proportion of time teachers spent with particular groups during the programming activities.

When switching to the Thymio VPL platform, the teacher's behavior changed drastically. More than half of the time was now spent attending to a specific group. In this activity, only two groups had direct supervision of the teacher and the time spent in overall classroom monitoring decreased with respect to the first activity. Teacher B, who first performed the activity with Thymio VPL showed a similar trend in behavior. In the VPL setup, the teacher spent almost all of their time with particular groups and only a small fraction of their time on overall classroom monitoring. Yet only 5 out of 6 groups received direct supervision. When switching to the PaPL platform, the teacher now spent significantly more time on overall classroom monitoring and in addition, all groups were now receiving teacher attention.

5.4.5 Discussion

The results of this second study demonstrated strong potential in regards to the use of the Thymio PaPL platform in classrooms. Even before the programming activities, students presented to the Thymio PaPL blocks were able to attain high scores in the Parson's Puzzle tests. After the activity, they further improved and all students achieved maximum scores. On the other hand, students working with the screen-based Thymio VPL blocks reached lower scores in both tests. Providing an intuitive platform with "low floors" was argued to be an essential design element in computing education (Resnick et al., 2009). This might be particularly important to promote motivation, interest, and self-efficacy in novices. The interview with the teachers suggested that introducing Thymio PaPL prior to Thymio VPL could help students build better foundations and may, therefore, represent a smoother entry point to computing education. The teachers further acknowledged the potential of the PaPL platform to promote collaboration and communication. Thanks to its simplicity and intuitiveness, the platform may also support inexperienced teachers in running and orchestrating computing education lessons. The observations of the teachers' activities showed that both teachers provided more balanced support to all groups when they were working with the PaPL platform. Furthermore, they also had more time for overall classroom monitoring. This allows them to diagnose the progress of all the groups and provide more support where necessary.

Nevertheless, the results of the studies presented in this work are subject to the limitations of the small sample sizes and particularly the brevity of the interventions. In order to draw more substantial conclusions, future studies should include larger samples, longer interventions as well as methods to evaluate the effective learning gains in the long-term. Such studies are needed to provide real evidence about the advantages of paper-based programming languages for computing education. Moreover, it should be investigated to what extent such approaches can be favorable and when should the transition to more complex systems be considered. Finally, this work has explored the PaPL framework for event-based programming. However, it could also be studied how PaPL can be used for sequential programming paradigms involving control structures such as loops and conditions.

5.5 CreroBot - an accessible DIY educational robot

In this section we will present the development of CreroBot, an accessible educational robot that is based on a maker approach. In the past decade, the maker movement has gained a lot of attention, fueled by the evolution of constructive technology and widespread access to the internet. Researchers and educators have discussed the role of learning by making (i.e., building knowledge by building things) to transform the educational landscape, particularly in combination with ER. Describing ER as the pioneer of the maker movement, Eguchi (2017) has emphasized its potential to "provide all students with the opportunity to learn the skills and knowledge that they need to become effective members of the workforce and future innovators and creators". While making is not new in informal settings, Eguchi emphasized that it is important to also bring these approaches to formal education, in order to make it accessible to all students. According to her, it is however, important "to make sure that the materials and tools are accessible to all children".

Focusing on the maker approach provided an ideal framework for our CreroBot study. The restrictions on personal meetings caused by the Covid-19 pandemic, encouraged us to prepare a distance learning unit for in-service teacher participating in a continuous professional development program about ER based on the ideas of making. The project-based and self-regulated learning approach underlying maker activities allowed to leverage the do-it-yourself character of CreroBot and hence provided a good framework for the instructional alignment of the activities. In this context, the ERLS framework was used to guide the development of

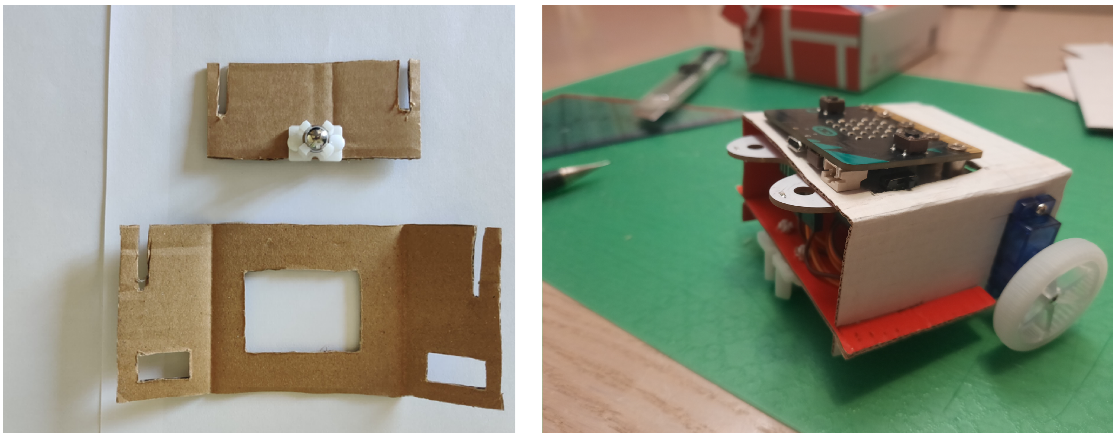

Figure 5.15 – Main cardboard construction pieces (left) and example of a CreroBot in the basic version (right).

the CreroBot design aiming at an ER tool that could be used for distance learning purposes.

5.5.1 Introducing the CreroBot

To allow for the use in a distance learning context, the CreroBot system had to be easily reproducible by individuals. Therefore, a particular emphasis was laid on sourcing minimal components. Instead of basing the design on materials and tools usually used in FabLabs and maker spaces, the CreroBot project aimed at exploiting materials that are easily available. Cardboard is such a material. It is a structure that both teachers and students are familiar with, and it is moreover, almost infinitely available at practically no cost. Thus, it was decided upon as the material for building the external body frame of CreroBot. To further facilitate the construction procedure, the body of the CreroBot was based on only two cardboard pieces that could be cut out and combined using a slit system (Fig. 5.15).

Meanwhile, the BBC micro:bit micro-controller has gained popularity in the domain of physical computing education due to its affordable price tag (around 15 CHF), compact design and variety of inbuilt electronics (e.g magnetometer, temperature sensor, light sensor, accelerometer, LED matrix, etc.). Many schools have started to use the micro:bit as a tool to teach coding, making it an ideal hardware base for the CreroBot. Basing the design of CreroBot on a micro:bit does not only enable educators to simply integrate CreroBot as another possible application of the micro:bit, it also allows for a continuity of the activities, which are all based on the same electronic system. This may significantly reduce financial expenses as well as efforts for educators and students to get used to a different systems every time anew.

The micro:bit can be transformed into CreroBot by adding simple components. The most basic version of CreroBot includes two micro-servo motors, two wheels (possible to craft with cardboard), jumper cables, a battery holder and an universal wheel (Fig. 5.15). More complex version of the CreroBot can be implemented with help of micro:bit expansion boards and

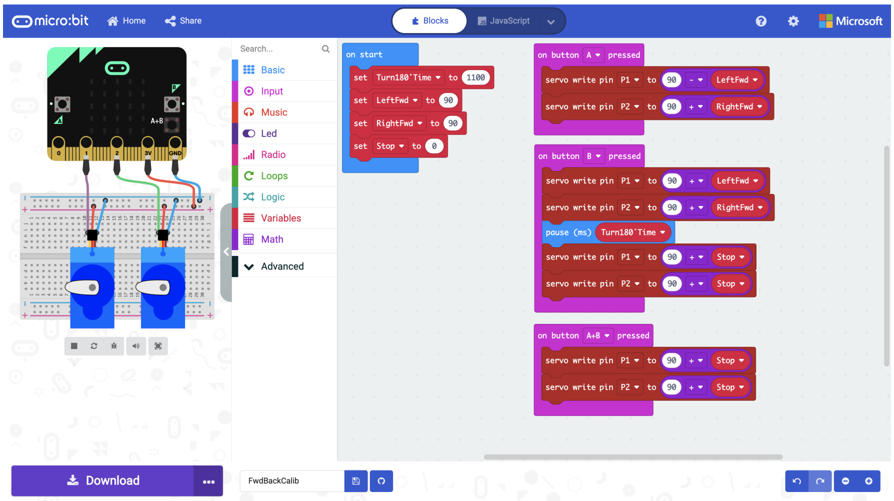

Figure 5.16 – The Microsoft Makecode programming platform used with CreroBot.

additional electronic components.

The CreroBot can be programmed with any programming interface compatible with the micro:bit. However, in the following we will focus on two specific platforms, that have been used in the pilot study with the in-service teachers:

1. Microsoft Makecode for micro:bit: The browser-based platform provides block-based programming similar to the interfaces such as Scratch or Blockly. The platform also provides a way to display and edit the code in JavaScript, allowing for the creation of more complex programs (Fig. 5.16).

2. CreroBot PaPL: A tangible programming language developed based on the PaPL framework and specifically adapted to CreroBot. Using arrows printed on paper and pasted on cardboard support pieces that can be connected in a puzzle-like fashion, the robot can be instructed to move in different directions. This platform follows a sequential programming paradigm with the option of using loops and nested loops. For CreroBot PaPL an Android version of the PaPL framework has been developed, which provides a wireless conncetion to the micro:bit via Bluetooth Low Energy (BLE). To send the commands to the robot, a photo of the arranged programming tiles can be taken and directly sent to the CreroBot (Fig. 5.17).

5.5.2 Study design

A pilot user study was organized to evaluate the usability and the educational potential of the CreroBot system. Due to the Covid-19 pandemic that accompanied the entire course of the project, it was not possible to perform any studies in presence. Instead, alternative ways to

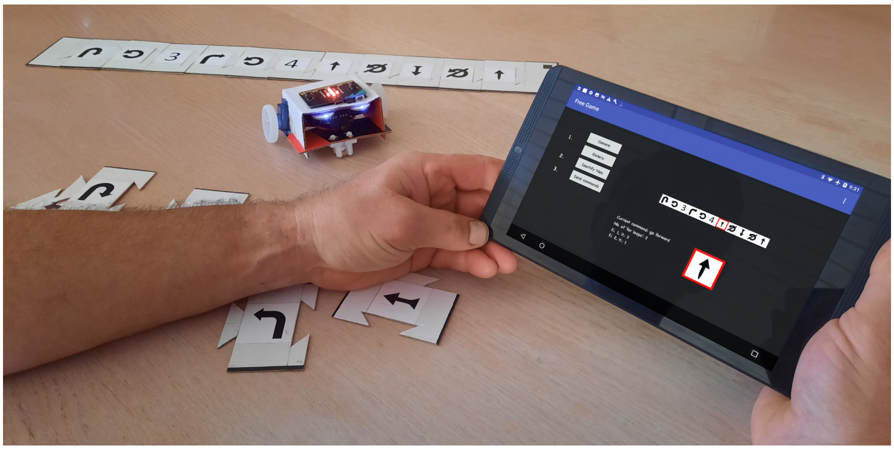

Figure 5.17 – The CreroBot PaPL programming platform used with CreroBot.

conduct user studies in remote had to be explored. In this regard, the results of this study may also be interesting from the viewpoint of distance education, a field that was already gaining more and more importance even before the Covid-19 pandemic.

Participants

The pilot user study was conducted with a group of five in-service teachers (two males and three females) enrolled in a Certificate of Advanced Studies (CAS) in Educational Robotics, as part of their continuous professional development. Three teachers worked in elementary schools (with children between 6 and 11 years old) and two teachers in middle schools (with children between 11 and 15 years old). The activities of the study were presented as part of the training program, substituting lessons that had to be canceled due to the Covid-19 pandemic. At the time of the study, the teachers were half way through the training program, which typically requires around one and a half years of part-time commitment. Through the preceding lessons, the teachers were already acquainted with other educational robots, namely the Thymio and the Lego EV3 robots, and their respective programming interfaces. All teachers gave their informed consent to participate in the study and to being video-recorded during the final interviews.

Study protocol

Due to the restrictions on personal meetings during the Covid-19 pandemic situation, the entire study was administered virtually using video recordings and online conference software. The main idea was to guide the teachers through different activities with the CreroBot, from its physical assembly to the programming of an autnomous navigation behavior. A kit with all necessary materials was delivered to the participants at their respective residences, which included the electronic components, cardboard material, printouts with CreroBot PaPL block icons, and an Android tablet. A set of video lectures was prepared by one of the researchers that presented the activity to the participants and guided them through the activities. Using the ERLS framework as a guidance, particular emphasis was put on the alignment of the intended outcomes and the distance learning activities with the affordances of the CreroBot and its programming interfaces. Specifically, the activities were framed by the concept of making, an approach that is often used with ER activities. By implementing making activities, we aimed at achieving instructional alignment with the do-it-yourself character of CreroBot and CreroBot PaPL. The activities, which in total were estimated to take around four hours to be completed, were structured in six incremental stages:

1. Introductory video lecture (5 min): In this video pre-recorded by one of the researchers, the participants were welcomed to the activities and the goals of the CreroBot project were presented.

2. Making the CreroBot and the CreroBot PaPL tiles (60 min): In this activity, the teachers had to prepare the physical elements with cardboard by cutting and pasting the material. The activity was guided by step-by-step video tutorials that were recorded by the researchers beforehand (Fig. 5.18)

3. Getting familiar with micro:bit and Makecode (40 min): In this activity, the teachers were asked to complete a set of simple programming tasks with the Makecode platform, such as controlling the on-board LED matrix, reading the values from the integrated compass and programming the robot to move in different directions on pressing the buttons. The activity was introduced by a short video lecture at the beginning of the activity.

4. Calibrating the motors (20 min): In this activity, the teachers had to calibrate the continuous micro-servo motors of the CreroBot, to render its movements as precise as possible. The activity was introduced by a short video lecture at the beginning of the activity.

5. Programming CreroBot with PaPL (30 min): In this activity, the teachers were asked to program the CreroBot with the CreroBot PaPL tiles and the Android application to make it move on grid cells. The activity was introduced by a short video lecture at the beginning of the activity.

6. Bonus task (30 min): Motivated participants were proposed to complete a final bonus task, in which they had to program the CreroBot that follows a specific cardinal direction, while autonomously correcting for deviations using readings from the integrated compass.

The teachers were given two weeks to complete the activities, while being able to organize their time at their own discretion. For all programming activities, basic code structure were provided to help the teachers get started. Moreover, all participants were instructed to contact the researchers in case of issues, whereupon one-to-one video conference sessions were organized. After the teachers completed all activities, final interviews were organized with one

Figure 5.18 – Screen captures of the tutorial videos for the construction of the CreroBot (left) and the CreroBot PaPL tiles (right).

of the researchers. The interviews were conducted individually using online video conference tools and lasted around 30 minutes each.

Measures

Semi-structured personal interviews were conducted using online video conference tools by the same researcher with each participant (in Italian). The interview questions were crafted to give the participants the freedom to express their views within a predetermined theme or context. The themes (and consequently the questions) eventually converged toward specific aspects of making the CreroBot and its usability for classroom use. All the interviews were recorded after permission was obtained from the interviewees and later transcribed. The interviews were then translated to English and all data was processed in accordance with appropriate research and ethics standards.

5.5.3 Results and discussion

In the following we will present the results from the CreroBot study with the teachers (herein after called T1-5), and discuss them with respect to two perspectives. First, we will discuss the specific use of CreroBot to implement ER activities for distance learning scenarios. Subsequently, we will focus on the general potential of CreroBot as an ER tool to introduce maker approaches in formal education.

CreroBot as a tool for distance learning

The idea to create the CreroBot was an indirect consequence of the restrictions imposed during the Covid-19 pandemic, which forced us to find alternatives to allow students at EPFL to continue their semester project as well as in-service teachers to complete the coursework of their continuous professional development program. Therefore, we decided to aim for the design of an accessible do-it-yourself educational robot, developed by the students and then tested with the teachers. The use of low-cost components as well as easily available materials for the body of CreroBot allowed the students to develop the system from home, without the need for sophisticated tools. Likewise, the teacher could participate in the study by receiving a little parcel which included all the materials they needed to build their own CreroBot. The special circumstances also emerged as a possibility to use the ERLS framework in a new context, namely that of distance learning. The development of CreroBot, the definition of the intended outcomes and the choice of the instructional and assessment activities had to be aligned to fit this new format of instruction. To leverage the do-it-yourself character of the system, we decided to frame the activity in the context of the maker approach, an idea which is inherently closely related to ER activities. Indeed, Eguchi (2017) has described ER as the pioneer of the maker approach. The intended outcome of the study was therefore, to introduce the teachers to the maker approach using the CreroBot system. The idea of making builds strongly on constructionist and constructivist ideas and they often embody project-based and self-regulated learning approaches. In such approaches, the teacher usually assumes the role of a tutor or mentor, rather than the one of a traditional lecturer. To align with these ideas, we decided to propose the activities in the framework of a mini-project, for which the teachers had two weeks of time to complete them. The activities were supported by video tutorials that provided the teachers with some guidance throughout the activities. Moreover, participants always had the possibility to contact one of the researchers to ask for individual help (this support was provided to two of the five teachers). As a way to assess the outcomes of this activity, we decided to organize final video interviews with each participant. Video interviews provided a feasible way to get in-depth insights into the individual experiences of the participants, especially in light of the restrictions imposed during the Covid-19 pandemic. The results from the interviews showed, that overall, the ERLS designed using the CreroBot appeared to be a suitable for implementing distance learning activities with ER. All teachers participating in the study mentioned that they enjoyed the process of building the CreroBot, programming it and using CreroBot PaPL to command the movements with a high degree of excitement. Some teachers displayed their enthusiasm by also attempting the bonus challenge. The participants highly appreciated the proposed activities and acknowledged that it was personally useful to them, as for instance, stated by the following teachers:

> *"This experience has concretely helped me to understand that there are easily structured making activities, that you can do even without great experience in assembling the components."* (T5)

> *"I don't consider myself to be an inventor of new activities, but when I see them, ideas pop up. [...] This was great in the context of this activity."* (T4)

CreroBot as a tool to introduce the maker approach in formal education

The individual interviews with the teachers also illustrated that CreroBot has potential as an ER tool to introduce maker approaches in formal education. One main advantage of the system appears to be its ease of use, allowing users to have a positive experience in autonomously building a robot with few components and accessible materials, as highlighted by one of the teachers:

> *"I have found it simple enough to assemble ... it made me curious ... I liked when it worked ... I speak for myself, for primary school, it is very important for the students to do something manual."* (T2)

> *"It brings together the advantages of various systems. It embodies the tangible programming like the Bluebot, and to a certain extent, the functionality and movements of the Thymio. Moreover it also gives you flexibility, as it could be used to build what you want. So in my opinion it combines several positive aspects in one system, that is moreover, easy to use."* (T2)

With respect to the systems they have already worked with, all teachers identified personalizability as one the most interesting affordances of the CreroBot. Furthermore, the use of the micro:bit as the hardware base of the CreroBot enables the implementation of many different learning activities, an aspect that was appreciated by another teacher:

> *""Once I went to look around a bit, I saw that with micro:bit the students can really do many things, for instance, make a compass. [...] One moment someone is using it like this and in the next moment, for something else. All components are such that one can disassemble and redo other things. You can use the same microchip to do many things."* (T5)

The same teacher highlighted that following a maker approach for ER activities could be considered favorable, since it is a concept that teachers and students are generally familiar with, even though it is named differently:

> *"I liked building the robot ... I think instead of just programming, it is good that we also build a robot. You call it making. I call it learning-by-doing, which is an idea often applied in schools"* (T3)

Indeed, all teachers acknowledged the potential of maker approaches using CreroBot to better understand underlying mechanisms of the system. Not only did it help the teachers themselves to understand what is inside a robot, they also believed that this learning experience could be beneficial for their students:

> *"At the elementary school level, I find it interesting to make students understand what's inside the robot. And with middle school, this is even more important."* (T2)

> *"In my opinion, it is interesting from the point of view of the construction. Thoroughly understanding how it is built, how the robot works. I personally have a 3D printer that I built and I'm glad I did, because I understand every piece well, what it's for. [...] I am of the idea that in any case, it is important not only to press the buttons and make something happen but also understand what is behind [...] I think it is something that is important to know, therefore, making can certainly help you to get a better understanding, in particular for the physical and mechanical parts, not only the logical part."* (T2)

Better understanding the underlying mechanisms also encouraged the teachers to find alternative materials and designs to improve the structure of their CreroBot:

> *"These small imperfections, which I guess are typical of the making approach, have stimulated me to look for different solutions. For example, to search for alternative materials for the construction."* (T4)

> *"I made this modification ... I used a toothpick to stabilize the structure of the robot. Also I thought that maybe other materials could be used, like the ones the children work with in arts classes."* (T1)

As a matter of fact, all teachers introduced some modifications to the originally proposed design to make their CreroBot more robust or functional. Some teachers used different types of cardboard to ease the construction, while others introduced some changes to the design of the body to make it more robust. With respect to the potential use of CreroBot in formal education, the teachers highlighted the aspects of multi-disciplinarity, facilitating for instance collaborations with arts teachers:

> *"In short, at the primary school level it would be interesting to put it as a collaborative project with arts teachers"* (T3)

> *"You could collaborate with colleagues who teach arts with recycled materials. Our children, they are already used to working with recycled materials."* (T5)

However, some teachers also raised some concerns about the feasibility of the approach. This concerns for instance the construction part, which some teachers thought would take too much time, especially for younger children. Moreover, concerns were raised about the use of craft knives which are required to cut out the cardboard pieces. One teacher proposed as a solution to precut the pieces and provide them to the students. Another teacher suggested that CreroBot could be proposed as a joint project with older students that could handle these kinds of tasks:

> *"The work in small groups could be useful to integrate the different skills of the students, for instance, technical, creative or manual, improving the self-esteem of each. I think it could also be useful to make students of different ages and with different levels of interest and skills interact with each other."* (T5)

This could, as a consequence, help students in exploring new interests:

> *"I can see a creative student starting with a focus on decorative tasks during the construction, but then in the course of the activity taking an interest also in the technical aspects."* (T5)

Finally, some teachers emphasized that it is important to fix some of the technical issues that were encountered before moving on to classroom studies. Indeed, the study helped to identify a few of such issues, that were not known before the study. Such issues included for instance, connection issues between the CreroBot and the Android application or crashes of the CreroBot PaPL app. Nevertheless, in summary, the teachers acknowledged the educational potential of the devised CreroBot system. As a matter of fact, three out of the five teachers have already asked whether they and their classes could participate in future classroom studies. This exploratory study therefore provides some initial evidence that CreroBot could serve as an interesting ER tool for classroom use, especially with regard to the introduction of the maker approach in formal education. The use of accessible materials as well as its simple design allows to leverage many facets of making without the need for sophisticated technologies such as 3D-printers or laser-cutters. It may therefore provide an important stepping stone for maker approaches, allowing to introduce the basic principles of making in formal education, at least to a certain extent. However, it should be acknowledged that the results of this study are of a very preliminary nature and have to be confirmed in larger studies, that particularly need to involve students interacting with the system. In this context, it would be also interesting to explore how far CreroBot can be leveraged as a tool for making education and when the transition to more sophisticated tools should be considered.

5.6 Conclusion

In this chapter we have discussed two examples of ER tools and outlined how they have been designed to be aligned within the ERLS framework. The first example described the development of the Thymio PaPL platform, a tangible programming interface for the educational robot Thymio that is based on accessible materials such as paper and cardboard. The development of the platform was performed in several steps, which to a certain extent can be considered similar to the alignment-centered design approach proposed by Lauwers (2010). However, in this project, there were no possibilities to involve educators in participatory design and therefore, the initial constraint finding process had to be done by the developers alone. In this context, the HERLS heuristics appeared to be a useful guiding tool, that helped the developers in identifying both usability and utility issues in their systems before the next design iteration. The revised version of the Thymio PaPL platform was finally tested in a teacher-run pilot test with two teachers and their classes. The results obtained from this study illustrated the potential of the platform to provide a smoother entrance to programming activities with Thymio as compared to its screen-based counterpart. Moreover, the use of Thymio PaPL was appreciated by the teachers as way to facilitate classroom orchestration. The results of this pilot study therefore provide some promising findings with regard to the potential of Thymio PaPL as an ER tool for classroom activities, that may motivate further studies in this direction.

In the second example, we presented the development of CreroBot, a do-it-yourself educational robot, relying on low-cost components and constructed based on a cardboard structure. The idea for the project originated from the restrictions imposed during the Covid-19 pandemic and it emerged as a possibility to explore the use of the ERLS framework also in the context of distance learning. The CreroBot was devised with the aim of using it in a distance learning activity with in-service teachers to introduce them to the idea of making. The CreroBot, the intended outcomes, as well as the instructional and assessment activities were devised using the ERLS framework to align the distance learning activity. The results of the study showed, that the devised activities were much appreciated by the participating teachers and helped them to discover the basic ideas of making. Moreover, the individual interviews with the teachers showed that CreroBot may also represent an interesting ER tool to introduce the maker approach to formal education. However, it should be acknowledged that these results are of very preliminary nature and larger studies are needed to consolidate them.

6 Alignment of ER classroom activities

Disclaimer
The content of this chapter has been adapted from the following works - with permission of all co-authors and publishers:

- Chevalier, M., Giang, C., Piatti, A., & Mondada, F. (2020). Fostering computational thinking through educational robotics: a model for creative computational problem solving. *International Journal of STEM Education*, *7*(39). https://doi.org/10.1186/s40594-020-00238-z

  As one of the two equally contributing main authors of this publication, my contribution to this work involved: conceptualization, methodology, formal analysis, investigation, data curation, visualization and writing - original draft preparation.

- Giang, C., Chevalier, M., Negrini, L., Peleg, R., Bonnet, E., Piatti, A., & Mondada, F. (2018). Exploring escape games as a teaching tool in educational robotics, In *International conference edurobotics 2016*. Springer. https://doi.org/10.1007/978-3-030-18141-3_8

  As the main author of this publication, my contribution to this work involved: conceptualization, methodology, formal analysis, investigation, data curation, visualization and writing - original draft preparation.

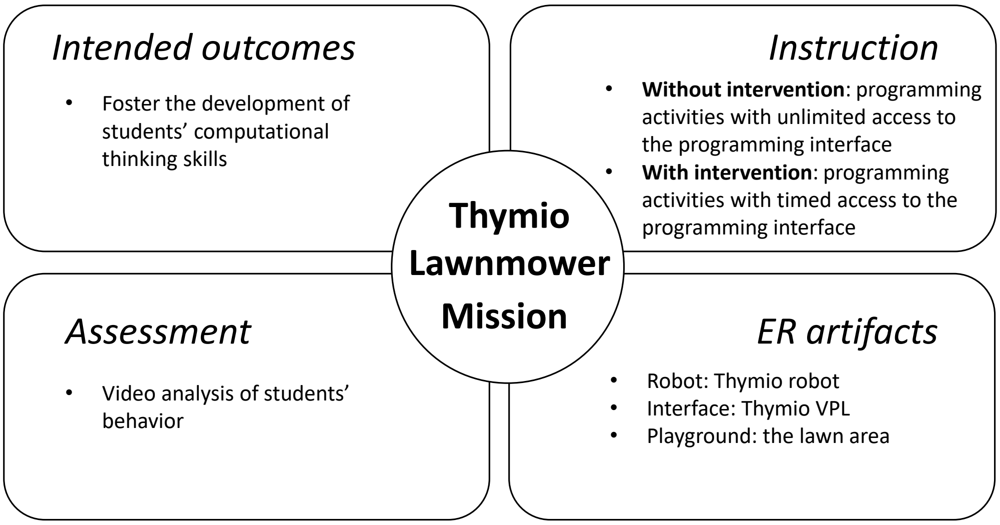

Figure 6.1 – ERLS implemented in the classroom study conducted with the Thymio lawnmower mission.

6.1 Summary

In this chapter, we will present two examples that illustrate how alignment can be attained by appropriately adapting classroom activities to the learning goals and the affordances of ER tools. The ER tools used in both examples are the Thymio robot and its graphical programming interface VPL.

The first example is the Thymio lawnmower mission, a learning activity designed to promote the development of students' computational thinking (CT) skills. In a classroom study with primary school students we will illustrate the effects of a misalignment of classroom activities with the learning goals and the affordances of ER tools. Moreover, we will present how simple instructional interventions can be applied to remedy this situation, leading to a better aligned ERLS and consequently, more favorable learning experiences. A summary of the ERLS implemented for this classroom study is depicted in Fig. 6.1.

The second example presented in this chapter is the Thymio Escape Game, an immersive ER activity, that has taken inspiration from escape room experiences. In the following, we will outline how the escape game format may represent a favorable approach to attain alignment of ER classroom activities with the learning goals and the affordances of ER tools. Two iterations of the Thymio Escape Game activity will be discussed: first, we will introduce a prototype version, which was developed without deliberate aim for classroom use. Subsequently, we will present a revised version, specifically developed as a classroom activity and redesigned based on the findings from a heuristic evaluation with HERLS. The final version was tested in a user study with students to evaluate whether it can increase their interest, attitude and self-efficacy related to robotics. A summary of the ERLS implemented in this study is depicted in Fig. 6.2.

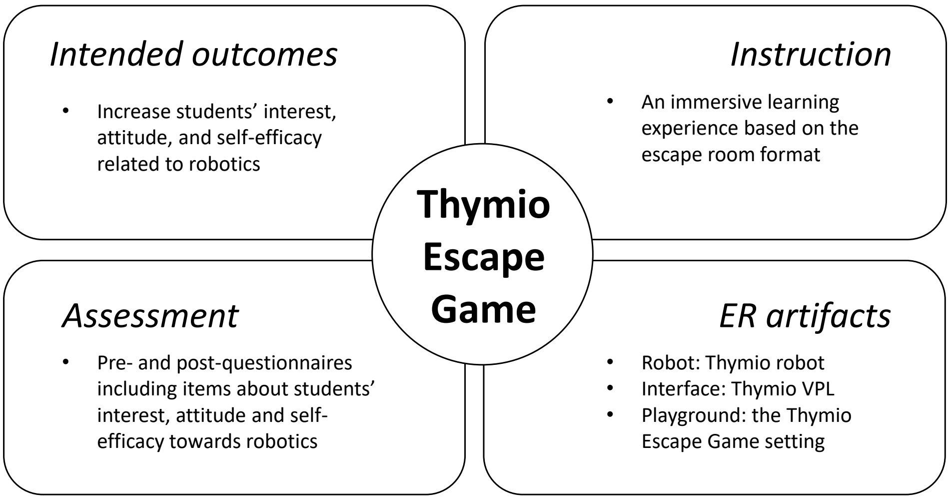

Figure 6.2 – ERLS implemented in the study conducted with students to test the classroom version of the Thymio Escape Game.

6.2 Aligning the Thymio Lawnmower mission

In this section we will discuss the Thymio Lawnmower mission, an ER classroom activity developed to foster the development of students' computational thinking (CT) competencies. In recent years, teaching CT competencies in formal education has become more and more a topic of interest and in this context, ER activities have been suggested as an promising approach (Atmatzidou & Demetriadis, 2016; Chen et al., 2017). However, so far, only few works (Bers et al., 2014; Sullivan et al., 2017) have discussed how ER activities should actually be designed and implemented to effectively foster the development of CT competencies. Consequently, ER activities intended for this purpose, may not succeed in eliciting the desired learning effect, due to a misalignment of the ER classroom activities with the learning goals and the affordances of the used ER tools. In the following, we will use the example of the Thymio Lawnmower mission, to illustrate the consequences of such a misalignment and we will present how a simple instructional intervention has been applied to remedy the situation, resulting in a better aligned ERLS and consequently, more favourable learning expriences.

6.2.1 Description of the activity

In the Thymio lawnmower mission, small groups of two to three students are asked to program a Thymio robot, that simulates an autonomous lawnmower behavior. The playground of the Thymio lawnmower mission consists of a fenced lawn area of 45 cm x 45 cm size (Fig. 6.3). The fence is constructed using wood, and the lawn area is represented by eight squares of equal size with an imprinted lawn pattern. A ninth square is imprinted with a brick pattern and placed at the bottom right corner of the area, representing a garage, i.e., the starting point of the Thymio lawnmower robot.

Figure 6.3 – Playground of the Thymio lawnmower mission. Starting position (left) and a programmed robot executing the mission (right).

In this ER classroom activity, the students have to program a lawnmower behavior, which autonomously drives the robot out of its garage and in the best case, makes it pass over all eight lawn squares while avoiding any collision with the fence. The interest of using the Thymio robot to carry out this mission is twofold: on the one hand, this robot has many sensors and actuators providing students with various possibilities to approach the mission. On the other hand, among the different programming languages that can be used with Thymio, one is the block-based graphical language VPL (Shin et al., 2014). The VPL programming platform (cf. Fig. 5.4) represents parts of the robot's language by graphical icons that can be directly interpreted by the students, hence providing an iconic representational cognitive artifact. This facilitates the programming of the robot, since it does not require students to learn complex syntax beforehand. Students can implement their solutions by simple drag-and-drop actions, matching event-blocks with one or multiple action blocks. However, in contrast to sequential programming languages, the robot cannot simply be instructed to move a certain distance towards a given direction. Instead, in the event-based programming language VPL, students have to reflect on how to use the robot's sensors and actuators to generate a desired behavior. The openness and uncertainty of the task thus requires the students to leverage many competences related to computational thinking, which is the main learning goal of this activity.

6.2.2 Study design

Using the Thymio lawnmower mission, an experimental study was conducted to explore the effects of a misaligned ERLS and to study possible instructional interventions to remedy the situation. Therefore, the activity was performed by two groups of primary school students under varying conditions.

Participants

A total of 29 primary school students (13 girls and 16 boys between 9 and 10 years old) participated in the experimental study. Prior to the study, all students have been introduced to the Thymio robot and the VPL programming interface through several school lessons (one hour per week for twelve weeks). The participation of the students in this study was approved by their guardians (parents) and class and school leaders (teachers and principal). A statement on ethics approval and consent was issued by The Coordination Committee for Educational Research in the Canton of Vaud (Switzerland).

Study protocol

At the beginning of the experimental session, all students were randomly assigned to groups of two or three. Each group of students was then randomly assigned to one of the two experimental conditions (control or test). The experimental procedures for the groups in each condition were different:

- **Control groups.** The activity for the control groups started with a short introduction, where the goal and the rules of the mission were explained by one of the experimenters. The students were then given 40 minutes to implement their lawnmower robot. During the whole time period they were allowed to use all ER artifacts that were provided to them: the playground, the Thymio robot and the VPL programming interface. No additional constraints were imposed on the students.
- **Test groups.** The experimental procedure for the test groups differed in the structure of the activity. Following the introductory speech, the activity started with 10 minutes of blocking of the programming interface. The students were given access to the playground and the Thymio robot, but they were not allowed to use the VPL programming platform. After this phase, the blocking was released, and the students were allowed to use all ER artifacts for 10 minutes. This was followed by a partial blocking phase of 10 minutes, where the students had access to everything including the VPL platform, but they were not allowed to execute any code on the robot. For the last 10 minutes, the blocking was released again, and the students were allowed to use everything that was provided to them.

The study was conducted in two consecutive sessions of 45 minutes, one for each experimental condition. The test group (7 girls and 8 boys) started the mission first, while the control group (6 girls and 8 boys) went on a guided museum exhibition. After the completion of the first session, both groups switched. During each session the five groups of the same experimental condition worked on the Thymio lawnmower mission simultaneously. Each group was provided a playground, a Thymio robot and a computer with the VPL platform installed. The sessions were supervised by two experimenters who provided technical support and addressed the students' questions regarding the task assignment. However, the experimenters did not provide any support regarding the solution of the lawnmower mission. Each group, as well as their interactions with the VPL platform and the playground, were recorded on video for later analysis.

Video analysis

Based on a socio-constructivist approach, this study relied upon in situ observations to capture the interactions of the students with each other as well as with the different ER artifacts (the robot, the interface and the playground). Therefore, the videos recorded from the experimental sessions were analyzed in several steps.

Prior to individual analyses, the two principle experimenters (herein after called evaluators) met to discuss and agree on appropriate observables indicating transitions of students towards the different phases of the CCPS model (Chevalier et al., 2020). It describes students' cognitive processes related to CT competencies in ER activities as a temporal model of six phases: understanding the problem (USTD), generating ideas (IDEA), formulating the robot's behavior (FORM), programming the behavior (PROG), evaluating the behavior (EVAL) and off-task behavior (OFFT). At this point, we refrain from a in-depth discussion of the model, since it is considered to be out of the scope of this thesis. The interested reader is encouraged to read the work presenting the model (Chevalier et al., 2020). For the remainder of this section, we would only like to emphasize, that according to the definition of the model, it is possible for students to transition from any phase to another at any moment of the activity.

To identify appropriate observables (visual and verbal), the evaluators first analyzed various prerecorded ER activities together. The videos were recorded from different kinds of ER activities and allowed to establish criterion standards (Sharpe & Koperwas, 2003) that are not limited to one specific ER activity. The whole procedure was aimed at streamlining the way both evaluators would perform their individual analyses.

Subsequently, both evaluators performed the behavioral analysis independently, sequentially mapping the behaviors of the students during the robot lawnmower activity to the different phases of the CCPS model. The mappings were made under the assumption that a student can only be in one of the six phases of the CCPS model at a time. Each evaluator performed the mapping based on their interpretation of the behavior of the students, while considering the criterion standards that have been established beforehand. Transitions to the first three phases of the CCPS model were mainly mapped based on the students' verbalizations, such as *"How can we do that?"* (USTD), *"Ah, I have an idea!"* (IDEA) or *"If this sensor detects the wall, the robot turns left"* (FORM). In contrast, transitions to the last two phases were mostly based on visual observations (e.g. a student starting to use the computer (PROG) or a student watching the Thymio robot after executing the program (EVAL)). Students who were clearly not involved in the activity were mapped to the off-task behavior phase (OFFT). Two state graphs were created for each student (one by each evaluator) using a software dedicated to

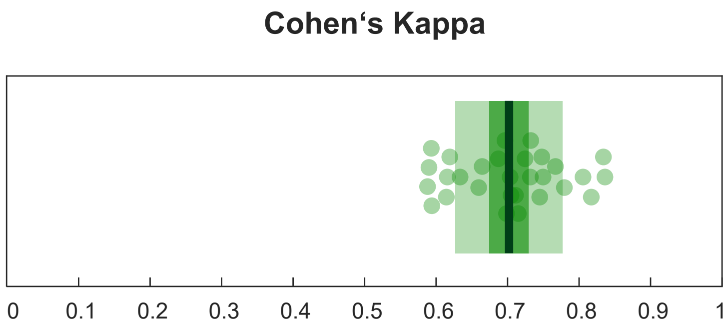

Figure 6.4 – Cohen's kappa values for inter-rater reliability computed for each student based on the independent observations of each evaluator. Dots indicate Cohen's Kappa values calculated for each student based on the independent observations of two evaluators. The dark vertical line indicates the mean value, dark gray areas one standard deviation, and light gray areas the 95% confidence intervals

the creation of activity chronicles such as Actograph (SymAlgo Technologies, Paris, France) and a numerical computing tool such as Matlab (MathWorks, Natick, Massachusetts, USA).

Following this step, both evaluators compared their state graphs against each other and discussed any major discrepancies between their evaluations. Major discrepancies were considered segments in the state graphs in which both evaluators did not agree on the same behavior for more than one minute. The corresponding video scene was reviewed by both evaluators together to achieve a mutual decision. Based on this decision, the state graphs of the evaluators were modified accordingly.

Subsequently, the continuous state graphs of both evaluators were discretized into equally spaced time segments of one second. Cohen's Kappa was computed for the discretized pair of state graphs of each student, in order to validate the inter-rater reliability of the performed videos analyses. Therefore, confusion matrices were created for the observations made by both researchers. Agreement between both evaluators were quantified by the number of times both evaluators agreed on mapping the same phase of the CCPS model to a student's behavior. The Kappa values were then calculated for the observations made for each student, using the formula presented in the work of Bakeman and Gottman (1997) and taking into account the proportion of agreement observed and the proportion expected by chance. The range of the values for Cohen's Kappa was $0.59 < k < 0.84$ (Fig. 6.4), which according to Landis and Koch (1977) can be interpreted as substantial agreement between both evaluators. Finally, the reviewed state graphs created by the first evaluator were used to compute the time spent in each state of the CCPS model by each student.

6.2.3 Results

When analyzing the total time spent in each phase of the CCPS model (Fig. 6.5), it was observed that in the first three quarters of the activity, students of the control group predominantly spent

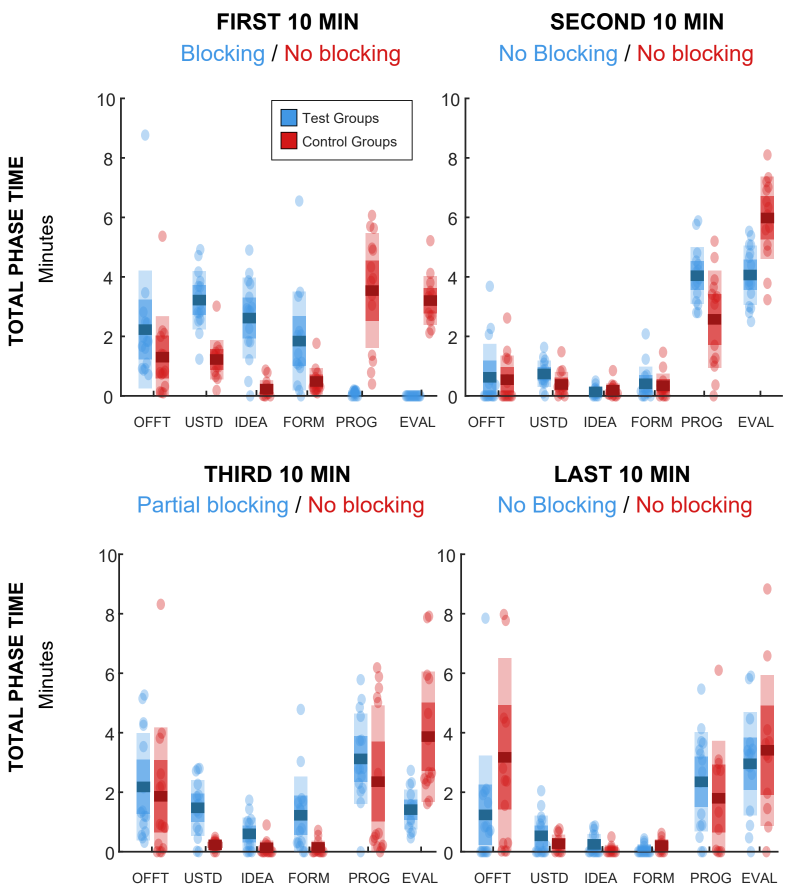

Figure 6.5 – Total time spent in each phase by both groups for the four quarters of the activity. Colored dots show data points for each student of the test (blue) and control (red) groups. Dark horizontal lines indicate the mean values, dark-colored areas one standard deviation, and light-colored areas the 95% confidence intervals

their time in PROG and EVAL phases (on average 22 out of 30 minutes). In contrast, USTD, IDEA, FORM and OFFT phases were observed less frequently (8 out of 30 minutes). Moreover, in the last quarter of the activity, PROG and EVAL remained more prevalent compared to USTD, IDEA and FORM phases, however, a similar amount of time was now also spent on off-task behavior (OFFT, 3 out of 10 minutes). On the other hand, students from the test groups displayed a more even distribution among the phases for the first three quarters of the activity. On average, students spent 13 out of 30 minutes in PROG and EVAL phases and 12 out of 30 minutes in USTD, IDEA and FORM phases. Moreover, during the first three quarters, the times spent on off-task behavior (OFFT) by the test groups was very similar to the ones by the control groups. However, in contrast to the control groups, for which an increase of off-task behavior was observed in the last quarter, test groups displayed a similar level of off-task behavior until the end of the activity.

The performance of each group's lawnmower was quantified by the highest number of lawn squares that the robot managed to cover without collision. The results showed that three groups (two test and one control) managed to complete the task, covering all eight lawn squares with their lawnmower robot. Five groups (three test and two control) covered six squares and two groups (both control) covered only 4 squares. The number of squares was only quantified for trajectories that started from the garage and hence were not based on random movements.

In order to illustrate the dynamics at individual levels, the state graphs for one exemplary student of each group are presented (Fig. 6.6). The data depicts the activities of each student during entirety of the 40 minutes. It can be observed that the student from the control group immediately started the activity by jumping into the PROG phase. Throughout the activity, the student spent most of the time only in PROG and EVAL phases, sporadically transitioning to one of the other phases, that were then followed by transitions back to the PROG-EVAL loop. The student from the test group on the other hand, showed a more balanced distribution among the five main phases of the CCPS model. From the state graphs it can also be observed that the student from the test group also performed more playground interactions (11 times), i.e., interactions with the robot or the lawn area, compared to the student from the control group (2 times). A similar result was observed when analyzing the overall data for playground interactions of each experimental group (in total 93 interactions for the test groups and 59 interactions for the control groups).

6.2.4 Discussion

In this study we have studied how two different implementations of the same ER activity can have different effects on students' cognitive processes related to computational thinking (CT). For students in the control group, the Thymio lawnmower mission was proposed without any constraints, allowing the groups to freely work on the mission based on their personal planning. Based on the video analysis, it was observed that almost all groups working under this condition spent most of the 40 minutes in activities related to the programming of the Thymio robot. Plunged in a programming-evaluating loop, students mostly followed a blind trial and error approach, neglecting other cognitive processes required to successfully solve the mission. However, as Shute et al. (2017) have suggested before, "considering CT as knowing how to program may be too limiting". Indeed, to successfully accomplish the Thymio lawnmower mission, students are also required to appropriately understand the problem situation, generate idea solutions and formulate the robots desired behavior before programming. However, the unconstrained organization of the activity encouraged the students to directly dive into the programming of the robot without a proper planning of their steps. With respect to the main learning goal of this activity, i.e., the development of CT competencies, the outcome of such approaches may therefore be questionable. Moreover, we noticed that many students in the control groups started to become frustrated and bored, after a series of unsuccessful trials. Those students showed an increase of off-task behavior in the last quarter of the activity, an effect that is certainly not desired from a pedagogical point of view. The activity as performed with the control groups therefore illustrates how a misalignment of the classroom activity

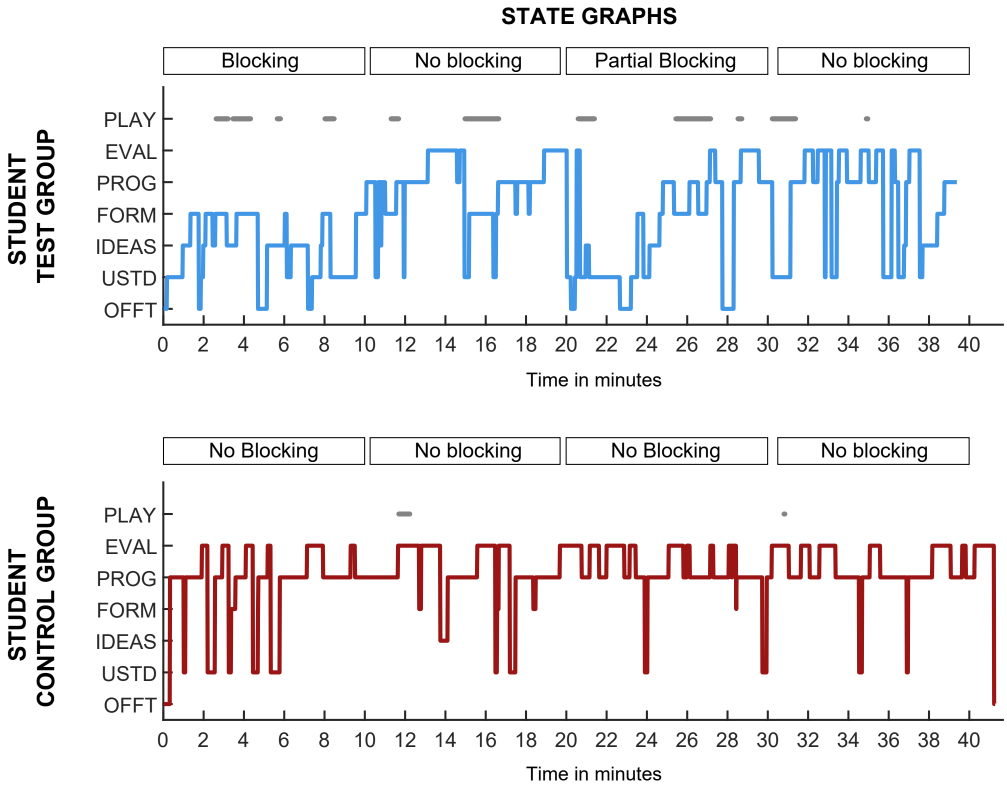

Figure 6.6 – State graphs for two example students. The figure shows the data for two students, each exemplifying what was observed in the test and control groups, respectively. It depicts the complete state graphs for the student from the test (first row) and the control group (second row), displaying in which phase each student was at each moment of the activity. Moreover, the students' interaction with the playground is highlighted (PLAY).

design with the affordances of the used ER tools can hamper the achievement of desired learning goals (in this case the development of CT competencies). Indeed, student mostly interacted with the programming interfaces, not making use of the affordances provided by the robot and the playground. Moreover, even their interaction with the VPL platform may not be considered optimal, since students mostly followed a blind trial and error approach, not being able to take advantage of one of the main affordances provided by the interface, i.e., the close representation of the robot's properties by the iconic programming blocks.

For the test groups, the same Thymio lawnmower mission was conducted with relatively minor adaptations, which however, appeared to have a strong effect on how students approached the activity. Instead of allowing students to freely access all ER artifacts available, some temporary constraints were imposed on the groups. In the first ten minutes of the activity, students were not allowed to use the VPL programming interface. This constraint encouraged students to shift their attention away from the programming of the robot, drawing it to the cognitive processes required prior to the programming step. Indeed, we observed that students made increased use of the ER artifacts available (i.e., the robot and the playground) to thoroughly explore and generate different solution ideas. Based on these ideas, students were able to fully harness the programming interface in the subsequent ten minutes, when the access to VPL platform was released. With more elaborated ideas in mind, students showed more targeted approaches with respect to the programming of the robot. However, to prevent them from eventually diving into the blind trial and error approach as observed from the control groups, another instructional intervention was introduced half way through the activity. After ten minutes of being able to use the VPL programming interface, students were instructed to not load any new programs to the robot. While still being able to see their last programs loaded to their robots, this partial blocking of the VPL platform encouraged students in thoroughly debugging their programs. Indeed, debugging has been considered an important activity related to CT (Shute et al., 2017), and this intervention provided students with a possibility to reflect on their current implementations in order to identify necessary modifications. When all constraints were released for the last ten minutes of the activity, we observed that students in the test groups were still as engaged as in the first three quarters of the activity and in contrast to the control groups, no increase in off-task behavior was observed. The two instructional interventions introduced at the beginning and half way through the activity, appeared to allow for a better alignment of the Thymio lawnmower mission with the affordances provided by the robot, the playground and the programming interface, eventually providing better learning experiences to the students and facilitating the achievement of the desired learning goal. Not only did students in the test groups display more cognitive processes related to CT competencies, they also tended to be more successful in accomplishing the Thymio lawnmower mission.

The example of the Thymio lawnmower mission illustrated how simple instructional interventions can make significant contributions to the alignment of classroom activities with the affordances of ER tools. Indeed, the temporary blocking of the programming interface is an approach that is not limited to activities involving the Thymio robot, but it can be also applied to many other ER activities. The interventions allowed students to capitalize on the affordances provided by the robot and the playground, helping them to develop more structured approaches to solve the Thymio lawnmower mission. Moreover, the temporary blocking of the programming interface encouraged students to make more efficient use of the interfaces' affordances when the access was released. While this study only described one example to effectively adapt an ER classroom activity, other types of interventions can be considered. In the following we will present a second example of a ER activity, in which multiple adaptations have been made to allow for an effective classroom use of an ER activity.

6.3 Introducing the Thymio Escape Game

This section describes the development and evaluation of the first prototype of the Thymio Escape Game. We first introduce the concept of educational escape games and discuss why this format might particularly be an interesting approach for the alignment of ER classrooms activities with the affordances of ER tools. Subsequently, we will present the different elements and the structure of the activity. The first prototype was then tested with 61 volunteers, who evaluated the playability as well as the educational potential of the activity.

6.3.1 Adopting the escape game format to ER classroom activities

Entry into and adherence to learning has always been a recurrent theme in education. In this context, learning motivation has been widely discussed (Deci et al., 1991; Vallerand et al., 1992) and game-based learning (GBL) approaches, capitalizing on the engaging character of games, have become more and more popular. In the past decades, particular attention has been devoted to digital GBL, a field popularized by Prensky (2003), aiming at integrating educational elements into computer games and/or digital simulations. Researchers and educators have agreed on the potential of such solutions, used for both instructional teaching (Tobias et al., 2014) as well as for the exploration of new topics (Steinkuehler & Squire, 2014; Whitton, 2014). Nevertheless, GBL also encompasses non-digital examples, such as educational card, board or role-playing games, that allow the implementation of more tangible and human-centered activities. In recent years, educational escape games have become increasingly popular (Clarke et al., 2017) and there have been attempts to exploit this concept for educational purposes (Nicholson, 2015).

In escape games, players usually team up to discover clues helping them to solve a series of tasks and puzzles. The game is typically (but not necessarily) played in a room locked from the outside, and the main objective is to find a way to escape from it by solving the presented puzzles (this is the reason why they are often also referred to as escape rooms). Often, the game is embedded in a narrative and the tasks have to be solved within a given time limit (typically between 30-60 minutes). Besides the captivating character of escape games, which puts the players into an immersive problem-solving situation, the need for collaboration and communication to solve the puzzles as a team has been considered as a main appeal (Nicholson, 2015). From a pedagogical point of view, escape games are a methodology based on a social-constructivist approach (Vygotsky, 1980). The learner is called to face new, rather difficult problems, which can be solved thanks to the interaction with peers and the support of the teacher. The role of the teacher in this approach is to structure the learning environment and provide instructional scaffolding to the learners, "facilitating the students' interaction with the material and with each other", as described by King (1993). Moreover, very often ER activities are proposed to small groups of students, prompted to collaborate and communicate to solve a given problem situation. Due to these aspects, educational escape games have gained attention as a new way of instruction, especially in the context of problem-based learning.

As a matter of fact, there have been previous approaches to use educational escape games as a mean of instruction. For instance, Vörös and Sárközi (2017) developed an escape game to teach the physics of fluids to a group of gifted students, while Kinio et al. (2017) devised an escape game for the surgical education of medical students. Both works reported high engagement and enjoyment on the part of the participants and suggested escape games to be a promising approach to break out from the traditional classroom routine. Similar results were found by Hou and Chou (2012), who developed a digital escape game (i.e., it was played on a computer) to teach the basics of electromagnetism to high school students. In addition to high engagement, the authors reported that participants experienced flow - a state of mind, which has been considered as beneficial for successful learning (Shernoff et al., 2014). Nevertheless, no previous works have reported the adoption of the escape game concept in the context of ER. As presented in Chapter 3 of this thesis, socio-constructivism is, together with constructionism, one of the main learning theories guiding alignment in the ERLS framework. We therefore argue, that adopting the escape game format to create ER classroom activities may generate interesting and favorable learning situations, capitalizing on their common pedagogical grounds. To this end, we here present the prototype of an escape game using the educational robot Thymio and the graphical programming interface VPL. This exploratory study aims at investigating how escape games can be effectively used as a teaching tool, particularly in the context of ER.

6.3.2 A first prototype for the Thymio Escape Game

The development of the first prototype of the Thymio Escape Game was based on design guidelines proposed in previous works about learning games (Boller & Kapp, 2017) and educational escape games (Clarke et al., 2017). The main learning goal of this activity was to introduce the players to the Thymio robot and the graphical programming interface VPL. Participants played in groups of 3-5 people and had 30 minutes to solve the quest. The game started by bringing the group into a dark room and showing them a video to introduce the scenario: the local energy supply has been attacked and the task for the group was to reactivate the main energy source, a reactor powered by mobile robots. In order to accomplish this task, the players had to find three Thymio robots hidden in the room, each characterized by a different color (yellow, blue and red). After the robots were found, they had to be programmed using VPL. The goal was to make them move inside the three compartments of the battery, where they had to stop and lighten up in green (Fig. 6.7).

Each robot had to be programmed differently, in order to make it move from its starting position (indicated by colored spots on the ground map, cf. Fig. 6.7) into the compartments of the battery. The blue robot had to programmed to follow a black line using its two infrared ground sensors and stop inside the compartment when the line ended. Using the same ground sensors, the red robot had to be programmed to drive drive straight ahead and to automatically stop when it reached the black spot inside its compartment. The yellow robot instead, had to be programmed to drive straight on and to stop when its infrared front sensors detected the wall of its compartment.

Several hints for solving the tasks were hidden in the room, each one tagged with a color marker (yellow, blue or red, corresponding to the robot the hint was referring to). The hints

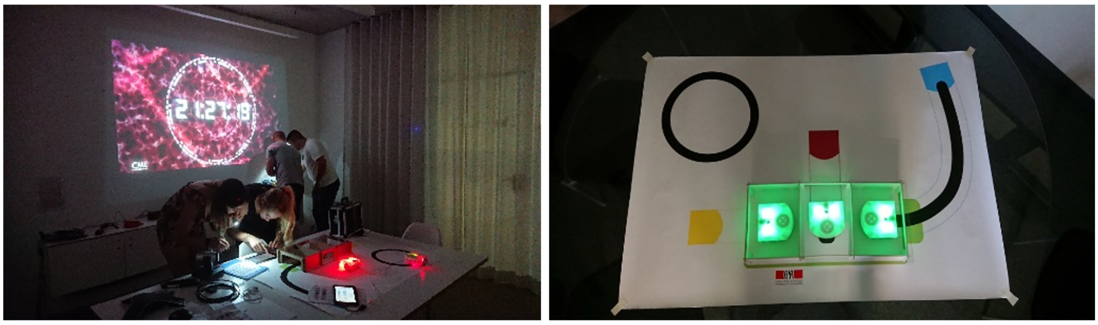

Figure 6.7 – Illustration of an in-game situation (left) and the ground map with the final winning state (right).

were placed based on the value of the information they contained (i.e., simple hints were easy to discover, while more valuable hints, e.g., the full solution of a program code, were well hidden). The players were instructed about the game rules beforehand (i.e., no use of violence, no manipulation of the battery, no placing of the robots by hand), but they were not informed about the operating principle of the battery. The winning state had to be discovered by themselves, supported by various notes hidden in the room. Moreover, each robot was already preprogrammed with an initial, incomplete behavior at the moment it was found: for instance, putting the blue robot on a black line, would make it follow it. However, if the robot was put on the line leading to the right compartment of the battery, it would eventually enter the compartment, shortly lighten up in green, but then turn and leave.

In order to win the game, the players had to add modifications to the program code of each robot. A countdown timer with visual effects and suspenseful music was used to convey a crisis situation. Each game session was supervised by an experimenter who took the role of the game master, providing further guidance to players. In order to not interfere with the players actions, the game master stayed outside of the playing area, and guidance was only provided if it was perceived as absolutely necessary. Moreover, to not interrupt the game flow, assistance was provided in a subtle way (i.e., by short comments) and clues were only given with reference to the hints that could be found in the room. If the group succeeded to place the three robots inside the compartments of the battery, a final video was launched, indicating that the task was accomplished successfully.

6.3.3 Study design

Participants

The game was tested with 61 participants (31 males, 30 females) of various age groups (Table 6.1) and varying prior experience in robotics as classified by the participants' self-assessment (Table 6.2). As the aim of this exploratory study was to study the perception of the game by people of various backgrounds (i.e., by a heterogeneous population), participation was

Table 6.1 – Demographics of study participants

Age group	Males	Females	Total
Under 15	2	1	3
15-24	5	14	19
25-44	18	13	31
45-64	5	2	7
Over 65	1	0	1
Total	31	30	61

Table 6.2 – Self-assessment of study participants with regard to prior experience with robotics

Experience	Males	Females	Total
Novice	6	19	25
Intermediate	9	8	17
Advanced	8	3	11
Expert	8	0	8
Total	31	30	61

not restricted based on age or experience. However, participants were only included in the study if they had none or limited experience with the Thymio robot and its VPL programming language. For the study, the participants signed up in groups, i.e., the members of each group knew each other before the experimental session. Participation was entirely voluntary, and all participants gave their informed consent to participate in the study. In the case of minor participants, informed consent was given by their parents or their legally authorized representative.

Study protocol

At the onset of an experimental session, the participants were briefly interviewed about their previous experience and their affinity to robotics and technology. The information gathered from this evaluation was used by the game master to determine the right balance of guidance in the activity. Subsequently, the participants were brought into the playroom and the game started. In contrast to most commercial escape games, the aim of this activity was to achieve at a high success rate, assuming that this would elicit a more positive learning experience. Therefore, the game master was given the possibility to provide extra guidance to the group, if it was perceived as necessary. The whole experimental session was followed by a second experimenter, who took the role of an observer to gather qualitative data by in-situ observations. The observer stayed in the background (i.e., in a separate area of the room, that was not accessible to the group) and did not interact with the players. The observations were recorded using written reports for later analysis. After the game, the participants were asked to complete a questionnaire and they were debriefed about the activity for around ten minutes.

The debriefing, a well-established concept to mentally process experiential learning activities (Dennehy et al., 1998; Nicholson, 2012), was moderated by the game master. The aim was to facilitate a discussion about the players' performed actions and their consequences, in order to emphasize the learning objectives of the game. Moreover, the participants had the possibility to ask questions about the Thymio robots, the VPL programming language and the escape game activity.

Measures

Quantitative data. To conduct the survey, a questionnaire was designed based on the one developed by Hou and Chou (2012). In their study, they used an evaluation questionnaire to determine the acceptance of their escape game by the participants. The user acceptance was grounded on two dimensions: perceived usefulness and perceived ease of use. Questions related to these two dimensions have been shown to be a reliable measure for the user acceptance of information technology (Davis, 1989). In their questionnaire, Hou and Chou also utilized the model proposed by Kiili (2006) to determine the flow state of their participants. For this study, their questionnaire was adapted to determine the usefulness, ease of use and the flow state perceived by the participants during the Thymio Escape Game. The study was conducted under the assumption that participants, and particularly novices, would distinguish between two possible uses of the activity: initial exploration and comprehension. Therefore, the dimension of the perceived usefulness included questions related to both sub-dimensions. The final questionnaire (Table 6.3) comprised eighteen statements related to the following dimensions: perceived usefulness for initial exploration (items EXP1-4), perceived usefulness for comprehension (CPR1-4), perceived ease of use (USE1-3) and perceived flow experience (PFL1-6). For the FLOW statement (*"I experienced a clear flow experience during playing"*), participants had to directly state whether they experienced flow during the game. In order to prevent ambiguities, a description of flow taken from the literature (Kiili, 2006) was presented in the questionnaire. The participants had to respond to each statement using a 4-point Likert scale *(Strongly disagree / Disagree / Agree / Strongly agree)*.

Qualitative data. Qualitative data was collected by means of three open-ended questions in the questionnaire (Table 6.3). Moreover, it was collected by means of in-situ observations made by a second experimenter who followed the entire experimental sessions (i.e., the introduction, the game and the debriefing). The observations were recorded using written reports and they were subsequently discussed among the experimenters in order to determine the significance of each observation.

Table 6.3 – Items of the questionnaire distributed at the end of the game

Item	Statement
EXP1	The game raised my interest in robotics
EXP2	The game raised my interest in Thymio and VPL
EXP3	I think the game allows me to better explore Thymio and VPL compared to conventional methods (books, lectures, etc.)
EXP4	The game helped me explore Thymio and the pro-gramming interface VPL
CPR1	The game helped me understand the basic concepts of robotics
CPR2	I think the game allows me to better understand Thymio and VPL compared to conventional methods (books, lectures, etc.)
CPR3	The game helped me understand how the sensors and actuators of Thymio work
CPR4	The game helped me understand the basics of the programming language VPL
USE1	The goal of the game is easy to understand
USE2	The rules of the game are easy to understand
USE3	I understand the educational goals of the game
PFL1	I felt in total control of my playing actions
PFL2	The way time passed seemed to be different from normal
PFL3	I really enjoyed the playing experience
PFL4	I found the experience extremely rewarding
PFL5	The challenge that the game provided, and my skills were at an equally high level
PFL6	I had total concentration while playing the game
FLOW	I experienced a clear flow during playing
Q1	About the game, I liked...
Q2	About the game, I did not like...
Q3	General suggestions

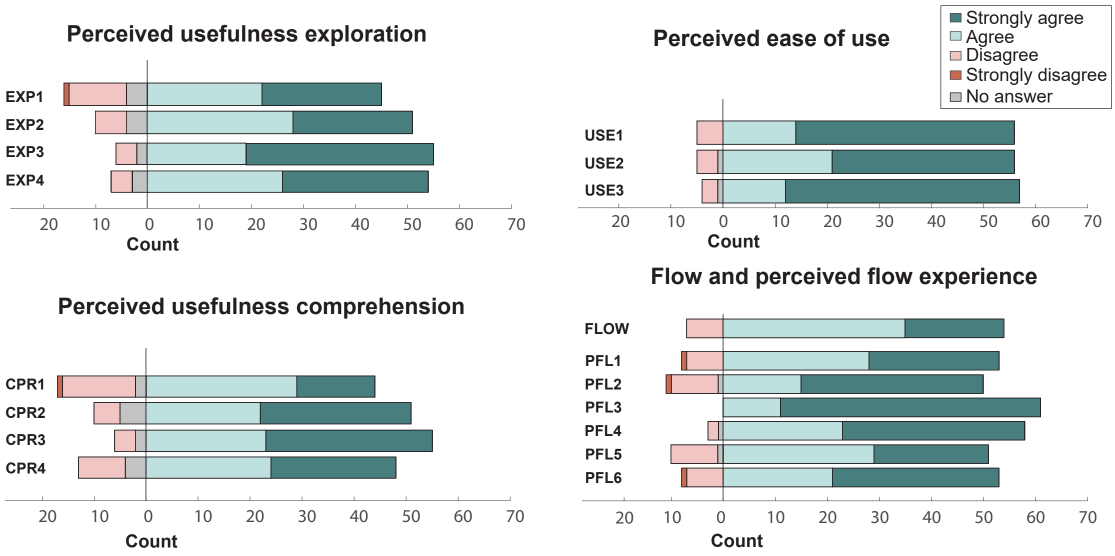

Figure 6.8 – Summarized results of the questionnaires by each dimension.

6.3.4 Results

Quantitative results

The analysis of the answers given in the questionnaires showed that overall, most participants highly appreciated the Thymio Escape Game (Fig. 6.8). On average, high approval was found for the EXP (84% agreed or strongly agreed), CPR (82%), USE (92%) and PFL (89%) dimensions. Moreover, 89% of the participants agreed or strongly agreed on the statement FLOW. However, it was also found that particularly statements related to robotics education (i.e., EXP1 and CPR1) found less approval: 21% and 18% of the participants respectively, disagreed on those statements. A more detailed analysis, showed that these answers were mostly given from people with robotics experience - a population for whom those statements are most likely not suited. From the open-ended answers of the questionnaire, it was found that for similar reasons, a few participants did not provide answers to some of the statements in the EXP and CPR dimensions. Similarly, participants who have already heard about Thymio and VPL beforehand or had limited experience with the robot, chose to not answer to the corresponding statements (i.e., EXP2-4).

A comparison between the groups with different robotics experiences (i.e., novices, intermediate, advanced and expert participants) was performed by analyzing the mean Likert scale values for each dimension. The analysis was done for males and females of each group (Tables 6.4-6.6). In general, similar trends were observed for all groups. The only noticeable differences were found for male experts, who, as mentioned before, agreed less on CPR statements than the rest - and more interestingly, also for female novices who agreed less on CPR and FLOW statements. The results illustrated that although the FLOW item was less approved by females in all groups, it was particularly different for female novices: 32% (6 out of 19) of

Table 6.4 – Mean Likert scale values and standard deviation for novices (n=25)

Dimension	Males (6)		Females (19)	
	mean	*sd*	*mean*	*sd*
EXP	3.38	0.77	3.27	0.79
CPR	3.58	0.50	3.15	0.81
USE	3.56	0.61	3.54	0.66
PFL	3.43	0.61	3.54	0.66
FLOW	3.33	0.52	2.89	0.73

Strongly disagree = 1, Disagree = 2, Agree = 3, Strongly agree = 4

Table 6.5 – Mean Likert scale values and standard deviation for intermediate participants (n=17)

Dimension	Males (9)		Females (8)	
	mean	*sd*	*mean*	*sd*
EXP	3.45	0.71	3.52	0.57
CPR	3.45	0.67	3.38	0.62
USE	3.54	0.64	3.61	0.66
PFL	3.28	0.71	3.38	0.68
FLOW	3.22	0.44	3.13	0.64

Strongly disagree = 1, Disagree = 2, Agree = 3, Strongly agree = 4

female novices disagreed on the FLOW item, while only 2% (1 out of 42) of the remaining participants disagreed. Due to the unbalanced group constellations, statistical tests were only performed for gender differences for participants of all experience levels grouped together (Table 6.7). A two-sample t-test showed that the only significant difference between males and females was indeed found for the FLOW item ($p = 0.015$). This is in particular interesting, since no significant difference ($p > 0.44$) was found for any of the other four dimensions of the questionnaire, including PFL, which includes statements linked to the flow state.

Finally, it was studied whether there were any correlations between the different dimensions. By calculating Pearson's correlation coefficient between the eighteen items of the questionnaire, it was found that the strongest correlations appeared for the items within the dimensions of perceived usefulness for exploration and comprehension. Interestingly, strong correlations were also found across the two sub-dimensions (i.e., EXP3 with CPR2 and EXP4 with CPR4, $r = 0.75$ for both). Moreover, more than half (9 out of 16) of the correlations between the statements of the EXP and CPR dimensions were at least moderate ($r > 0.42$). These unexpected results were opposed to the initial assumption that participants would consider initial exploration and comprehension as two separate aspects of the escape game activity. Indeed, the results illustrated that if participants accepted the escape game as educationally valuable, they believed it was useful for both initial exploration and actual comprehension.

Table 6.6 – Mean Likert scale values and standard deviation for advanced and expert participants (n=19)

Dimension	Males (16)		Females (3)	
	mean	*sd*	*mean*	*sd*
EXP	3.32	0.54	3.42	0.69
CPR	3.14	0.71	3.55	0.52
USE	3.69	0.55	3.89	0.33
PFL	3.50	0.62	3.78	0.43
FLOW	3.50	0.52	3.33	0.58

Strongly disagree = 1, Disagree = 2, Agree = 3, Strongly agree = 4

Table 6.7 – Mean Likert scale values and standard deviation for all participants together (n=61)

Dimension	Males (31)		Females (30)		Two sample t-test	
	mean	*sd*	*mean*	*sd*	*T-statistic*	*p-value*
EXP	3.37	0.64	3.35	0.74	0.198	p=0.843
CPR	3.32	0.68	3.25	0.74	0.763	p=0.446
USE	3.62	0.59	3.60	0.63	0.264	p=0.792
PFL	3.42	0.66	3.46	0.74	-0.457	p=0.648
FLOW	3.39	0.50	3.00	0.70	2.512	p=0.015

Strongly disagree = 1, Disagree = 2, Agree = 3, Strongly agree = 4

Qualitative results

Qualitative data was collected by means of three open-ended questions in the questionnaire and by in-situ observations of a second experimenter present during the game sessions. The open-ended questions were completed by almost half (28 out of 61) of the participants. For the first item Q1, participants mostly indicated that they liked "the atmosphere of the game". This feedback, together with the vivid group dynamics observed, reinforced the results from the quantitative data, indicating that most participants were experiencing flow. Indeed, group members actively shared their discoveries, lively exchanged ideas and verbalized satisfaction: *"Yes we found it!"* (a player after finding a robot in a locked suitcase), *"Ah, now I understand! Look... "* (a player explaining a program code to another), *"Good job! Now the next one!"* (a player after one robot successfully entered the battery). However, it was also observed that the dynamics were not the same for all groups. Especially groups with members of varied level of technical proficiency seemed to be less communicative. In these cases, members with lower proficiency appeared to be more passive and rather focused on non-technical tasks, e.g., searching for hints, while the "experts of the groups" dealt with the robots. For the items Q2 and Q3 of the questionnaire many participants indicated that they desired a more self-contained game, with more hints provided by the game itself rather than having active guidance by the game master. Moreover, some participants felt that the game was too short: *"A little short if the objective is to understand the basics of robotics"*, *"To fully understand the code, more time is needed. But the objective of discovery is very good and makes you want to go further"*. Although a more passive game master was desired during the actual quest, his role as an active facilitator during the debriefing was highly appreciated. For many participants the debriefing was an eye-opener, allowing them to reflect on their actions and better comprehend the functioning of the robots and the program codes. One participant summarized this observation with the following verbalization: *"After the debriefing, I completely changed my mind about the usefulness of this activity"*.

6.3.5 Discussion

By introducing the escape game format to educational robotics, this study sought to explore the experience perceived by the players and the educational potential of such a framework. In the following, we will discuss the results obtained from this exploratory study in view of these two aspects.

Players' perception of the game

From the viewpoint of entertainment, the Thymio Escape Game was a great success, since all participants (61 out of 61) stated to have "really enjoyed the game experience". This was further supported by qualitative data obtained from direct observations, which illustrated high engagement and enjoyment of the participants during the activity. Moreover, a great majority of the participants agreed on the game's ease of use and its usability as a teaching tool.

Unexpectedly, we found strong correlations between EXP and CPR items, refuting our initial assumption that participants would distinguish the game's usability. In contrast, the results showed that if participants found that the game was a valuable teaching tool, they believed it could be useful for both initial exploration and comprehension. However, the results also demonstrated that a positive game and learning experience was not sufficient to ensure a flow experience for all participants. Indeed, 11% (7 out of 61) of the participants disagreed on having "experienced a clear flow during playing". We found a significant difference for the FLOW statement across genders: while all males (100%) agreed or strongly agreed on having experienced flow during the game, this was true for only 73% of the females. This observation is in line with the results of Hou and Chou (2012), who also reported that males experienced more flow than females in their studies. However, when considering the statements of the PFL dimension, we observed that the perceived flow experience was similar for both genders. We presume that this result could be related to the presentation of the FLOW statement, which in our case may not have been sufficient to fully capture the meaning of flow. Moreover, we found that the major contribution to the disparity in the FLOW statement was caused by female novices. Indeed, 6 out of the 7 participants who disagreed on the FLOW statement, belonged to this group. Looking at the gender group sizes for each cluster based on robotics experience, we found that females were predominant in the novices' group. Therefore, another way to interpret these results could be that disparities in the FLOW statement were actually not based on gender differences, but on the proficiency level of the participants. However, further inquiries with larger and more balanced sample sizes would be needed to draw meaningful conclusions.

Educational potentials of escape games

As presented in the previous section most participants perceived the escape game as educationally valuable. Indeed, our observations illustrated that it provides a framework in which people strongly collaborate to achieve a common goal. We observed that participants were self-driven to understand the functioning of the setup and lively exchanged ideas and knowledge among each other. Indeed, the answers given in the questionnaire indicated the desire for a more self-contained game, imposing a less active role on the game master and more autonomy for the players. These results strongly support the idea, that escape games could be a favorable instructional approach in the field of ER to put self-regulated and collaborative learning into practice. In contrast to other constructivist approaches, escape games provide a well-framed environment, which allows the learners to autonomously explore and construct knowledge without having to allocate cognitive resources to activities that are not related to the learning goals, alleviating one main criticism on constructivism found in literature (P. A. Kirschner et al., 2006). The well-defined common goal (i.e., reaching the winning state), that can only be achieved following a pre-determined sequence of tasks, not only provides a more structured framework to implement constructivist methodologies – it also imposes a positive social interdependency among the players, a condition known to be beneficial for collaborative learning scenarios (Johnson & Johnson, 2003). The possibility to give hints, whether indicated by clues in the room or communicated by the game master, additionally provides means to implement and continuously adjust instructional scaffolding. We, therefore, believe that the presented perspectives on escape games may not only apply in the context of ER but can potentially be extended to many other disciplines. However, we suggest that ER may particularly be suited for the integration of escape game activities, due to its inherent constructivist roots. Escape game activities may therefore represent a promising approach for the alignment of classroom activities with the affordances of ER tools. However, whether the escape game format is a effective way to implement ER classroom actvities, still has to be demonstrated. In this context, it is especially important to assess the activity in authentic contexts, i.e., with players of the target age group and in classroom environments. In the next section, we will therefore outline the efforts made to create a classroom version of the Thymio Escape Game.

Table 6.8 – Usability and utility issues identified in the heuristic evaluation of the Thymio Escape Game prototype

ID	Description usability issue	Refers to heuristic
ESC.1	The activity can only be played by max. 6 players at a time.	HERLS-7
ESC.2	The activity requires a significant amount of time for set up and dismantling (around 30 minutes).	HERLS-7
ESC.3	In the current implementation, players are encouraged to follow a blind trial-and-error approach, instead of reflecting on the performed actions.	HERLS-3
ESC.4	When players load wrong program codes on the robots, they may crash into the energy reactor causing material damage and unsettling the students. It also creates a loss of immersion in the game.	HERLS-2
ESC.5	Arranging multiple parallel games of the current version, could promote undesired competition between groups.	HERLS-6
ESC.6	The current role of the game master causes a loss of immersion in the game when hints are given.	HERLS-2
ESC.7	The incomplete preprogrammed behaviors of the robots are confusing and intimidating to the players and hamper an immediate interaction with the robots.	HERLS-2

6.4 Aligning the Thymio Escape Game with classroom activities

Though the prototype of the Thymio Escape Game was mostly well received by the players, some essential changes had to be applied to make it feasible for classroom use. Indeed, the design of the prototype was strongly inspired by commercial escape rooms, involving elements that would hamper the implementation in classrooms. To identify the prototype's points of improvement, a heuristic evaluation with HERLS was performed. Following the heuristic evaluation, a revised classroom version of the Thymio Escape Game was developed. Finally, the revised version was tested in user studies with pre-service teachers and pupils to evaluate its compatibility with classroom settings.

6.4.1 Heuristic evaluation of the Thymio Escape Game prototype

The heuristic evaluation was performed by the principal developer of the activity. After the identification of potential usability issues, the principal developer discussed the results with other project members to compile a list with the most important usability issues (Table 6.8). Based on this list, the following design goals were defined for the next iteration of the Thymio Escape Game activity:

- The activity should allow the simultaneous participation of all students of an entire class (addressing ESC.1)

- The time needed for the preparation of the activity should not exceed 15 minutes, allowing to set up and dismantle the materials of the activity during recess (ESC.2).
- New game rules should prevent blind trial and error approaches (ESC.3).
- The activity should provide students with possibilities to extensively test and improve their solutions before trying them with the energy reactor (ESC.3 and ESC.4).
- The activity should incorporate rules that promotes collaboration between students, instead of competition: the game can only be won, if all reactors are activated at the same time. Groups are allowed to support other groups, if help is demanded (ESC.5).
- The role of the game master should be revised to allow for an immersive experience in which it is still possible to give hints to the players when needed (ESC.6).
- The robots should not be provided with incomplete preprogrammed behaviors (ESC.7).

6.4.2 Developing a classroom version of the Thymio Escape Game

Based on the design goals presented in the previous subsection, a revised version of the Thymio Escape Game was developed, specifically adapted for classroom use. The general game idea was not changed: the goal was still to solve different puzzles to find the Thymio robots, which then had to be programmed to drive into the different compartments of the energy reactor. To allow for the simultaneous participation of all students of a class, multiple identical game setups were prepared. Therefore, the classroom was divided into workspaces of equal sizes and each group of up to six students was assigned to one of the workspaces. Depending on the venue, different arrangements of the room were possible. In any case, it was important that the room arrangement allowed each group to work on their setup without interfering with other groups. Two examples of tested room arrangements are the diamond-shaped (Fig. 6.9) or the parallel (Fig. 6.10) setup of the reactors.

While the diamond-shaped setup is generally preferred, since it allows for a better separation of the groups, it may not always be feasible, especially when the venue is rather small or more than four setups have to be prepared. In these situations, the parallel setup can provide an alternative solution. In either case, both setups embodied a new element introduced for the new version of the activity: the separation of a testing area and the reactor area. While the testing area contained all the materials needed to complete the activity (puzzles, laptops, robots, printed material etc.), the reactor area only contained the energy reactors with their ground maps. The main idea of this spatial separation was to encourage students to first extensively test and debug their program codes before trying their solutions with the reactors. To facilitate this procedure, test maps for each robot (Fig. 6.11) were provided in the testing areas, representing replicas of the corresponding part of the ground map. For the yellow robot, additional cardboard pieces were provided, that could be assembled together to simulate the walls of the reactor. To further prevent students from taking a blind trial and error approach,

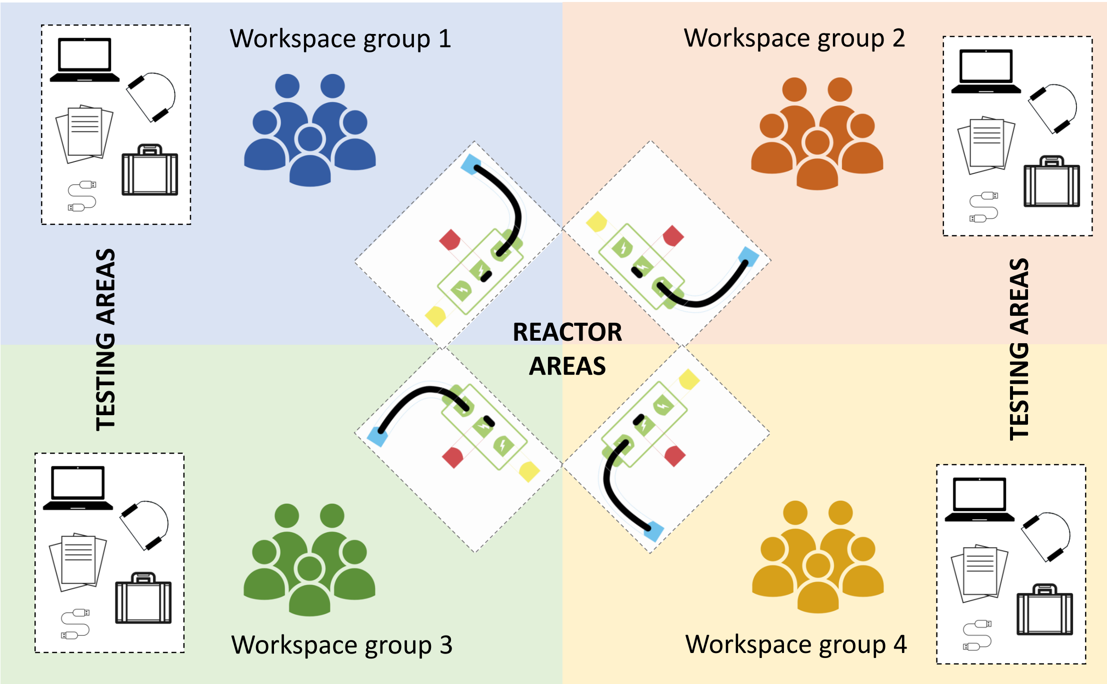

Figure 6.9 – Room arrangement for diamond-shaped setup of energy reactors.

they were also instructed that they would only have one single trial to test their solution with the reactor. If this trial was not successful, the game was over.

In the new version of the game, students were still able to find printed hints in their workspaces helping them with the programming tasks. However, the new hints were now entirely printed on standard DIN A4 sheets and then hidden in a locked suitcase. Moreover, it was decided to completely remove the incomplete preprogrammed behaviors of the robots and instead provide them without any initial behavior. These simplifications allowed to significantly reduce the preparation time right before the activity. The preparation procedure was minimized to the following steps:

1. Rearranging the tables in the room
2. Placing the ground maps and energy reactors
3. Locking and hiding the suitcases containing the unprogrammed robots and the printed materials
4. Hiding the USB cables
5. Placing the laptops and torches
6. Placing the initial puzzles (printed on A4 sheets) to unlock the laptops and the suitcase

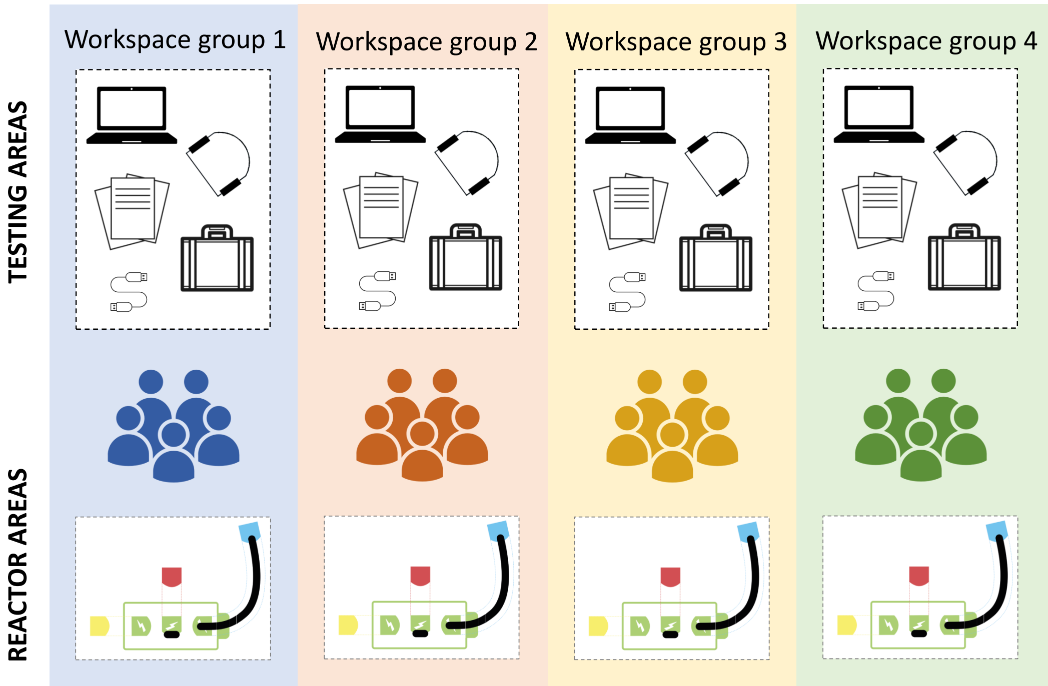

Figure 6.10 – Room arrangement for parallel setup of energy reactors.

Once familiarized with this new procedure, it was possible to set up as well as to dismantle the whole game within 15 minutes. Increasing the efficiency of the setting up and dismantling procedure was an essential improvement to make the Thymio Escape Game feasible for classroom use.

The new arrangement of having multiple games in parallel bore the risk of encouraging an undesired competitive thinking between groups. Indeed, it was expected that different groups would exhibit different completion times for the activity. These performance differences could potentially lead to situations where some groups get bored, while others get discouraged. In order to prevent these situations, some additional constraints were added to the rules of the game. Not only had each group only one trial to try their solution with the energy reactor. It was also imposed that the game could only be finished with success if all energy reactors were activated at the same time. This constraint was expected to enforce a positive social dependency among the different groups, creating a more collaborative atmosphere. In fact, groups that finished before others were allowed to help, if it was requested by the other groups. Finally, another means to equalize the progress of the different groups were the hints given by the game master. Groups that were already ahead received less hints than those who were lagging behind.

Finally, the role of the game master was revised to allow for a better immersion in the activity. The new game master now played the role of a scientist, that recruited the players for a secret research project. Instead of presenting a pre-recorded video, the introduction was now personally given by the game master using a set of presentation slides. To convey an authentic

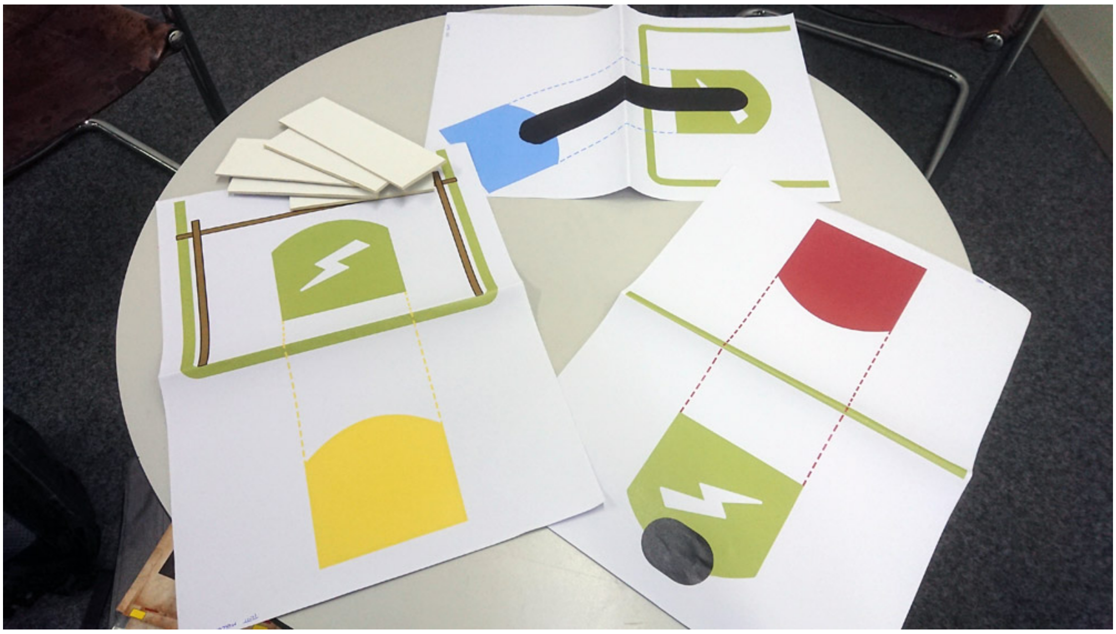

Figure 6.11 – Test maps for the red, blue and yellow robots.

character, the game master was disguised with a lab coat and safety glasses, and acted through role playing. In contrast to the game master in the prototype version, who was located outside of the playing area, the new role as a scientist, allowed the game master to get into direct contact with the players. As a result, the game master was not only able to adjust and target the guidance, but also playing a character that is part of the story, allowing to maintain the players' immersion in the activity.

6.4.3 Study design

Two user studies were organized to test and evaluate the revised version of the Thymio Escape Game. First, a pilot study with 81 pre-service teachers was conducted to study the feasibility of the new game structure. Particularly the parallel organization of multiple setups for large number of participants needed to be tested, before bringing the activity to classroom environments. Following this pilot study, a user study with 76 pupils from Switzerland and the USA was organized to evaluate the new game design in authentic contexts.

Pilot test with pre-service teachers

Participants. The pilot test was conducted with 81 participants enrolled in a Bachelor study program to become teachers in preschools (for children between 3 and 6 years old) and elementary schools (for children between 6 and 11 years old). Participants were predominantly females (n=72), with males being in the minority (n=9). Participants were on average 23.19±3.55 years old. Based on a self-assessment questionnaire, the great majority of partici-

Table 6.9 – Questions for teachers' self-assessment with respect to robotics and Thymio

Statement	Novice	Intermediate	Advanced	Expert
In robotics (and related disciplines, e.g. computer science) I am ...	62	17	2	0
With respect to Thymio and VPL I am...	75	6	0	0

Table 6.10 – Short survey presented to pre-service teachers

Item	Statement *(Strongly disagree / Disagree / Agree / Strongly agree)*
tESC1	I prefer the escape game as an introduction to Thymio and VPL compared to conventional methods (books, lectures, etc.)
tESC2	I think the game alone is enough to understand the basics of Thymio and VPL
tESC3	I think a combination of the escape game and conventional methods is necessary to understand the basics of Thymio and VPL

pants reported no or little prior experience with the Thymio robot and with robotic in general (Table 6.9). The Thymio Escape Game activity was part of a series of workshops about technology and media attended by the pre-service teachers. All participants gave their informed consent to participate in the study.

Study protocol. Before the study, the participants were divided into three groups:

- Group A: 18 pre-service teachers for preschools
- Group B: 30 pre-service teachers for elementary schools
- Group C: 33 pre-service teachers for elemantary schools

The three groups participated in different game sessions, and in each session, the revised version of the Thymio Escape Game was organized with multiple setups at the same time (4 setups for group A, 6 setups for groups B and C). Within each group, participant were asked to autonomously form subgroups and to choose one of the setups for the activity. After a short introduction on the general idea of the activity, the game was launched. For this study, all the modifications described in the previous section were applied. Each session was accompanied by two researchers, one taking the role of the game master and the other providing support in the case of technical issues. The activity lasted for 45 minutes and was concluded with a debriefing and a short survey with three questions on the teachers perception of the educational potential of the activity (Table 6.10). Moreover, participants were given the possibility to provide feedback by means of a free text field.

Table 6.11 – Summary of participant information from the second user study

ID	Gender *(m/f)*	Age *(mean±sd)*	Prior experience with educational robotics *(< 1 time / > 5 times)*	Recruited from
FEST-1	5/6	12.0±0.94	5/6	Public science festival in the USA
FEST-2	7/11	11.50±2.55	11/7	Public science festival in the USA
SCHOOL-1	6/5	11.45±0.52	1/10	Middle school in the USA
SCHOOL-2	6/8	11.46±0.52	5/9	Middle school in the USA
SCHOOL-3	11/10	11.19±0.60	11/10	Secondary school in Switzerland
Total	36/40	11.47±1.37	33/43	

User study with pupils

Participants. The second user study was conducted with 76 pupils from five different groups (Table 6.11). Two groups were recruited from a public science festival in the USA, while the other three groups were participating as a class from schools in Switzerland and the USA. The majority of the participants were from the same age group (around 11-12 years old). Informed consent for the students' participation was obtained by their parents or their legally authorized representative prior to the study.

Study protocol. One goal of the Thymio Escape Game activity is to deliver a positive learning experience, that could have benefical effects on students' interest, attitude and self-efficacy related to robotics. Therefore, a pre-post study was conducted, using a questionnaire with Likert-scale questions related to these dimensions (Table 6.12). The students were asked to complete the questionnaire before and after the Thymio Escape Game activity. Moreover, the questionnaires included questions related to the students demographics and their prior experience with educational robots.

The activity was conducted in the same way for all five groups: they were first welcomed by the researchers in front of the room prepared with the Thymio Escape Game activity. Outside of the room, they were then asked to complete the pre-questionnaires. Subsequently, subgroups were formed - in the sessions organized at the schools, the subgroups were arranged with the help of the supervising teachers, while for the sessions at the science festival the decision was up to the participants. In all five sessions, four reactor setups were prepared in the diamond-shaped configuration (see Fig. 6.9), resulting in subgroups sizes of 2-6 students, depending on the session. Once assigned to a subgroup, the students were allowed to enter the room, in which one researcher, disguised as the game master, introduced them to the activity. The game was then conducted as the revised version according to the descriptions presented in the previous section. During the activity another researcher was present to provide support in case of technical issues. The time limit for all sessions was set to 30 minutes, which were then followed by the debriefing session. The sessions concluded with the post-tests distributed after the debriefing.

Table 6.12 – Questions in the pre- and post-tests presented to pupils

Item	Statement *(Strongly disagree / Disagree / Agree / Strongly agree)*
pINT1	Robotics is interesting
pINT2	Learning more about robotics is interesting
pINT3	It is interesting to solve robotics questions in this learning activity
pINT4	A robotics course is more interesting to me in comparison with other courses
pATT1	A robotics course is valuable and worth studying
pATT2	It is worth learning about robotics
pATT3	It is important to learn more about robotics, including testing the robot
pATT4	It is important to learn about programming the robot
pSE1	I am sure that I can learn about robotics
pSE2	I don't do well in robotics
pSE3	Even before I begin a new topic in robotics, I feel confident I will be able to understand it
pSE4	I think I have good skills and strategies to learn robotics

Statistical analysis. Data were pooled together for all participating students from the five sessions (n=76). The students' answers were analyzed for pre- and post-tests separately and according to three dimensions: interest, attitude and self-efficacy towards robotics. Each dimension was quantified by four Likert-scale items in the pre- and post-tests. To analyze the reliability of the test items, Cronbach's alpha values were calculated for each dimension. Descriptive statistics are reported as mean ± one standard deviation, for all students' data pooled together as well as for two types of clusters: one cluster was separated by gender (males and females) and the other by prior experience with robotics (with and without). One-sample t-tests were performed for within-group comparisons and two-sample t-tests were performed for between-group comparisons. The level of statistical significance was set to $\alpha = 0.05$. The statistical tests were corrected for multiple comparisons using a Bonferroni correction for, resulting in a corrected significance level of $\alpha_{corr} = 0.025$ (for two comparisons).

6.4.4 Results

Pilot study with pre-service teachers

The main objectives of this pilot study was to assess the feasibility of the revised game design and to collect some first feedback regarding the educational potential of the revised activity. In all three sessions organized for the pre-service teachers, no major issues occurred, illustrating that the new game design even allowed for the simultaneous participation of up to 33 players at the same time. All groups successfully completed the activity and general feedback was predominantly positive. The short survey conducted at the end of the activity, allowed to explore the perceptions of the participating pre-service teachers. Two thirds of the participants agreed or strongly agreed that they would prefer to have the escape game as an introduction to Thymio and the VPL interface (tESC1, Fig. 6.12). The participants appreciated the activity as a playful and motivating way to discover the Thymio robot and its programming interface VPL, as for instance described by one of the participants in the free comments section of the survey:

> *"The escape game was very well organized and motivating, the prepared materials were very helpful for solving the exercises."*

Another participant praised the escape game approach, however, at the same time advocating for more time to complete the activity:

> *"It would take more time for everyone to experiment, but I think escape game is the best way to do it."*

In another comment, a participant suggested that a formal introduction should precede the escape game:

> *"The escape game is fun and allows you to understand how Thymio works but it's stressful. It would be better to make an introduction first."*

Indeed, the teachers seemed less convinced about the potential of the activity to completely replace conventional introductions. More than two thirds disagreed or strongly disagreed on the statement that the game alone would be enough to understand the basics of Thymio and VPL (tESC2). This viewpoint was summarized by a comment of a participant who argued as follows:

> *"I believe that having conventional activities before the escape game promotes the understanding of the basics of Thymio during the game."*

As a matter of fact, most of the comments given by the participants were related to the integration of a formal introduction before the escape game, to take full advantage of the activity. This

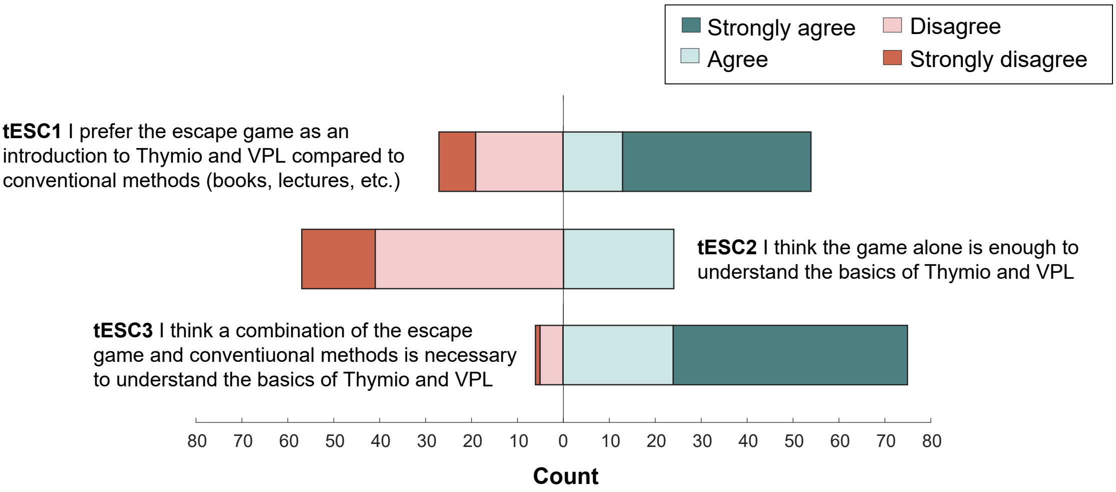

Figure 6.12 – Results from survey with pre-service teachers.

was also reflected by the answers given to the third question of the survey, for which a large majority (93%) indicated that a combination of the escape game and conventional methods would be the most effective way to learn about the basics of Thymio and VPL (tESC3).

User study with pupils

The main objective of this user study was to assess the potential of the activity in an authentic context, i.e., with participants of the target age group and under conditions similar to the ones in classrooms. In all five sessions, four parallel game setups were prepared. Similar to the pilot study conducted with the pre-service teachers, no major issues were observed and the game was successfully completed by all groups. Overall, the activity was received very well, and the participating students seemed to have enjoyed the Thymio Escape Game experience. As a response to the question whether students would like to redo this kind of activity, 63 out of the 76 students answered with yes, as opposed to 13 students who responded no.

The results obtained from the pre- and post questionnaires were analyzed according to the three dimensions explored by the test items: students' interest, attitude and self-efficacy towards robotics. Each dimension was measured by four different items in the questionnaires (see Table 6.12). To analyze the reliability of the test items, Cronbach's alpha was calculated for each dimension in both pre- and post-tests (Table 6.13). The Cronbach's alpha values were above 0.8 for all dimensions, which according to existing literature can be considered a reasonable value (Gliem & Gliem, 2003).

As quantified by the pre- and post-tests, the average Likert-scale scores of students' interest, attitude and self-efficacy toward robotics increased after the activity (Table 6.14). On average, values increased by 0.29 for interest, by 0.10 for attitude and by 0.34 for self-efficacy. However, a one-sample t-test confirmed statistically significant differences between pre- and post-tests

Table 6.13 – Cronbach's alpha values for each dimension in pre- and post-tests

Dimension	Pre-tests	Post-tests
Interest (pINT1-4)	0.858	0.863
Attitude (pATT1-4)	0.867	0.879
Self-Efficacy (pSE1-4)	0.842	0.821

Table 6.14 – Descriptive statistics from pre- and post-tests for all students (n=76)

Dimension	Pre-tests	Post-tests		One-sample t-test	
	mean±sd	*mean±sd*	*delta*	*T-statistic*	*p-value*
Interest	3.07±0.66	3.36±0.62	+0.29	-5.893	1.00e-07
Attitude	3.28±0.60	3.38±0.67	+0.10	-1.596	0.11
Self-Efficacy	2.88±0.69	3.22±0.67	+0.34	-5.500	5.04e-07

Strongly disagree = 1, Disagree = 2, Agree = 3, Strongly agree = 4

only for the dimensions interest and self-efficacy. All post-test scores were above 3.0, indicating a positive self-assessment of the students with respect to the items of the questionnaire. While scores for interest and attitude were already above 3.0 in the pre-tests, the activity helped to also push self-efficacy scores towards the positive side. Indeed, the greatest gain between pre- and post-tests was found for the self-efficacy dimension (+0.34).

Furthermore, the scores were analyzed for a separation by gender (Table 6.15). Both males and females showed statistically significant increases in scores for interest and self-efficacy. While females did not further improve (+0.0) their attitude scores (which were already on a considerably high level with a mean of 3.41), males improved them a bit (+0.17). However, the changes were not statistically significant. Two-sample t-tests between males and females showed no statistically significant differences between the groups, in any of the dimensions and neither in pre- nor in post-tests (Table 6.16). However, between pre- and post-tests, differences between males and females decreased for interest (delta between groups from 0.15 to 0.02) and attitude (from 0.26 to 0.09). Differences for self-efficacy, which were small already in the pre-tests, increased slightly (from 0.03 to 0.04).

The scores were also analyzed for a separation by the students' prior experience with robotics (Table 6.17). Based on the participants' answers given in the questionnaire, participants were clustered into students without (less than one experience with robotics) and with prior experience (more than 5 experiences with robotics). Also here, statistically significant improvements were found for both groups only for interest and self-efficacy. For the attitude dimension, a tendency of improvement was found for students with prior experience. Between-group comparisons showed that mean scores for all dimensions were statistically significantly different for both groups in the pre-tests (Table 6.18). Between pre- and post-tests, the differences

Table 6.15 – Descriptive statistics from pre- and post-tests by gender

Dimension	Pre-tests	Post-tests		One-sample t-test	
Males (n=36)	*mean±sd*	*mean±sd*	*delta*	*T-statistic*	*p-value*
Interest	2.99±0.76	3.38±0.66	+0.39	-4.916	2.07e-05
Attitude	3.15±0.67	3.32±0.71	+0.17	-1.835	0.075
Self-Efficacy	2.90±0.74	3.24±0.70	+0.34	-3.718	6.98e-04
Females (n=40)	*mean±sd*	*mean±sd*	*delta*	*T-statistic*	*p-value*
Interest	3.14±0.56	3.36±0.59	+0.22	-3.449	0.001
Attitude	3.41±0.50	3.41±0.64	+0.0	-0.198	0.844
Self-Efficacy	2.87±0.66	3.20±0.66	+0.33	-4.005	2.70e-04

Strongly disagree = 1, Disagree = 2, Agree = 3, Strongly agree = 4

Table 6.16 – Two-sample t-tests between males (n=36) and females (n=40)

Dimension	Males	Females		Two-sample t-test	
Pre-tests	*mean±sd*	*mean±sd*	*delta*	*T-statistic*	*p-value*
Interest	2.99±0.76	3.14±0.56	+0.15	1.001	0.320
Attitude	3.15±0.67	3.41±0.50	+0.26	1.930	0.057
Self-Efficacy	2.90±0.74	2.87±0.66	+0.03	-0.212	0.833
Post-tests	*mean±sd*	*mean±sd*	*delta*	*T-statistic*	*p-value*
Interest	3.38±0.66	3.36±0.59	+0.02	-0.087	0.931
Attitude	3.32±0.71	3.41±0.64	+0.09	0.596	0.553
Self-Efficacy	3.24±0.70	3.20±0.66	+0.04	-0.231	0.818

Strongly disagree = 1, Disagree = 2, Agree = 3, Strongly agree = 4

Table 6.17 – Descriptive statistics from pre- and post-tests by prior experience

Dimension	Pre-tests	Post-tests		One-sample t-test	
W/o Exp (n=33)	*mean±sd*	*mean±sd*	*delta*	*T-statistic*	*p-value*
Interest	2.80±0.72	3.07±0.72	+0.27	-2.771	0.009
Attitude	3.10±0.66	3.13±0.81	+0.03	-0.196	0.846
Self-Efficacy	2.49±0.61	2.86±0.71	+0.37	-4.027	3.39e-04
With Exp (n=43)	*mean±sd*	*mean±sd*	*delta*	*T-statistic*	*p-value*
Interest	3.27±0.55	3.60±0.41	+0.33	-5.931	4.99e-07
Attitude	3.42±0.52	3.57±0.49	+0.15	-2.902	0.058
Self-Efficacy	3.20±0.59	3.49±0.51	+0.29	-3.596	8.44e-04

Strongly disagree = 1, Disagree = 2, Agree = 3, Strongly agree = 4

increased for interest (delta between groups from 0.47 to 0.53) and attitude (from 0.32 to 0.43), and they decreased for self-efficacy (from 0.71 to 0.63). Also in the post-tests, differences between the groups were statistically significant for all dimensions.

In a next step, students were then clustered by their changes between pre- and post-tests . Students' scores were classified either as increasing (changes in scores greater than 0.1), unchanged (changes in scores between -0.1 and 0.1) or decreasing (changes in scores smaller than -0.1). The analysis showed that for interest and self-efficacy, almost two thirds of the students (47 out of 76) increased their scores in the post-tests (Table 6.19). Furthermore, over one third of the students (27 out of 76) increased their scores in the attitude dimension. The distribution of students who improved their scores, appeared to be balanced by gender for all dimensions. However, with regard to the interest and attitude dimensions, students with prior experience appeared to be predominant. For the self-efficacy dimension instead, both students with and without prior experience increased their scores.

For most students (33 out of 76) the activity did not change their scores in the attitude dimension. Similarly, for around one fifth (14 out of 76) and a bit more than one quarter of the students (22 out of 76), the activity did not change their scores for self-efficacy and interest, respectively. The profiles for these groups were similar to the group of increasing students, with most students having prior experience with robotics. For interest and attitude gender distribution appeared to be balanced, for self-efficacy a higher proportion of males was observed.

Finally, there were also some students who showed a decrease in their post-test scores following the activity. For attitude and self-efficacy this was the case for around one fifth of the students (15 and 16 out of 76 students, respectively). For one tenth of the students (7 out of 76 students) a decrease in scores for the interest dimension was observed. Students who decreased their scores were more likely females, and they were more likely to not have

Table 6.18 – Two-sample t-tests between students without (n=33) and with prior robotics experience (n=43)

Dimension	Without Exp	With Exp		One-sample t-test	
Pre-tests	*mean±sd*	*mean±sd*	*delta*	*T-statistic*	*p-value*
Interest	2.80±0.72	3.27±0.55	+0.47	-3.213	0.002
Attitude	3.10±0.66	3.42±0.52	+0.32	-2.373	0.020
Self-Efficacy	2.49±0.61	3.20±0.59	+0.71	-5.039	3.28e-06
Post-tests	*mean±sd*	*mean±sd*	*delta*	*T-statistic*	*p-value*
Interest	3.07±0.72	3.60±0.41	+0.53	-3.986	1.57e-04
Attitude	3.13±0.81	3.57±0.49	+0.43	-2.970	0.004
Self-Efficacy	2.86±0.71	3.49±0.51	+0.63	-4.508	2.439e-05

Strongly disagree = 1, Disagree = 2, Agree = 3, Strongly agree = 4

prior experience with robotics. The only exception was found for self-efficacy, for which the proportion of both students with and without experience was balanced. However, it should be considered that more students participating in this study had prior experiences (n=43) as compared to students without prior experiences (n=33).

Prior experience with robotics, to some extent, also appeared to influence whether students enjoyed the activity or not (Table 6.20). With regard to the question whether students would like to redo this kind of activity, 63 students answered with yes, as opposed to 13 students who responded with no. Students who responded positively mostly had prior experience with robotics, whereas students negating the question were predominantly novices. On the other hand, gender did not appear to influence whether students wanted to redo these kinds of activities or not.

6.4.5 Discussion and future work

In this study we have presented the design and evaluation of the revised version of the Thymio Escape Game. With respect to the prototype version presented in the previous section, particular emphasis was put on aligning the game design with classroom activities as described in the ERLS framework. After a heuristic evaluation performed using the HERLS heuristics, several changes were applied to the game design and structure to better align it with the educational goals and to make it compatible with classroom use. Indeed, the introduction of the spatial separation of testing and reactor areas as well as the additional constraint of having only one trial with the reactor, appeared to encourage students to take more structured approaches. Similar to the constraint imposed on the students in the Thymio lawnmower activity, they helped students to make use of other cognitive artifacts available, such as the test maps. Using the test maps, students did not have to fear breaking the reactor and hence showed more

Table 6.19 – Profiles of increasing, unchanged or decreasing student scores

Dimension	Gender (m/f)	Experience (with / without)	Mean change
Increasing			
Interest (n=47)	22/25	29/18	+0.55±0.34
Attitude (n=27)	15/12	18/9	+0.56±0.46
Self-Efficacy (n=47)	22/25	25/22	+0.66±0.35
Unchanged			
Interest (n=22)	12/10	13/9	0.00±0.02
Attitude (n=33)	15/18	21/12	0.00±0.00
Self-Efficacy (n=14)	9/5	10/4	0.00±0.00
Decreasing			
Interest (n=7)	2/5	1/6	-0.46±0.17
Attitude (n=16)	6/10	4/12	-0.52±0.29
Self-Efficacy (n=15)	5/10	8/7	-0.38±0.27

Table 6.20 – Profiles of students wanting to play the activity again or not

Play again?	Gender (m/f)	Experience (with/without)
Yes (n=63)	29/34	41/22
No (n=13)	7/6	2/11

rigorous testing behaviors. They also allowed students to better collaborate in subgroups of 2-3 students, since each subgroup could work with one robot and one test map. The new role assumed by the game master helped to provide scaffolding, without interrupting the game flow, and hence further leveraged socio-constructivist learning approaches.

Two user studies were organized to demonstrate the feasibility of the revised activity in authentic contexts as well as to have an initial assessment of its educational potential. In a pilot study with 81 pre-service teachers, the simultaneous arrangement of multiple setups allowed the game to be organized for groups as large as 33 people. Demonstrating the feasibility of the game for such big groups was important to ensure that the activity can also be organized in classroom settings. The three sessions with the pre-service teachers were completed without any major issues and hence illustrated the feasibility to implement the revised activity with big groups. In a short survey presented at the end of the activity the participants indicated general appreciation for the activity. However, many participants also highlighted that the Thymio Escape Game activity would be more useful if it was used in combination with conventional approaches as an introduction, that would provide the players with basics about the Thymio robot and the programming interface VPL.

The importance of prior experience for the effective use of the Thymio Escape Game activity was also a main finding of the second user study performed with 76 pupils. In this study, students' interest, attitude and self-efficacy towards robotics were analyzed by pre- and post-tests right before and after the activity. Overall, we found that the activity helped students to increase their interest and self-efficacy, whereas increases in attitude were not statistically significant. Indeed, the intervention conducted in this study was rather short (around 90 minutes in total for game and debriefing) and changes in attitude may require more time to show significant changes.

When analyzing the data clustered by gender and prior experience, increases for interest and self-efficacy were observed for both males and females, as well as for students with and without prior experience. Between-group comparisons showed that there were no statistically significant differences between males and females in all dimensions of the pre-tests. The same was observed for the post-tests, with differences in mean scores becoming even smaller following the activity. This finding is particularly interesting in light of previous work, where self-evaluations of students showed statistically significant differences after a sequence of more conventional educational robotics activities (Negrini & Giang, 2019). Moreover, between-group comparisons were performed for students with prior experience in robotics and those without. Not surprisingly, we found statistically significant differences between groups in all dimensions of the pre-tests. Students with prior experience showed high scores for interest, attitude as well as self-efficacy. Though both groups increased their scores following the activity, the post-test scores of students with prior experience were still statistically significantly higher compared to the ones of students without prior experience.

These results suggest that the Thymio Escape Game had, at least on the short term, a positive effect on students' interest and self-efficacy towards robotics, independent of gender or prior experience. By adopting the escape game format, it is possible to implement ER classroom activities, leveraging the appeals of escape room experiences, that are equally appreciated by boys and girls, as well as novices and experienced students. However, the analysis also showed that implementing an ER escape game activity did not automatically guarantee for a positive learning experience for all the students. As a matter of fact, up to a fifth of the students showed decreased scores for the three dimensions in the post-tests. A deeper analysis showed that these participants more likely appeared to be females and in the case of interest and attitude dimensions, were more likely to be students that had no prior experience with robotics.

Prior experience also seemed to be a factor influencing how much the students enjoyed the activity. From the 13 students indicating that they would not like to redo these kind of activities, 11 had no prior experience with robotics. On the other side, however, there were also 22 students with no prior experience that seemed to have enjoyed the activity, indicating that they would like to redo these kind of activities. From the data available it is not possible to draw further conclusions on why certain students did not seem to enjoy the activity as much as others or why some students showed a decrease of scores in the post-tests. A possible factor, as in any collaborative activity, could have been the group constellations, that were not controlled for in this study. Previous works have investigated collaboration in commercial escape games (Pan et al., 2017) and likewise, the effect of the group composition in collaborative learning has been discussed before (Slavin, 1996). Bringing together these evidences could thus further improve the learning experiences induced by educational escape games. Indeed, it might be that students with decreasing scores found themselves in unfavorable group constellations, that were negatively influencing their learning experience.

To draw more substantial conclusions, however, further studies are needed. Though the current study presents promising results with regard to the potential of adapting the escape game format to ER classroom activities, it also comes with limitations. First of all, it has to be acknowledged that quantifying interest, attitude and self-efficacy using Likert-scale questionnaires may not fully reflect the complexity of such constructs. In this regard qualitative research methods, such as in-situ observations or interviews, may be considered, allowing to provide a more exhaustive assessment of these constructs. Moreover, measures were performed within a relatively short time frame (right before and right after the activity), not allowing to capture whether the observed positive effects would also be retained in long term. In the current study, the Thymio Escape Game activity was organized as a one time event and it was not embedded in a broader context. However, especially in formal education it is important to understand where and how this kind of activities can be integrated. As found in the study with the pre-service teachers, it seems like a conventional introduction could help to more effectively take advantage of the escape game format. Without a formal introduction, the activity may appear overwhelming and stressful to some participants, leading to less enjoyment and appreciation of the activity. This may have also been the case for the students who decreased their scores in the post-tests or those students who indicated that they would not like to redo this kind of activity. Future studies should investigate these aspect in order to provide more evidence for the educational relevance of these approaches.

6.5 Conclusion

In this chapter we have discussed two examples of ER activities and illustrated the actions taken to account for instructional alignment in each case. Using the Thymio lawnmower mission, we have shown how a simple instructional intervention, i.e., a temporary access restriction to the programming interface, has created a more favorable learning experience with respect to the students' development of computational thinking skills. In contrast, we have observed that unrestricted access to the programming interface, may lead to a misaligned learning system, in which students are less likely to achieve the intended outcomes. In the second example, we have presented the development of the Thymio Escape Game. The activity was devised in several iterations following a procedure that, to a certain extent, can be considered similar to the alignment-centered design process proposed by Lauwers (2010). While we did not have the possibility to involve educators in participatory design approaches, the ERLS framework and the HERLS heuristics showed to be useful tools in guiding the development of the activity. The final result was an immersive ER classroom activity that has been appreciated by students and teachers likewise. A user study showed that the activity can, at least in the short term, raise students' interest and self-efficacy related to robotics. Overall, we therefore, suggest that the Thymio Escape Game can be considered a successful implementation of an ER classroom activity. Until today, it has been played by more than 600 players in different countries around the world (Switzerland, Italy, Germany, USA, France and Belgium). The game material as well as documentations in four languages (English, German, French and Italian) have been made publicly available to disseminate the concept also outside of the academic realm. Finally, a good indicator for the suitability of the Thymio Escape Game as a classroom activity are last but not least the teachers that have already independently adopted the game concept for their purposes.

7 General conclusion

7.1 Summary

In this final chapter, we will present a synthesis of the results presented in the previous chapters and highlight the main contributions of this thesis. Subsequently, we will discuss the limitations of the studies conducted and indicate possible directions for future work. The work will conclude with some final words of the author.

7.2 Main contributions of this thesis

This thesis sought to study the question of instructional alignment for educational robotics activities. In three main stages, we explored this subject from different perspectives, with the aim of providing a comprehensive view on this matter. In the following, we will summarize the main contributions of each stage and relate them to the research questions that were formulated at the end of Chapter 2 of this thesis.

7.2.1 Contributions to alignment theory

At the first stage, we sought to address the research question *"How can instructional alignment in the context of ER be conceptualized?"*. The discussion was initiated by a definition of the type of ER activities that have been considered in this thesis. We have specified them as those activities, which are implemented in formal education settings and in which robots are used as cognitive artifacts to teach both technical and non-technical topics. Based on this definition, we identified the main components of such activities and specified their distinct roles as cognitive artifacts in situated learning systems. Furthermore, we discussed the main learning theories underlying such activities. Based on these discussions, we then introduced the notion of Educational Robotics Learning Systems (ERLS). It represents an extension of the alignment model previously proposed by Lauwers (2010), that has been specifically adapted to the context of educational robotics. It may serve as an alignment framework for developers designing ER tools as well as for educators designing and implementing ER classroom activities. Finally, it may also be useful for educational researchers who are interested in studying the question of instructional alignment for ER.

With the aim of further facilitating the application of the ideas presented in the ERLS framework by developers and educators, we have then proposed a backwards design approach that was supported by HERLS, a set of heuristics for the design and evaluation of ERLS. The HERLS heuristics were developed in multiple iterations involving teachers, developers and educational researchers. An exploratory user study with twelve primary and lower secondary school teachers illustrated that a heuristic evaluation using HERLS was more appropriate in reflecting the teachers' needs than evaluations based only on technical characteristics of ER tools. Moreover, focus groups with developers illustrated their appreciation of HERLS as a guiding tool for the design of ER tools, helping them to uncover issues in the designs of their systems that they would not have found based only on their intuition.

In short, the main contribution of this thesis with respect to the alignment theory of ER can be summarized as follows:

- We have devised the ERLS alignment framework and its complement, the HERLS heuristics, aimed at facilitating the instructional alignment of educational robotics activities. Studies with teachers, developers and educational researchers have illustrated their potential as a useful tool to conceptualize and guide instructional alignment in the context of ER.

7.2.2 Contributions to the design of ER tools

Using the examples of two development projects, we then sought to address the second research question *"How can developers align the designs of their ER tools?"*. In the PaPL project, the goal was to develop a tangible programming interface for the educational robot Thymio. A main emphasis was laid on the use of accessible materials, such as paper and cardboard to facilitate the use of the interface in classroom environments. The design of the interface generally followed the alignment-centered design process presented by Lauwers (2010). However, the whole process was now additionally supported by the ERLS framework and the HERLS heuristics. Since the PaPL project did not have the possibilities to involve educators in participatory design of the interface, the initial design and requirements evaluation had to be performed by the developers on their own. We observed that it was especially during these initial steps, that the proposed approach seemed to be particularly useful for the developers. The ERLS framework and the HERLS heuristics allowed them to autonomously perform evaluations of their initial ideas and prototypes and helped them to identify weak spots of the system that could then be addressed in the next development iteration. The resulting Thymio PaPL programming interface was finally tested in a classroom study run by two teachers. The results of the study illustrated the potential of the interface to facilitate the introduction of programming activities with the Thymio robot. In comparison with an introduction using the screen-based alternative, students achieved higher scores in their Parson's Puzzle quizzes. Moreover, the study provided some initial evidence about the potential of Thymio PaPL to support teachers in classroom orchestration. While it is undisputed that the proposed approach using the ERLS framework and the HERLS heuristics cannot fully replace participatory design approaches, we believe it can be considered a valuable complement to the alignment-centered design process proposed by Lauwers. It may be especially valuable for cases in which participatory methods are not available or feasible.

In the second example, the CreroBot project, we extended the approach taken in the PaPL project to develop an accessible do-it-yourself educational robot. Due to the Covid-19 pandemic, however, it was not yet possible to test the CreroBot in classroom studies. Nevertheless, the circumstances happened to open up a possibility to apply the ideas of the ERLS framework in a new context, namely that of distance learning. In this regard, a whole teaching unit was designed for in-service teachers taking part in a continuous professional development course on ER. The unit was designed as a remote lesson, in which the teachers worked with CreroBot and its programming interfaces to discover the concept of making. Finally, personal interviews were scheduled with the teachers to evaluate their learning experience. The results of the interviews illustrated that the teachers appreciated the proposed distance learning activities with CreroBot and considered them to be educationally meaningful. As a matter of fact, many teachers have already expressed their interest in participating in classroom studies with CreroBot, once the epidemiological situation would allow it.

In short, our contributions with respect to the development of ER tools can be summarized as follows:

- We have illustrated how the proposed ERLS alignment framework and its complement, the HERLS heuristics, can support existing alignment-centered design processes. Especially in development settings, where participatory methods are not available or feasible, it may help developers in aligning their ER tools with educational needs.

- We have devised a framework that allows developers to create tangible programming languages based on computer-vision methods and accessible materials such as paper and cardboard. So far, the framework has been used to implement two interfaces: Thymio PaPL and CreroBot PaPL. In the near future, the code of the PaPL project as well as a documentation will be made publicly available, allowing other developers to adopt the approach for their purposes.

- We have devised a low-cost design for a do-it-yourself robot based on accessible materials. The use of cardboard as the base for the robot's structure allows for a high degree of accessibility and personalization, leveraging the ideas of maker education. Though the system has still to be assessed in classroom studies, an exploratory study with in-service teachers has highlighted its potential as a valuable learning tool.

- Finally, we have also provided an example of how educational robotics education can be organized in distance learning. In this regard, we also illustrated how the ERLS framework can be leveraged for the design of ER tools and instructional activities aimed at the specific case of distance learning.

7.2.3 Contributions to the design of ER activities

In the last stage of this thesis we finally presented two examples of ER activities with the aim of addressing the research question *"How can educators align their instructional activities in the context of ER?"*. As a first example, we presented the Thymio lawnmower mission, a learning activity with the educational robot Thymio aimed at fostering the development of students' computational thinking competencies. With two groups of primary school students, we illustrated the effect of a misalignment in the instructional component of the ERLS framework and suggested a simple intervention that can be applied to remedy the situation. While one group had unlimited access to the graphical programming interface, the other was subject to temporary restrictions on the use of the interface. Video analyses of the students' behaviors showed that the latter displayed a much broader variety of cognitive processes related to computational thinking competencies, whereas the former appeared to be plunged in a blind trial-and-error approach. By introducing a temporary access restriction to the programming interface, we hence encouraged students to follow more rigorous approaches in planning and structuring their ideas and strategies, resulting in a more constant engagement of the students in the activity as well as more successful solutions for their lawnmower robots.

As a second example, we presented the development and evaluation of the Thymio Escape Game activity. Here, the main idea was to leverage the escape game format to create an immersive learning experience with the Thymio robot, aiming at increasing students' interest, attitude and self-efficacy towards robotics. Also in this project, we followed the alignment-centered design process proposed by Lauwers (2010). However, this time it was applied to design an ER classroom activity using existing ER tools. Similar to the PaPL project, we did not have the possibilities to involve teachers in participatory design, especially in the initial phases of the project. Once again, we therefore, used the ERLS framework and the HERLS heuristic to guide the development. The process resulted in a classroom-adapted version of the activity that was finally tested in a user study with students from Switzerland and the USA. As measured by pre- and post-questionnaires, we observed that the activity helped to increase students' interest and self-efficacy, independently of gender or prior experience with robotics. Although the great majority of players appreciated the experience, the surveys also showed that there were some students that did not enjoy the activity and even decreased their scores following the activity. While this certainly deserves further research to be understood, we can nevertheless conclude that overall, the Thymio Escape Game can be considered an example of a successfully implemented ER activity. Until today, the activity has been played by more than 600 people and the material as well as documentations in four languages (English, German, French and Italian) have been made available to the public. As a matter of fact, educators from France, Belgium, Italy and different Cantons of Switzerland have already independently adopted the approach to implement ER classroom activities.

In short, our contributions with respect to the development of ER activities can be summarized as follows:

- We have illustrated how a misalignment of the instructional component of the ERLS framework can negatively impact the achievement of intended outcomes. In the specific case of ER activities for the development of computational thinking skills, we have presented a simple instructional intervention (i.e., a temporary access restriction to the programming interface) to remedy the situation.
- We have illustrated how an existing alignment-centered design process, originally devised for the design of ER tools, can also be applied for the design of ER activities. Moreover, we have illustrated how the ERLS framework and its complement, the HERLS heuristics can support this process, especially in cases for which participatory design is not feasible.
- We have devised an immersive ER classroom activity based on the concept of escape games, that, to a certain extent, has had an impact on educational practice. Until today, the activity has been played by over 600 people in different countries and it has been adopted by several teachers to independently implement ER classroom activities. The game material as well as documentations in four languages (English, German, French and Italian) have been made publicly available.

7.3 Limitations and future work

In this section, we would like to present a discussion on the limitations of the studies presented in this thesis. The discussions should allow the reader to appropriately interpret the results of this thesis and they will moreover, guide us towards possible directions for future works.

First of all, it has to be emphasized that the ERLS framework as well as the HERLS heuristics, are not claimed to be exhaustive. As outlined in Chapter 3, this thesis was concerned with a specific subset of possible ER activities, i.e., those activities in which robots are used as cognitive artifacts to teach both technical and non-technical topics in formal education. While this represents a very specific subset of activities, we still believe that our findings have relevance, since this kind of activities are nonetheless quite common in classroom settings. However, whether the proposed approach also applies to other types of ER activities, remains open and could be investigated in future research.

To assess the validity of the ERLS framework and the HERLS heuristics, focus groups with external teachers, developers and educational researchers were organized. Though the number of participants was comparatively high for this kind of research (five focus groups with 45 participants), we acknowledge that generalizability is naturally limited by the number of focus groups that can be organized. To provide further evidence about the usefulness of the devised approach, we therefore proposed the methodology to an external group of developers currently involved in the design of a new ER tool (CFG2 in Chapter 4). The findings from the study with this group further supported our proposed approach. However, the methodology would have to be applied by more external design groups and researchers to further consolidate its relevance. We acknowledge that in this context, the ERLS framework as well as the HERLS heuristic, may be subject to further revisions.

Though the proposed approach proved to be useful for both the design of ER tools as well as ER activities, it should be emphasized, that its main value lies in the role of complementing existing alignment-centered design processes, rather than replacing them. The ERLS framework and the HERLS heuristics may provide some useful guidance, especially in the early design stages and particularly when participatory design approaches are not feasible. However, for projects aiming at having a strong educational impact, user studies and co-design approaches are still essential, since they provide insights that cannot be covered by the presented methodology. Nevertheless, we believe that complementing existing design approaches with the ERLS framework and the HERLS heuristic could provide some additional guidance, potentially making the design process more targeted and purposeful. This is another direction of research that could be further explored in future studies.

Finally, it should be acknowledged that the studies performed to assess the devised ER tools and ER activities were implemented as one time interventions. While this does not necessarily attenuate the relevance of all findings (especially not those based on interviews or in situ observations), it certainly limits the power of the results obtained from pre-post tests. To provide stronger evidence on the impact of the instructional alignment design based on the proposed methodology, long-term interventions and evaluations have to be considered. This would also allow to include more high-level learning objectives as the intended outcomes component of the ERLS framework.

7.4 Final words

This thesis has sought to make a contribution to the field of educational robotics by addressing the following global research question: *How can instructional alignment be attained for educational robotics activities?* In three main stages, we studied that topic from different perspectives. Starting from a theoretical point of view, we first devised a conceptual framework for the alignment of ER activities and in this context, we introduced the notion of Educational Robotics Learning Systems (ERLS). Based on this framework we proposed a backwards design approach that was complemented by HERLS, a set of heuristics for the design and evaluation of ERLS. In the second stage, we presented two projects that were concerned with the development of ER tools and illustrated how the devised methodologies can be useful to align the design of such tools. Finally, in the last stage of this thesis, we presented two examples of ER activities to illustrate how alignment can be attained by appropriately adapting the instructional activities. By going from theory through development up to the implementation of classroom activities, we intended to provide diverse perspectives on this topic and we hope that the findings from this dissertation will prove useful for future research in this field.

Bibliography

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

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

Antonenko, P. D., Dawson, K., & Sahay, S. (2017). A framework for aligning needs, abilities and affordances to inform design and practice of educational technologies. *British Journal of Educational Technology*, *48*(4), 916–927.

Anwar, S., Bascou, N. A., Menekse, M., & Kardgar, A. (2019). A systematic review of studies on educational robotics. *Journal of Pre-College Engineering Education Research (J-PEER)*, *9*(2), 2.

Arai, T., Machii, K., & Kuzunuki, S. (1995). Retrieving electronic documents with real-world objects on interactivedesk, In *Proceedings of the 8th annual acm symposium on user interface and software technology*. ACM.

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.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: an introduction to sequential analysis*. Cambridge university press.

Barbosa, M. B., Rego, A. B., & de Medeiros, I. (2015). Heeg: heuristic evaluation for educational games, In *Xiv brazilian symposium of games and digital entertainment. teresina-pi, brazil*.

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: a review. *Science robotics*, *3*(21), eaat5954.

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

Bers, M. U., Flannery, L., Kazakoff, E. R., & Sullivan, A. (2014). Computational thinking and tinkering: exploration of an early childhood robotics curriculum. *Computers & Education*, *72*, 145–157.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher education*, *32*(3), 347–364.

Biggs, J. (2003). Aligning teaching and assessing to course objectives. *Teaching and learning in higher education: New trends and innovations*, *2*(April), 13–17.

Blikstein, P. (2019). Project bloks: embodied and collaborative learning with tangible interfaces for young children. *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings*.

Blikstein, P., Sipitakiat, A., Goldstein, J., Wilbert, J., Johnson, M., Vranakis, S., Pedersen, Z., & Carey, W. (2016). Project bloks: designing a development platform for tangible programming for children. *Position paper, retrieved online on*, 06–30.

Bloom, B. S. Et al. (1956). Taxonomy of educational objectives. vol. 1: cognitive domain. *New York: McKay*, *20*, 24.

Bocconi, S., Chioccariello, A., Dettori, G., Ferrari, A., & Engelhardt, K. (2016). Developing computational thinking in compulsory education.

Bocconi, S., Chioccariello, A., & Earp, J. (2018). The nordic approach to introducing computational thinking and programming in compulsory education. *Report prepared for the Nordic@ BETT2018 Steering Group. doi: https://doi. org/10.17471/54007*.

Boller, S., & Kapp, K. (2017). *Play to learn: everything you need to know about designing effective learning games*. Association for Talent Development.

Bower, M. (2008). Affordance analysis–matching learning tasks with learning technologies. *Educational Media International*, *45*(1), 3–15.

Bower, M., & Sturman, D. (2015). What are the educational affordances of wearable technologies? *Computers & Education*, *88*, 343–353.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational researcher*, *18*(1), 32–42.

Bruner, J. (1978). The role of dialogue in language acquisition. *The child's conception of language*, 241–256.

Cascalho, J., Mendes, A., Ramos, A., Pedro, F., Bonito, N., Almeida, D., Augusto, P., Leite, P., Funk, M., & Garcia, A. (2019). Azoresbot: an arduino based robot for robocup competitions, In *Iberian robotics conference*. Springer.

Catlin, D. (2017). 29 effective ways you can use robots in the classroom (D. Alimisis, M. Moro, & E. Menegatti, Eds.). In D. Alimisis, M. Moro, & E. Menegatti (Eds.), *Educational robotics in the makers era*, Cham, Springer International Publishing.

Catlin, D., & Blamires, M. (2010). The principles of educational robotic applications (era): a framework for understanding and developing educational robots and their activities. The 12th EuroLogo conference.

Catlin, D., Kandlhofer, M., Cabibihan, J.-J., Angel-Fernandez, J., Holmquist, S., & Csizmadia, A. P. (2019). Edurobot taxonomy, In *Smart learning with educational robotics*. Springer.

Chen, G., Shen, J., Barth-Cohen, L., Jiang, S., Huang, X., & Eltoukhy, M. (2017). Assessing elementary students' computational thinking in everyday reasoning and robotics programming. *Computers & Education*, *109*, 162–175.

Chevalier, M., Giang, C., Piatti, A., & Mondada, F. (2020). Fostering computational thinking through educational robotics: a model for creative computational problem solving. *International Journal of STEM Education*, *7*(39). https://doi.org/10.1186/s40594-020-00238-z

Clarke, S., Peel, D. J., Arnab, S., Morini, L., Keegan, H., & Wood, O. (2017). Escaped: a framework for creating educational escape rooms and interactive games for higher/further education. *International Journal of Serious Games*, *4*(3), 73–86.

Clarkson, E., & Arkin, R. C. (2007). Applying heuristic evaluation to human-robot interaction systems., In *Flairs conference*.

Cohen, S. A. (1987). Instructional alignment: searching for a magic bullet. *Educational Researcher*, *16*(8), 16–20.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319–340.

Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: the self-determination perspective. *Educational psychologist*, *26*(3-4), 325–346.

Dennehy, R. F., Sims, R. R., & Collins, H. E. (1998). Debriefing experiential learning exercises: a theoretical and practical guide for success. *Journal of Management Education*, *22*(1), 9–25.

Desurvire, H., Caplan, M., & Toth, J. A. (2004). Using heuristics to evaluate the playability of games, In *Chi'04 extended abstracts on human factors in computing systems*.

Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education*, *69*, 485–492.

Eguchi, A. (2014). Educational robotics for promoting 21st century skills. *Journal of Automation Mobile Robotics and Intelligent Systems*, *8*(1), 5–11.

Eguchi, A. (2017). Bringing robotics in classrooms, In *Robotics in stem education*. Springer.

Elia, J. S. I. (1988). An alignment experiment in vocabulary instruction: varying instructional practice and test item formats to measure transfer with low ses fourth graders.

Elkin, M., Sullivan, A., & Bers, M. U. (2016). Programming with the kibo robotics kit in preschool classrooms. *Computers in the Schools*, *33*(3), 169–186.

Fahey, P. A. (1988). Learning transfer in main ideas instruction: effects of instructional alignment and aptitude on main idea test scores.

Fraser, N. (2015). Ten things we've learned from blockly, In *2015 ieee blocks and beyond workshop (blocks and beyond)*. IEEE.

Gall, M., & Breeze, N. (2005). Music composition lessons: the multimodal affordances of technology. *Educational Review*, *57*(4), 415–433.

Giang, C. (2019). An open development environment for paper-based programming languages. *Research proposal for small projects approved by the Hasler Stiftung*.

Giang, C., Chevalier, M., Negrini, L., Peleg, R., Bonnet, E., Piatti, A., & Mondada, F. (2018). Exploring escape games as a teaching tool in educational robotics, In *International conference edurobotics 2016*. Springer. https://doi.org/10.1007/978-3-030-18141-3_8

Giang, C., Piatti, A., & Mondada, F. (2019). Heuristics for the development and evaluation of educational robotics systems. *IEEE Transactions on Education*, *62*(4), 278–287. https://doi.org/10.1109/TE.2019.2912351

Giang, C., Piatti, A., & Mondada, F. (2020). Aligning the design of educational robotics tools with classroom needs. *Manuscript submitted for publication*.

Gibson, J. J. (1979). The ecological approach to visual perception.

Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting cronbach's alpha reliability coefficient for likert-type scales. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education.

Guinchard, A. (2020). *Development of a mobile application for computer-vision based tangible programming languages* (semester project report). École Polytechnique Fédérale de Lausanne.

Heersmink, R. (2013). A taxonomy of cognitive artifacts: function, information, and categories. *Review of philosophy and psychology*, *4*(3), 465–481.

Heft, H. (1989). Affordances and the body: an intentional analysis of gibson's ecological approach to visual perception. *Journal for the theory of social behaviour*, *19*(1), 1–30.

Horn, M. S., Solovey, E. T., & Jacob, R. J. (2008). Tangible programming and informal science learning: making tuis work for museums, In *Proceedings of the 7th international conference on interaction design and children*. ACM.

Hou, H.-T., & Chou, Y.-S. (2012). Exploring the technology acceptance and flow state of a chamber escape game-escape the lab© for learning electromagnet concept. *ICCE 2012*, *38*.

Hu, F., Zekelman, A., Horn, M., & Judd, F. (2015). Strawbies: explorations in tangible programming, In *Proceedings of the 14th international conference on interaction design and children*. ACM.

Inkpen, K. M., Ho-Ching, W.-l., Kuederle, O., Scott, S. D., & Shoemaker, G. B. (1999). This is fun! we're all best friends and we're all playing: supporting children's synchronous collaboration, In *Proceedings of the 1999 conference on computer support for collaborative learning*. International Society of the Learning Sciences.

Jerzak, N., & Rebelo, F. (2014). Serious games and heuristic evaluation–the cross-comparison of existing heuristic evaluation methods for games, In *International conference of design, user experience, and usability*. Springer.

Johnson, D. W., & Johnson, R. T. (2003). *Assessing students in groups: promoting group responsibility and individual accountability*. Corwin Press.

Jung, S. E., & Won, E.-S. (2018). Systematic review of research trends in robotics education for young children. *Sustainability*, *10*(4), 905.

Karim, M. E., Lemaignan, S., & Mondada, F. (2015). A review: can robots reshape k-12 stem education?, In *2015 ieee international workshop on advanced robotics and its social impacts (arso)*. IEEE.

Khanlari, A. (2013). Effects of robotics on 21st century skills. *European Scientific Journal*, *9*(27).

Kiili, K. (2006). Evaluations of an experiential gaming model. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*.

King, A. (1993). From sage on the stage to guide on the side. *College teaching*, *41*(1), 30–35.

Kinio, A., Dufresne, L., Brandys, T., & Jetty, P. (2017). Break out of the classroom: the use of escape rooms as an alternative learning strategy for surgical education. *Journal of Vascular Surgery*, *66*(3), e76.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational psychologist*, *41*(2), 75–86.

Kirschner, P., Strijbos, J.-W., Kreijns, K., & Beers, P. J. (2004). Designing electronic collaborative learning environments. *Educational technology research and development*, *52*(3), 47.

Koczor, M. L. (1986). Effects of varying degrees of instructional alignment in posttreatment tests on mastery-learning tasks of fourth-grade children.

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (tpack)? *Contemporary issues in technology and teacher education*, *9*(1), 60–70.

Köffel, C., & Haller, M. (2008). Heuristics for the evaluation of tabletop games, In *Evaluating user experiences in games, workshop at the 2008 conference on human factors in computing systems (florence, italy, 5–10 april 2008). chi*.

Kölling, M., & McKay, F. (2016). Heuristic evaluation for novice programming systems. *ACM Transactions on Computing Education (TOCE)*, *16*(3), 1–30.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159–174.

Lauwers, T. (2010). Aligning Capabilities of Interactive Educational Tools to Learner Goals. https://doi.org/10.1184/R1/6714650.v1

Lee, I., Martin, F., Denner, J., Coulter, B., Allan, W., Erickson, J., Malyn-Smith, J., & Werner, L. (2011). Computational thinking for youth in practice. *Acm Inroads*, *2*(1), 32–37.

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

Malone, T. (1981). What makes computer games fun?, In *Proceedings of the joint conference on easier and more productive use of computer systems.(part-ii): human interface and the user interface-volume 1981*.

Mannila, L., Dagiene, V., Demo, B., Grgurina, N., Mirolo, C., Rolandsson, L., & Settle, A. (2014). Computational thinking in k-9 education, In *Proceedings of the working group reports of the 2014 on innovation & technology in computer science education conference*.

Mehrotra, A., Giang, C., Duruz, N., Dedelley, J., Mussati, A., Skweres, M., & Mondada, F. (2020). Introducing a paper-based programming language for computing education in classrooms, In *Proceedings of the 2020 acm conference on innovation and technology in computer science education*. https://doi.org/10.1145/3341525.3387402

Melcer, E. F., & Isbister, K. (2018). Bots & (main) frames: exploring the impact of tangible blocks and collaborative play in an educational programming game, In *Proceedings of the 2018 chi conference on human factors in computing systems*. ACM.

Mishra, P., & Koehler, M. (2009). Too cool for school? no way! using the tpack framework: you can have your hot tools and teach with them, too. *Learning & Leading with Technology*, *36*(7), 14–18.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: a framework for teacher knowledge. *Teachers college record*, *108*(6), 1017–1054.

Mitchell, F. M. (1999). All students can learn: effects of curriculum alignment on the mathematics achievement of third-grade students.

Mohamed, H., & Jaafar, A. (2010). Development and potential analysis of heuristic evaluation for educational computer game (pheg), In *5th international conference on computer sciences and convergence information technology*. IEEE.

Mondada, F., Bonani, M., Riedo, F., Briod, M., Pereyre, L., Rétornaz, P., & Magnenat, S. (2017). Bringing robotics to formal education: the thymio open-source hardware robot. *IEEE Robotics & Automation Magazine*, *24*(1), 77–85.

Morgan, D. L. (1996). *Focus groups as qualitative research* (Vol. 16). Sage publications.

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.

Mussati, A., Giang, C., Piatti, A., & Mondada, F. (2019). A tangible programming language for the educational robot thymio, In *2019 10th international conference on information, intelligence, systems and applications (iisa)*. IEEE. https://doi.org/10.1109/IISA.2019.8900743

Negrini, L., & Giang, C. (2019). How do pupils perceive educational robotics as a tool to improve their 21st century skills? *Journal of e-Learning and Knowledge Society*, *15*(2).

Nelson, L., Ichimura, S., Pedersen, E. R., & Adams, L. (1999). Palette: a paper interface for giving presentations, In *Proceedings of the sigchi conference on human factors in computing systems*. ACM.

Nicholson, S. (2012). Completing the experience: debriefing in experiential educational games, In *Proceedings of the 3rd international conference on society and information technologies*.

Nicholson, S. (2015). Peeking behind the locked door: a survey of escape room facilities. *White paper*.

Nielsen, J. (1995a). How to conduct a heuristic evaluation. *Nielsen Norman Group*, *1*, 1–8.

Nielsen, J. (1995b). Severity ratings for usability problems. *Papers and Essays*, *54*, 1–2.

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems, In *Proceedings of the interact'93 and chi'93 conference on human factors in computing systems*.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces, In *Proceedings of the sigchi conference on human factors in computing systems*.

Norman, D. A. (1988). *The psychology of everyday things*. Basic books.

Osborne, R. (2014). An ecological approach to educational technology: affordance as a design tool for aligning pedagogy and technology.

Pachidis, T., Vrochidou, E., Kaburlasos, V., Kostova, S., Bonković, M., & Papić, V. (2018). Social robotics in education: state-of-the-art and directions, In *International conference on robotics in alpe-adria danube region*. Springer.

Pan, R., Lo, H., & Neustaedter, C. (2017). Collaboration, awareness, and communication in real-life escape rooms, In *Proceedings of the 2017 conference on designing interactive systems*.

Pane, J. F., & Myers, B. A. (1996). *Usability issues in the design of novice programming systems* (tech. rep.). CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER SCIENCE.

Papert, S. (1980). Mindstorms: computers, children, and powerful ideas. *NY: Basic Books*.

Papert, S., & Harel, I. (1991). Situating constructionism. *Constructionism*, *36*(2), 1–11.

Parsons, D., & Haden, P. (2006). Parson's programming puzzles: a fun and effective learning tool for first programming courses, In *Proceedings of the 8th australasian conference on computing education-volume 52*. Australian Computer Society, Inc.

Piaget, J., & Cook, M. T. (1954). The construction of reality in the child.

Prensky, M. (2003). Digital game-based learning. *Computers in Entertainment (CIE)*, *1*(1), 21–21.

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J. S., Silverman, B., Et al. (2009). Scratch: programming for all. *Commun. Acm*, *52*(11), 60–67.

Rétornaz, P., Riedo, F., Magnenat, S., Vaussard, F., Bonani, M., & Mondada, F. (2013). Seamless multi-robot programming for the people: aseba and the wireless thymio ii robot, In *2013 ieee international conference on information and automation (icia)*. IEEE.

Riedo, F., Chevalier, M., Magnenat, S., & Mondada, F. (2013). Thymio ii, a robot that grows wiser with children, In *2013 ieee workshop on advanced robotics and its social impacts*. IEEE.

Sabuncuoğlu, A., & Sezgin, M. (2020). Kart-on: affordable early programming education with shared smartphones and easy-to-find materials, In *Proceedings of the 25th international conference on intelligent user interfaces companion*, Cagliari, Italy, Association for Computing Machinery. https://doi.org/10.1145/3379336.3381472

Sapounidis, T., & Demetriadis, S. (2013). Tangible versus graphical user interfaces for robot programming: exploring cross-age children's preferences. *Personal and ubiquitous computing*, *17*(8), 1775–1786.

Sapounidis, T., Demetriadis, S., & Stamelos, I. (2015). Evaluating children performance with graphical and tangible robot programming tools. *Personal and Ubiquitous Computing*, *19*(1), 225–237.

Scaradozzi, D., Screpanti, L., & Cesaretti, L. (2019). Towards a definition of educational robotics: a classification of tools, experiences and assessments, In *Smart learning with educational robotics*. Springer.

Seau Yoon, F., Ho, J., & Hedberg, J. G. (2005). Teacher understandings of technology affordances and their impact on the design of engaging learning experiences. *Educational Media International*, *42*(4), 297–316.

Sharpe, T. L., & Koperwas, J. (2003). *Behavior and sequential analyses: principles and practice*. Sage.

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2014). Student engagement in high school classrooms from the perspective of flow theory, In *Applications of flow in human development and education*. Springer.

Shin, J., Siegwart, R., & Magnenat, S. (2014). Visual programming language for thymio ii robot, In *Conference on interaction design and children (idc'14)*. ETH Zürich.

Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educational researcher*, *15*(2), 4–14.

Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, *22*, 142–158.

Slavin, R. E. (1996). Research on cooperative learning and achievement: what we know, what we need to know. *Contemporary educational psychology*, *21*(1), 43–69.

Steinkuehler, C., & Squire, K. (2014). Videogames and learning. *Cambridge handbook of the learning sciences*, 377–396.

Strawhacker, A., Sullivan, A., & Bers, M. U. (2013). Tui, gui, hui: is a bimodal interface truly worth the sum of its parts?, In *Proceedings of the 12th international conference on interaction design and children*. ACM.

Sullivan, A. A., Bers, M. U., & Mihm, C. (2017). Imagining, playing, and coding with kibo: using robotics to foster computational thinking in young children. *Siu-cheung KONG The Education University of Hong Kong, Hong Kong*, *110*.

Sweetser, P., & Wyeth, P. (2005). Gameflow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, *3*(3), 3–3.

Tada, K., & Tanaka, J. (2015). Tangible programming environment using paper cards as command objects. *Procedia Manufacturing*, *3*, 5482–5489.

Tallarico, I. (1985). Effects of ecological factors on elementary school student performance on norm-referenced standardized tests: nonreading behaviors (testwiseness).

Terry, M., Cheung, J., Lee, J., Park, T., & Williams, N. (2007). Jump: a system for interactive, tangible queries of paper, In *Proceedings of graphics interface 2007*. ACM.

Tobias, S., Fletcher, J. D., & Wind, A. P. (2014). Game-based learning, In *Handbook of research on educational communications and technology*. Springer.

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.

Tremblay, M. C., Hevner, A. R., & Berndt, D. J. (2010). The use of focus groups in design science research, In *Design research in information systems*. Springer.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The academic motivation scale: a measure of intrinsic, extrinsic, and amotivation in education. *Educational and psychological measurement*, *52*(4), 1003–1017.

Vörös, A. I. V., & Sárközi, Z. (2017). Physics escape room as an educational tool, In *Aip conference proceedings*. AIP Publishing LLC.

Vygotsky, L. (1964). Thought and language.

Vygotsky, L. (1980). *Mind in society: the development of higher psychological processes*. Harvard university press.

Whitton, N. (2014). *Digital games and learning: research and theory*. Routledge.

Wiggins, G., Wiggins, G. P., & McTighe, J. (2005). *Understanding by design*. Ascd.

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

Wonder-McDowell, C., Reutzel, D. R., & Smith, J. A. (2011). Does instructional alignment matter? effects on struggling second graders' reading achievement. *The Elementary School Journal*, *112*(2), 259–279.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of child psychology and psychiatry*, *17*(2), 89–100.

Zufferey, G., Jermann, P., Lucchi, A., & Dillenbourg, P. (2009). Tinkersheets: using paper forms to control and visualize tangible simulations, In *Proceedings of the 3rd international conference on tangible and embedded interaction*. ACM.

Christian Giang

Contact information:

Chemin des Triaudes 4 A506
1024 Ecublens VD, Switzerland
+41 78 916 99 97
christian.giang@epfl.ch

Education

09/2018 – 10/2021	**ETH Zürich** • Teaching certificate for Electrical Engineering
10/2017 – 09/2020	**EPFL Lausanne** • Ph.D. student in Educational Robotics in collaboration with SUPSI - Scuola Universitaria Professionale della Svizzera Italiana. Advisors: Prof. Francesco Mondada and Dr. Alberto Piatti.
03/2015 – 08/2015 02/2014 – 07/2014	**Technion – Israel Institute of Technology** • Two exchange semesters during Bachelor's and Master 's degree
09/2011 – 09/2017	**ETH Zürich** • Bachelor of Science in Information Technology and Electrical Engineering • Master of Science in Information Technology and Electrical Engineering with specialization „Systems and Control" Tutor: Prof. John Lygeros

Teaching experience

10/2017 – present	**SUPSI – Dipartimento Formazione e Apprendimento** • Lecturer for the Certificate of Advanced Studies (CAS) in Educational Robotics in spring and fall semesters • Lecturer for the German version of the MOOC on the educational robot Thymio
10/2017 – 09/2020	**EPFL – Mobile Robotics Laboratory** • Head teaching assistant for the course "Application of Bayes filters to mobile robot localization" in spring semesters • Teaching assistant for the course "Basics of mobile robotics" in fall semesters
09/2015 – 12/2015	**ETH Zürich – Automatic Control Laboratory** • Teaching assistant for the "Control Systems I" course • Teaching assistant for the "Control Lab" course
09/2012 – 12/2012	**ETH Zürich – Integrated Systems Laboratory** • Teaching assistant for the "Electronic Circuits Lab" course

Work experience

04/2017 – 07/2017	**EPFL - Translational Neural Engineering Laboratory** • Full-time research assistant
10/2015 – 06/2016	**EW Höfe, Freienbach** • Student assistant in the power grid team
09/2014 – 02/2015	**ABB Switzerland Ltd., Zürich-Oerlikon** • Six-month industry internship • Supporting the R&D team in the development of HVDC GIS
08/2013 – 07/2014	**Asylorganisation Zürich** • Voluntary activities as a mentor in the AOZ Future Kids project
11/2012 – 08/2014	**ETH Zürich - Product Development Group** • Student assistant in IT administration
08/2010 – 08/2011	**Jovenes Agroecologistas de la Zona Norte, Costa Rica** • Voluntary social year in ecotourism

Technical skills

- Expertise in Educational Robotics, Learning Technologies, Rehabilitation Robotics and Control Theory
- Experience with data analysis, quantitative and qualitative research methods
- Programming languages and software: C++, Java, Python, MATLAB, OpenCV, Adobe Creative Cloud, Microsoft Office

Research grants

- Main applicant for a research grant from the Hasler Stiftung for the project entitled "*An open development environment for paper-based programming languages*" (26'000 CHF total, approved in February 2020)

Peer reviewing

- 16th IEEE/RAS-EMBS International Conference on Rehabilitation Robotics (ICORR 2019), Toronto, Canada
- 4th Annual ACM/IEEE International Conference on Human Robot Interaction (HRI 2019), Daegu, South Korea
- BioMedical Engineering OnLine
- IEEE Transactions on Education

Languages

- German: C2
- Vietnamese: C2
- English: C1
- Spanish: C1
- French: C1
- Italian: B2
- Portuguese: B1
- Hebrew: A2

List of publications

- Chevalier, M.*, Giang, C.*, Piatti, A., & Mondada, F., (2020). Fostering computational thinking through educational robotics: a model for creative computational problem solving. *International Journal of STEM Education*, in press. (*) equally contributing authors
- Mehrotra, A.*, Giang, C.*, Duruz, N., Dedelley, J., Mussati, A., Skweres, M., & Mondada, F. (2020). Introducing a Paper-Based Programming Language for Computing Education in Classrooms. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education* (pp. 180-186). (*) equally contributing authors
- Pierella, C., Pirondini, E., Kinany, N., Coscia, M., Giang, C., Miehlbradt, J., ... & Chisari, C. (2020). A multimodal approach to capture post-stroke temporal dynamics of recovery. *Journal of Neural Engineering*, 17(4), 045002.
- Giang, C., Pirondini, E., Kinany, N., Pierella, C., Panarese, A., Coscia, M., ... & Micera, S. (2020). Motor improvement estimation and task adaptation for personalized robot-aided therapy: a feasibility study. *BioMedical Engineering OnLine*, 19(1), 1-25.
- Mussati, A., Giang, C., Piatti, A., & Mondada, F. (2019). A Tangible Programming Language for the Educational Robot Thymio. In *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)* (pp. 1-4). IEEE.
- Negrini, L., & Giang, C. (2019). How do pupils perceive educational robotics as a tool to improve their 21st century skills?. *Journal of e-Learning and Knowledge Society*, 15(2).
- Giang, C., Piatti, A., & Mondada, F. (2019). Heuristics for the development and evaluation of educational robotics systems. *IEEE Transactions on Education*, 62(4), 278-287.
- Giang, C., Pirondini, E., Kinany, N., Panarese, A., Pierella, C., & Micera, S. (2019). European Patent Application No. WO2019154911.
- Pierella, C., Giang, C., Pirondini, E., Kinany, N., Coscia, M., Miehlbradt, J., ... & Micera, S. (2018). Personalizing Exoskeleton-Based Upper Limb Rehabilitation Using a Statistical Model: A Pilot Study. In *International Conference on NeuroRehabilitation* (pp. 117-121). Springer, Cham.
- Giang, C., Chevalier, M., Negrini, L., Peleg, R., Bonnet, E., Piatti, A., & Mondada, F. (2018). Exploring escape games as a teaching tool in educational robotics. In *International Conference EduRobotics 2016* (pp. 95-106). Springer, Cham.
- Giang, C., & Mirkin, L. (2016). Identifying systems from closed loops: Duality to robust control and an unsung opportunity. In *2016 European Control Conference (ECC)* (pp. 501-506). IEEE.