Thèse n° 10 097

Human-Robot Swarm Interaction: An Explorative Path to Foster Complex Systems Understanding

Présentée le 28 juillet 2023

Faculté informatique et communications
Laboratoire d'ergonomie éducative
Programme doctoral en robotique, contrôle et systèmes intelligents

pour l'obtention du grade de Docteur ès Sciences

par

Hala KHODR

Acceptée sur proposition du jury

Prof. A. Ijspeert, président du jury
Prof. P. Dillenbourg, Prof. B. Bruno, directeurs de thèse
Prof. S. Hauert, rapporteuse
Prof. A. Repenning, rapporteur
Prof. F. Mondada, rapporteur

École
polytechnique
fédérale
de Lausanne

2023

To my dear family for their ubiquitous love . . .

Acknowledgements

This work would not have been possible without the ultimate support system I was blessed with from different perspectives and all life axes.
During the past years, I definitely spent much more time in the lab than anywhere else. So I would like to start by saying a thank you to the CHILI lab globally for being my home away from home.
Thank you, Pierre, my "Doktorvater". I feel incredibly fortunate to have had you as my supervisor and mentor. Thank you for creating an excellent working environment and setting a remarkable example of excellence, humility, respect, and consideration. Pierre is full of brilliant insights and captivating stories. His caring nature extends beyond academics, as he genuinely prioritizes his students' well-being. His constructive guidance kept me focused on finding joy in the process, and his constant encouragement kept me motivated through challenging times.
Thank you, Barbara, also known as BB, Brbrbr, or any combination of B and rs, my "doktormutter". Your guidance and support ensured that I stayed focused on the topics that were "truly related to my research". I am incredibly grateful for the countless crazy, fun, engaging, and thought-provoking conversations we had. We covered everything from teenage tales to food discussions, from fascinating research talks to life's broader topics. This created an abundance of inside jokes that could sometimes get us in trouble, but they made the journey all the more enjoyable. Your humor and care are unwavering. Thank you from every single "perspective".
Thank you, Aditi, for enthusiastically coming on board with my thesis. Aditi, you have been my go-to person for any small questions or complaints about learning sciences. Your vast knowledge and expertise in the field have been invaluable to me throughout this journey. I am especially grateful for your motherly ways of encouraging me to write and stay on track, coming to my rescue with clementines at the end of a long working day and a much-needed boost when I needed it the most.
Thank you, Wafa, whose influence was instrumental in my decision to join the lab for my PhD. From working together during my semester project to her recommendation that led me to CHILI, she has been a guiding force. Wafa's support during my first year, including navigating daunting user experiments, and her continued presence, even amidst her own busy schedule, have been invaluable.
Thank you, Florence, known as the real boss, for managing all the bureaucracy very smoothly.
To my colleagues, it has been a great pleasure to share this journey with each one of you.
Thank you:

Ayberk, for providing the best project documentation and prompt and enthusiastic responses to all my questions.
Arzu, for her happy spirit, the fun collaborations, and the late funny nights in the lab.
Thibault, my neighbor, whose voice is quite a view and for being by my side, literally.
Jauwairia, for giving me the most beautiful view from my desk, which I truly missed in my last few months. Thank you for helping me grow, for being my to-go person for advice, for giving me the honor to be your roommate in skiminaires and conferences, for maintaining our friendship even when I act as a very annoying little sister, both in and outside the lab.
Utku, the nicest and most helpful person on earth.
Sina, the great orchestrator and the bro who knows how to make my day with the best watermelon and ice cream.
Louis, for introducing me to new board and card games.
Kevin, the wise man and my basketball rival.
Teresa, my phat thai partner and kind friend.
Ramtin, for sharing the joy of telling lame jokes.
Sven, for being a cheerful friend and for your valuable feedback on my UI design.
Lucas, the best teammate ever, who patiently listens to and appreciates my translated jokes.
Daniel, my fifth brother who I can count on in all sweet and bitter moments. Thank you for "sharing" your snacks with me, transferring all the joyful positive energy, all the dad jokes reels, the expensive Decathlon shopping trips, and the very "careful" camping adventures.
Victor, a great collaborator, for sharing the frustration of electronics shortages and bearing all the Star Wars sound effects in the streets of Kyoto during IROS.
Alan, the barbecue king, for being a brilliant student to supervise and work with.
Richard, the fastest typer I ever met, for your valuable support, feedback, and positive energy. Thank you for the fun stories and discussions over lunch.
Chenyang, for your valuable friendship, even though a wall of screens separates us.
Zhenyu, the great basketball coach whose concern is about keeping people in the lab healthy.
Raphael, for your fun spirit. It was nice crossing paths in the lab the last few months.
Jie, for your deep philosophical questions on research and life.
Anthony, for your kindness.
Laurent, for your humor and "nothing special".
And many thanks to those who were around during the first years of my PhD: Jenny, Stian, Catharine, Corinne, Thomas, Dorsa and Jade.
Thank you, Soheil, for a fruitful collaboration despite the challenges of managing meetings that spanned across Switzerland, Canada, and Australia.
Thank you, Kevin H., for quickly and enthusiastically accepting the initiation of an amazing year-long collaboration that not only led to academic enrichment but also fostered a great friendship with wonderful memories, lame jokes, and meetings turning into F1 discussions.
I would also like to thank all the brilliant students who did a semester project or an internship with me. It was a great joy to work with you - Harold, Alex, Mehdi, Ulysse, Gianni, Luca, Jerome, Yassine, Jose, Lisa, Hadrien, Antoine, Victoria, Paul, Daniel, and Julien. A special thanks to the courageous ones, Nicolas and Jean-Etienne, who stuck around for longer.

I would like to express my appreciation to NCCR Robotics for funding my PhD and providing a superb environment for interaction among brilliant researchers in the field of robotics from all over Switzerland. Thank you for fostering a great environment for personal and professional growth, as well as facilitating exposure and interaction.
I extend my gratitude to my esteemed jury committee: Auke Ijspeert, Francesco Mondada, Sabine Hauert, and Alexander Repenning for taking the time to evaluate my work and provide their valuable feedback.
All the experiments and results of this thesis would not have been possible without the participation of the students. Thank you to all the wonderful students who took part in my studies. I would also like to extend my thanks to the directors and teachers for their diligent efforts in facilitating the organization of the studies.
I also completed my Masters at EPFL, during which I had the opportunity to work with and meet amazing people. I would like to express my gratitude to Auke for giving me the opportunity to conduct my master's thesis in Biorob. The atmosphere in Biorob was exceptional, and I would like to extend special thanks to Mehmet and Simon, with whom I worked closely on my Master's project and from whom I gained valuable knowledge. Additionally, I had the chance to work in DISAL, for which I am thankful to Alcherio and Bahar for providing me with this enriching opportunity.
Thank you, Ece, for being the very first amazing friend I met in literally the first lecture at EPFL. We quickly formed a strong bond and shared both the joys and struggles throughout our Masters and PhD journeys.
Thank you Alaa and Ahmed, for your friendship and welcoming me into your team for course projects making the integration of a newbie less challenging.
Thank you Jawwad, my true Pakistani friend for your kindness, despite my mean jokes.
Big thanks to GMU for creating such a tight-knit community, with weekly Friday gatherings, hikes, and the great Ramadan iftars. Through GMU, I had the opportunity to meet some amazing people and form incredible groups of friends.
Thank you Azza, Souad, Nour, Imane, Sara, Hana, Mayssa, Alaa, and Yusr for adopting me into your crew. Your friendship holds a special place in my heart. Thank you for all the gatherings, outings, and iftars we shared. A special thanks for considering speaking Tunisian a little slower so that I could follow along.
Thank you Jauwairia, Zahraa, Kaouthar, Firdaous, Mayssa, Caroline, Alaa, Zeinab, and Ghewa for your valuable bond, for all our crazy gatherings, travel trips, and sleep overs. You've been my haven in overwhelming times.
Thank you Mayssa, for being the consistent friend among many group friends, as we experienced our groups dispersing across the globe. Your positivety is contagious, and I am grateful to have shared all the awesome memories with you since my first year in Switzerland.
Thank you Zeinab, my cutie friend, for showing your care by sharing your delicious homemade food when it's needed the most, as we share the ups and downs of this PhD journey.
Thank you to all my Lebanese friends I met in Lausanne and specifically to the Lebanese gang—Taha, Mira, Kassir, Wajeb, Mahdi, Ahmad, Zeinab, and Rania—for being my second home. Mira and I arrived in Switzerland around the same time. Our friendship, which started

at AUB, has only grown stronger here. She has quickly become the person I can rely on anytime, my go-to emergency contact. She also got me a new little friend, Youssof, whom I'm unapologetically transferring my sense of humor to.
Going to the mountains has become my gateway for mood booster, and despite the fact that Kassir stole my initials and can be quite annoying on purpose at times, he has become my great hiking buddy, and I am grateful for his friendship.
My journey in Switzerland wouldn't have been the same without Francine and Martin. They quickly agreed to rent me the studio in their home when I was almost ready to give up on finding housing. They were incredibly welcoming, genuinely caring, always very helpful and made sure everything is going well for me. They introduced me to the Swiss traditions, the delicious experience of raclette and fondue.
I am also grateful to AUB, my alma mater, for the exceptional education and the opportunities it provided me. Additionally, I cherish the memories and the wonderful people I encountered during my time there, as they have played a significant role in shaping the person I am today.
A special place in my heart is reserved for the robotics club, where we created cherished memories and worked together to build, develop and grow the club.
Thank you Wael, for being a kind friend since the first day in Math 201.
Thank you Qusai, for being an awesome friend, for always finding a way to make me laugh, for sharing the joy of using references to funny Egyptian movies and series, for teaching me that taking risks can also be fun.
Thank you, Hiba, for being my best friend. From the moment we first met at Zaatar w Zeit, and with her planning all my courses throughout my entire bachelor's degree, she has always been there for me. Although we may be physically apart now, it always feels like we can pick up right where we left off.
To Sara, my childhood friend, thank you for being my pampa who truly understands me.
Though our paths have diverged, I am grateful to those who left an indelible mark on my journey, shaping the person I am today. Their enduring influence is a testament to the memorable moments and valuable lessons we've shared along the way.
To Basma, my sister by choice, thank you for being by my side, even when physically miles away. You bring smiles and positive energy into my life. Thank you for being my best friend, even after enduring the countless forever snoozing alarms in the dorm.
To my backbones, Sami, Anass, Rami, and Radi, thank you for always having my back, no matter where or when, for being an infinite source of humor that brightens my days. Together, we have created treasured core memories that we share and hold dear.
Thank you to my beloved mom and dad, my supporters since day zero. I owe them everything, and no words will ever be enough to describe my love.
Finally, the ultimate praise is to the one and only Allah, the most merciful and the most compassionate. Throughout my life I have been showered with His blessings. I ask Allah to give me the ability to praise Him for His bounties and make me able to use what He bestowed upon me to contribute to the betterment of this world.

Hala Khodr

Abstract

Order, regularities, and patterns are ubiquitous around us. A flock of birds maneuvering in the sky, the self-organization of social insects, a global pandemic or a traffic jam are examples of complex systems where the macroscopic patterns arise from the microscopic interaction of individual agents. Not only is the study of complex systems an important domain on its own but also represents a “powerful idea” that cuts across disciplines and can lead to the understanding of a large class of physical and social phenomena.

Swarm robotics is a field within robotics that models the collective behavior of decentralized and self-organized systems, primarily inspired by social insects in nature. The adaptability, scalability and resilience of swarm robots make them an appealing solution for a wide array of problems and applications across scales including search and rescue missions, large-scale logistics, environmental monitoring, entertainment and nanomedicine. A promising yet under-explored area is education.

In this thesis, we investigate a novel approach to learning with swarms, about swarms. We design, implement, and validate a comprehensive framework that can be used in formal as well as informal educational settings to foster understanding of complex systems. This includes the creation of an assessment instrument to measure complex systems understanding, the development of a platform for users to interact with multiple virtual as well as robotic agents, and the design and implementation of learning activities and educational games. These components are validated with experimental studies and analyses. We conducted a series of experiments with a total of over 200 participants spanning different age groups from high school students to adults. Our primary activity, named *Cellulan World*, designed as an individual human-swarm interaction game, demonstrated significant positive learning gains in complex systems understanding, as well as high reported engagement and enjoyment. The experiments also revealed interesting insights into the learning process and the user interaction aspects. Furthermore, we explore group activities in a “double swarm" setting where multiple humans interact with multiple robots. We delve deeper into the agent-agent interaction, focusing on different communication affordances, such as local and global verbal communication, as well as visual and haptic non-verbal communications.

In the context of this thesis, we re-designed the educational table-top swarm robot Cellulo, originally conceived in our lab, to allow for a modular architecture and broaden the scope of interaction with the robot. Simultaneously, we developed a library for our robotic platform using the cross-platform Unity game engine, simplifying the design and development of activities and games involving both humans and individual or swarms of robots. This also

allows for the seamless integration of tangible and virtual elements.
By offering an innovative and engaging educational experience, our framework presents a promising avenue to foster understanding and appreciation of the fascinating world of complex systems.

Keywords: Complex Systems Understanding, Distributed Mindset, Learning Analytics, Swarm Robotics, Educational Robotics, Human-Swarm Interaction

Résumé

Le monde qui nous entoure est empli de régularités, d'ordres et de motifs récurrents. Que ce soit une nuée d'oiseaux dans le ciel, l'auto-organisation des insectes sociaux, une pandémie à l'échelle mondiale ou un embouteillage, ce sont tous des exemples de systèmes complexes issus de l'interaction entre les individus à une échelle microscopique. L'étude des systèmes complexes est essentielle non seulement en tant que discipline, mais aussi parce qu'elle représente une idée transversale qui permettra de comprendre une multitude de phénomènes physiques et sociaux.

Dans le domaine de la robotique, la robotique en essaim s'intéresse au comportement collectif des systèmes auto-organisés et décentralisés, souvent inspirés par les insectes sociaux. Les robots en essaim sont attrayants grâce à leur adaptabilité, leur capacité à passer à l'échelle et leur résilience, rendant leur utilisation pertinente pour diverses applications allant de la recherche et sauvetage, la logistique à grande échelle, la surveillance environnementale, jusqu'au divertissement et à la nanomédecine. Cependant, leur potentiel dans le domaine de l'éducation est encore peu exploité.

Cette thèse explore une approche innovante d'apprentissage qui utilise les essaims, à propos des essaims. Nous avons développé, mis en œuvre et validé un cadre complet qui peut être utilisé dans des environnements éducatifs formels et informels pour favoriser la compréhension des systèmes complexes. Nous avons mis en place un outil d'évaluation pour mesurer la compréhension des systèmes complexes, développé une plateforme pour que les utilisateurs interagissent avec plusieurs agents virtuels et robotiques, et conçu des activités d'apprentissage et des jeux éducatifs. Nous avons validé ces éléments grâce à des études expérimentales et des analyses. Nous avons mené une série d'expériences avec plus de 200 participants de différents âges, des lycéens aux adultes. Notre activité principale, "Cellulan World", est un jeu d'interaction homme-essaim qui a montré des améliorations significatives dans la compréhension des systèmes complexes et un haut niveau d'engagement et de satisfaction. Les expériences ont également révélé des aspects intéressants du processus d'apprentissage et de l'interaction utilisateur. Nous avons également exploré des activités de groupe dans un cadre de "double essaim" où plusieurs humains interagissent avec plusieurs robots. Nous avons exploré l'interaction agent-agent, en se concentrant sur différentes modalités de communication, comme la communication verbale locale et globale, ainsi que les communications non verbales visuelles et haptiques.

Enfin, dans le cadre de cette thèse, nous avons repensé le robot en essaim éducatif de table Cellulo, conçu originalement dans notre laboratoire, pour permettre une architecture modulaire

et élargir la portée de l'interaction avec le robot. Parallèlement, nous avons développé une bibliothèque pour notre plateforme robotique en utilisant le moteur de jeu multiplateforme Unity, ce qui a simplifié la conception et le développement d'activités et de jeux impliquant à la fois des humains et des robots, individuellement ou en essaim. Cette bibliothèque facilite également l'intégration d'éléments tangibles et virtuels dans notre plateforme.
En proposant une approche éducative novatrice et engageante, notre cadre ouvre une voie prometteuse pour améliorer la compréhension et l'appréciation des systèmes complexes. En résumé, cette thèse offre une exploration enrichissante du monde fascinant des systèmes complexes à travers la robotique en essaim, tout en offrant des outils tangibles pour leur compréhension.

Mots-clés : Compréhension des Systèmes Complexes, Mentalité Distribuée, Analytique de l'Apprentissage, Robotique en Essaim, Robotique Éducative, Interaction Humain-Essaim

Contents

List of Figures

List of Tables

1 Introduction

Worldwide, A global pandemic.
Billions of people are affected, and the virus spreads without regard for borders or social structures. No single entity controls the virus; it is an invisible, decentralized force. Each individual virus particle is relatively simple in structure, but their collective action results in a global health crisis, disrupting societies and economies on an unprecedented scale. As the virus spreads, it mutates and evolves, challenging our understanding and response strategies.

The Human Body, The microscopic battleground within our bodies.
Countless immune cells work together in a decentralized network to protect us from pathogens and maintain our health. Each cell plays a specific role, recognizing and responding to invaders, yet the overall system functions without a central command. Through complex interactions and chemical communication, the immune system adapts, learns, and mounts a coordinated defense against threats.

Global Marketplace, The interconnected economy.
Billions of individuals and organizations interact through countless transactions, shaping a vast economic network that transcends national borders. No single entity controls this system; it is a decentralized web of supply and demand, driven by individual choices and behaviors. From these myriad interactions emerges a complex, dynamic system that allocates resources, drives innovation, and influences the course of human progress.

South America, The dense rainforests.
Eciton army ants are on the march, displaying a fascinating example of collaboration. There is no leader controlling this army; it doesn't have any commanding officer. When faced with obstacles such as gaps or uneven terrain, the ants create living bridges using their own bodies, enabling the rest of the colony to traverse the gap with ease. Every single ant possesses limited vision and a minimal level of intelligence, but their collective ingenuity in forming these bridges is optimal.

Southeast Asia, The twilight of a warm summer evening.
Thousands of fireflies light up the night, their bioluminescent flashes synchronizing in a seemingly choreographed display. No single firefly leads the group; instead, each firefly reacts to their neighbors' flashes, adjusting their timing accordingly. From these simple interactions emerges a captivating, collective light show that delights observers.

Lausanne, The open skies above.
Thousands of birds soar together, creating mesmerizing patterns as they fly in a coordinated dance. Despite the absence of a leader, the flock moves as one, reacting to changes in direction and speed with breathtaking precision. Each bird follows simple rules, responding to its neighbors' movements, yet the resulting behavior is a stunning display of emergent order and synchronization.

Beirut, The congested streets.
In the heart of a bustling city, countless vehicles navigate the complex network of roads, lanes, and intersections. As drivers attempt to optimize their own travel times, their individual actions can contribute to a more or less efficient overall system. Small disturbances can propagate through the system and lead to a cascade of slowdowns and eventually result in a congestion, often difficult to predict or prevent. The driver, already late to work, is left to ponder the many mysteries of the perplexing origin of the traffic jam.

The diverse examples explained above- the COVID-19 pandemic, the immune system, the economic system, ants colonies, flocks of birds, synchronization of fireflies, traffic jams, and many more - all demonstrate the remarkable phenomena that can emerge from complex systems. Local interactions between simple agents give rise to complex behaviors, illustrating the power of self-organization and decentralized control.

Concretely, complex systems can be defined as

> an interdisciplinary field of research that seeks to explain how large numbers of relatively simple entities organize themselves, without the benefit of any central controller, into a collective whole that creates patterns, uses information, and, in some cases, evolves and learns. (Mitchell, 2009)

Given the ubiquitousness and importance of these systems in our world, it is no surprise then that in 2021, the Nobel Prize in Physics was awarded to three laureates for their pioneering contributions to our understanding of complex physical systems. Syukuro Manabe and Klaus Hasselmann jointly received recognition for their foundational work "for the physical modelling of Earth's cimate, quantifying variability and reliably predicting global warming". On the other hand, Giorgio Parisi was recognized for his revolutionary contributions to "the discovery of the interplay of disorder and fluctuations in physical systems from atomic to planetary scales" [1].

[1] https://www.nobelprize.org/prizes/physics/2021/popular-information/

As the examples above suggest, not only is this an important domain to learn about on its own, but it actually consists of a “powerful idea" (Papert, 1980) that cuts across disciplines and can lead to the understanding of a large class of physical and social phenomena taught as conceptually different subjects. Researchers argue that different complex systems in nature, such as insect colonies, brain activity, immune systems, economies, the atmosphere, and human social interactions, have much in common. Understanding complex systems is fundamental to understanding natural and social sciences. In a book discussing the vital cognitive skills for the 21st century, complex systems are highlighted as one of the critical topics in education, alongside computational and mathematical thinking and continuing education (Reed, 2020). This underscores the significance of the topic of complex systems in fostering a deeper understanding of the world around us.

1.1 Common Misconceptions

Complex systems can be found all around us in nature, from the smallest cellular organisms to the largest ecosystems. They are a fundamental part of our daily experiences and interactions with the world. Yet, although they are commonplace, thinking about complex systems is anything but intuitive, and experts and novices think and build their knowledge about complexity in significantly different ways.

Novices primarily concentrate on visible and static structures within subsystems, while experts consider not only the structures, but also causal behaviors and functions (Hmelo-Silver et al., 2007). This distinction in cognitive and ontological approaches to complex systems is also supported by Jacobson (2001), who found that novices typically view complex systems as centralized, predictable, and reductive, whereas experts describe them as decentralized, nonlinear, and non-reductive.

Learning about complex systems can be challenging (Hmelo-Silver & Azevedo, 2006). Numerous studies have been conducted to investigate the difficulties that novices face in comprehending complex systems and the misconceptions they often hold.

Linear and Direct Effects

Novices often assume that relationships within complex systems follow a simple cause-and-effect pattern (Gotwals & Songer, 2010). However, the reality is far more intricate, and understanding the dynamics of complex systems requires a nonlinear approach. In these systems, small actions can lead to widespread and exponential effects at various levels within the system due to the numerous connections among components. For example, novices may struggle to understand how a change in one part of a food chain can impact another part that isn’t directly connected (Gotwals & Songer, 2010).

One common issue is that novices usually focus on immediate effects, overlooking the more complex cascading or indirect effects (Grotzer et al., 2015). They often don't realize that a change in one population can have consequences for other populations not directly connected, through a series of complex interactions. Some phenomena, like climate change, span vast distances and involve relationships between causes and effects that are difficult for learners to grasp, as they exist in different attentional frames (Grotzer & Shane Tutwiler, 2014).

Moreover, novices may face challenges understanding complex causality and nonlinear dynamics (Chi et al., 2012). They often rely on direct-causal schemas, which attribute behaviors and outcomes to immediate one-to-one interactions, rather than considering emergent schemas that recognize the interconnected and web-like nature of systems producing nonlinear effects.

In summary, people often have difficulty comprehending indirect cascading effects, assuming instead that effects are immediate and directly related to causes (Grotzer et al., 2015). It is essential to recognize that complex systems involve emergent properties, which are interconnected and can have far-reaching impacts (Chi, 2005).

Deterministic Thinking

In deterministic systems, the same initial conditions consistently produce the same outcomes, enabling accurate prediction of future events based on these conditions. Novices generally perceive processes as direct and static phenomena, rather than emergent (Chi, 2005). This deterministic thinking leads to a bias in which novices assume that the behavior of agents within a system is predictable.

However, complex systems do not follow deterministic behavior. Instead, they frequently exhibit nondeterminism and probabilistic behavior. Nondeterminism implies that the outcome of a specific event or system is not predetermined. Probabilistic behavior pertains to the likelihood of various outcomes occurring. In complex systems, both nondeterminism and probabilistic behavior play crucial roles, as interactions between components can be dynamic, unpredictable, and influenced by chance. Probabilistic, non-deterministic behavior in systems is indeed a challenging concept for novices to grasp (Wilkerson-Jerde & Wilensky, 2015).

Figure 1.1, a comic picture of a bird selecting its spot in the flock, as if it were on a plane, humorously illustrates the contrast between deterministic behavior and the actual decentralized nature of bird flocking. In deterministic systems, the same initial conditions consistently produce the same outcomes, much like the assigned seating on a plane. However, bird flocking is a complex, emergent phenomenon that arises from the decentralized interactions of individual birds, with no predetermined positions.

Interestingly, complex systems can exhibit emergent properties that are not entirely predictable from the behavior of their individual components. From a complex systems perspective, it is essential to recognize that despite the presence of random factors in natural

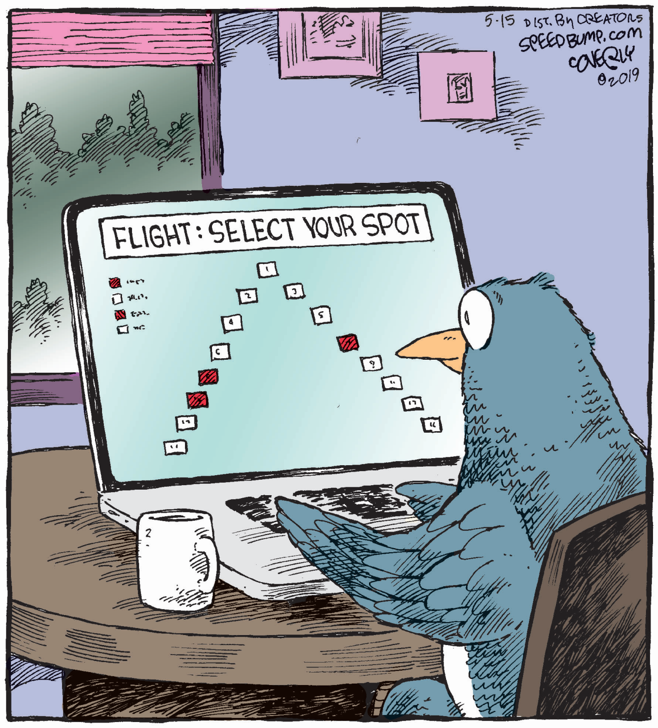


Figure 1.1: Comic - Birds' Boarding Pass: When deterministic thinking meets decentralized reality. Credits goes to Dave Coverly (https://www.speedbump.com), used with permission.

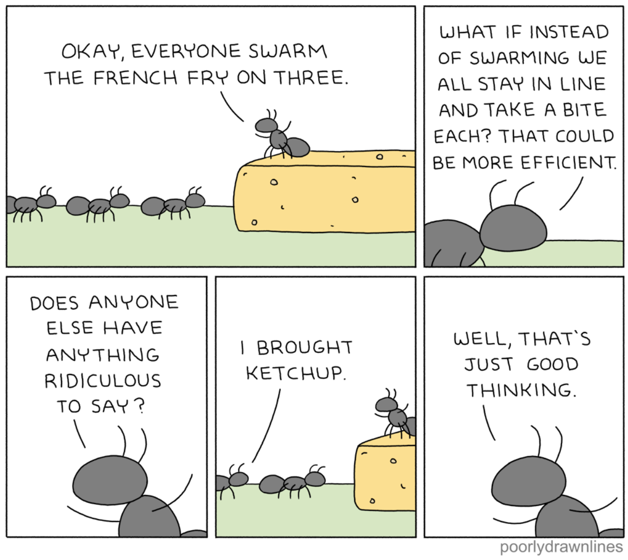


Figure 1.2: Comic - Ants' Misconception: When centralized control meets decentralized teamwork. Credits goes to Reza Farazmand (https://www.poorlydrawnlines.com), used with permission.

phenomena at all levels, emergent patterns and self-organization still occur. This realization underscores the importance of embracing a more nuanced understanding of systems and their behavior, moving beyond deterministic thinking to better comprehend the complexity and unpredictability inherent in many real-world scenarios.

Centralized Mindset

During one of my experiments, a 15-year-old student remarked, "*There is always a leader such as a principal in a school.*" This statement underscores a prevalent misconception where novices attribute a system's behavior to a single leader or pre-existing entity in the environment, rather than acknowledging the potential for decentralized interactions.

As defined by Resnick (1997), a centralized mindset involves the belief that global patterns of a system are governed directly by lead - a leader orchestrating the system's patterns - or seed - a single pre-existing entity in the environment. This perspective overlooks the role of decentralized interactions in shaping system behavior. Such a mindset can impede students' understanding of complex systems, as they may not recognize that control in systems often stems from decentralized processes, with higher-level structures or behaviors emerging from lower-level system activities. For example, the balance of predator-prey populations in healthy ecosystems arises from the collective interactions of predators and prey (Yoon et al., 2019). Figure 1.2 is another comic picture of an ant instructing the others to swarm the french fry on the count of three with a humorous follow up discussion between ants. It very nicely demonstrates the contrast between the idea of a centralized control that people imagine and the actual decentralized nature of ant swarming behavior. In reality, ants work collectively and communicate via pheromones to forage for food, without a designated "leader" giving orders.

It is crucial to emphasize that centralized ideas are not inherently wrong or harmful. Some phenomena can be accurately described by centralized theories, and centralized strategies may prove useful in developing new technologies or organizations. The problem emerges when individuals depend almost exclusively on centralized strategies, thereby undervaluing or overlooking decentralized approaches. Our objective should not be to replace the centralized mindset but to supplement it, leading to a richer set of mental models for interpreting systems in the world (Resnick & Wilensky, 1998).

1.2 Swarm Robotics for Learning

The Robots for Learning (R4L) concept was born out of the vision of Seymour Papert in the 1980s. Papert envisioned programmable entities, like the Logo Turtle, that could serve as "computational objects to think with", aligning with Piaget's constructivist theory that learners actively construct their knowledge. This notion was deeply rooted in the emerging "computer culture" of the era, leading to the idea of enabling learners to program simple robots, thus giving birth to the programmable learning robot movement (Papert, 1980).

The R4L domain has since branched into two distinct directions. The first, known as “robots as tools”, involves custom-designed robotic platforms primarily intended for teaching programming and related to Science, Technology, Engineering, and Mathematics (STEM) concepts (Karim et al., 2015). The second direction, developed alongside advances in the Human-Robot Interaction (HRI) field, envisions robots functioning as socially capable peers or tutors (Johal, 2020).

Expanding on the “robots as tools” approach, it involves providing learners with actual robots that they can physically construct and program. This concept is underpinned by several educational principles, such as learning with robots (Rogers & Portsmore, 2004), hands-on interaction (Mataric et al., 2007), understanding individual parts (Miglino et al., 1999), and evolutionary dynamics (Auerbach et al., 2014). These principles foster active engagement, experiential learning, and comprehension of complex systems through the prism of robotics. Additionally, the concept of “artificial organisms” introduces learners to evolutionary principles, enabling them to refine robot behaviors or solutions through processes that mimic natural selection and mutation (Auerbach et al., 2014).

Over time, a variety of programmable robotic platforms have been utilized as learning tools, catering to diverse age groups and skill levels. These robots range in their buildability, with some being completely constructible and sold as kits to stimulate learners’ creativity and innovation. The advent of the "maker movement" further broadens this approach by offering affordable rapid prototyping platforms and lightweight computers, primarily based on open-source hardware and software.

On the other hand, swarm robotics is a field within robotics focusing on the collective behavior of decentralized, self-organizing systems, and is largely inspired by social insects (Bayindir & Şahin, 2007). Their inherent qualities of adaptability, scalability, and resilience make swarm robots a versatile tool for numerous applications. These include but are not limited to, search and rescue operations, large-scale logistics management, environmental surveillance, entertainment, and the emerging field of nanomedicine (Dorigo et al., 2020).

A promising yet largely unexplored potential for swarm robots is in the realm of education (Johal et al., 2020). The collective behavior of these robots, dictated by individual interactions, could serve as an impactful teaching tool for imparting an understanding of complex systems, an essential skill in today’s digital age. As emphasized thus far, complex systems play a crucial role in numerous disciplines and real-world scenarios, rendering it vital for people to achieve a comprehensive understanding of them.

Swarm robotics offers a potentially enriching avenue for enhancing the understanding of complex systems. The inherent parallels between swarm robotics and complex systems, such as emergent behavior, nonlinearity, and adaptability, provide a natural fit for using swarm robots to model complex systems. Moreover, the hands-on, interactive nature of swarm robots can provide students with a tangible and engaging learning experience, allowing them to observe and manipulate the behaviors of individual agents in the swarm, and understand how these behaviors interact to produce emergent phenomena at the macro level.

1.3 Thesis Objectives

The primary objective of this thesis is to investigate the potential of utilizing swarm robots to foster complex systems understanding. By proposing a novel approach to learning with swarms about swarms, we aim to uncover the pedagogical opportunities that swarm robots offer in the domain of complex systems understanding.

To achieve this overarching goal, our methodological approach is to design, implement, test, and validate a comprehensive learning framework that can be applied in both formal and informal learning settings. This framework is designed to address misconceptions related to complex systems understanding, with a particular emphasis on the centralized/distributed mindset, while acknowledging the interconnected nature of these misconceptions.

Just as the notion of complex systems represents a potent and transferable idea that spans across multiple disciplines, this work contributes to and intersects with various research domains, including learning sciences, robotics, and human-computer interaction. In order to achieve the objective of enhancing complex systems understanding through swarm robotics, our methodology necessitates delving into all three domains for a comprehensive exploration.

In the field of learning sciences, we focus on designing and validating assessment and learning tools that effectively engage learners and foster their understanding of complex systems. This involves understanding the pedagogical principles and practices underlying effective learning experiences, as well as evaluating the impact of swarm robotics on students' comprehension of complex systems concepts.

In the realm of robotics research, we concentrate on the development and refinement of swarm robotics technology, specifically the modular version of Cellulo (Özgür, Lemaignan, et al., 2017), a tangible mobile educational robot - allowing for use of swarm in educational activities. This includes enhancing the robot's hardware and software interaction capabilities. Additionally, we explore the principles of swarm behavior, communication, and control, as well as the potential applications of swarm robotics in educational settings.

Lastly, within the domain of human-computer interaction, our work investigates the design and analysis of interactions involving humans, individual robots and swarms, both real and virtual. This exploration allows us to better comprehend the dynamics between humans and robots in educational settings and identify factors that contribute to effective swarm-based learning experiences.

By bridging these diverse fields, this thesis aims to advance our understanding of the pedagogical potential of swarm robots and their role in fostering a deeper comprehension of complex systems.

1.4 Thesis Structure

In the next chapter (Chapter 2), we present a literature review to provide context for our research. Specifically, we survey the use of swarm robots in education, we offer an overview of the tools and resources available in the literature for complex systems understanding and emphasize the primary learning strategies and theories that underpin our subsequent studies.

The remainder of this thesis is divided into three primary parallel sections. While Part I forms the core of the thesis, Parts II and III offer exploratory pathways into group activities, and enhancements and expansions of the robotic platform, respectively.

Part I presents and explores an individual learning activity where a single participant interacts with a swarm to learn about complex systems. This section includes four chapters (3-6). Chapter 3 introduces the assessment instrument developed to evaluate students' complex systems understanding and provides insights into this comprehension. Chapter 4 presents the iterative design of the framework used and the primary activity. Chapter 5 details the first study, examining the effectiveness of the designed activity. Chapter 6 reports a second study that investigates the efficacy of our activity with participants across different age groups.

Part II examines group activities involving two or more people and the potential interactions. Chapter 7 describes a classroom activity designed to investigate how diseases spread, an example of a complex system. Chapter 8 studies the impact of various communication affordances on the emergence of collaboration strategies though an online game.

Part III comprises the work conducted on the robotic platform, containing two chapters. Chapter 9 outlines the modular robot design and the SDK. Chapter 10 discusses the expansion of potential interactions through a VR connection, an online connection, and the possible extension to reconfigurable swarm robots.

Finally, in a synthesis chapter, we summarize the primary findings and contributions of this thesis, discuss its limitations and propose directions for future research.

2 Literature Review

In this chapter, we aim to establish the theoretical foundations of the thesis, by examining the state of the art and relevant literature. The chapter is organized into five sections. The first section focuses on the utilization of swarm robots as an educational tool, exploring their potential to enhance learning experiences in complex systems education. The second section introduces the robotic platform employed in this thesis. The third section delves into existing tools for learning complex systems, examining their strengths and limitations in facilitating students' understanding of such systems. In the fourth section, we evaluate learning strategies that can later inspire the design of our instructional approaches. Specifically, we will discuss game-based learning and tangibility in learning, both of which hold promise for engaging students and fostering deeper understanding of complex systems. Finally, particular focus is given, in the last section, to the learning theory of conceptual change. We will discuss the idea of conceptual change, which plays a pivotal role in addressing students' pre-existing alternate conceptions and facilitating the construction of accurate and coherent concepts. By examining these key areas, we aim to provide a comprehensive foundation for the development of effective learning activities featuring swarm robots and aiming to support the understanding of complex systems.

2.1 Swarm Robots in Education

Swarm robotics is a field of study that builds on the research of Multi-Robot Systems (MRS) by utilizing principles of communication, coordination, and collaboration among a substantial number of robots (Nedjah & Junior, 2019). In recent years, there has been a remarkable increase in interest towards swarm robotics. Numerous swarm platforms have emerged, focusing on both hardware and software system design. These platforms encompass three main types of robots: ground-based robots, robots capable of climbing vertical structures, and robots that fly within the environment, and spanning multiple applications from search and rescue to nanomedicine (Johal et al., 2020).

Following taxonomies and characterization from the literature (Farinelli et al., 2004), swarm robots are agreed to be under the category of strongly distributed multi-robot systems in terms of hardware design. However, in this thesis, we choose to employ the term “swarm robots” as a broad umbrella term, i.e., also including robots that, while not distributed in their architecture, exhibit a distributed behavior to the user. This decision is made to emphasize the perception of "swarm-ness" from the user’s perspective rather than the developer’s standpoint. Depending on the specific activity and its objectives, the collection of robots can intentionally exhibit behaviors that resemble those of a swarm, regardless of the way such behaviors are generated and encoded in their control architectures, be it truly distributed or not.

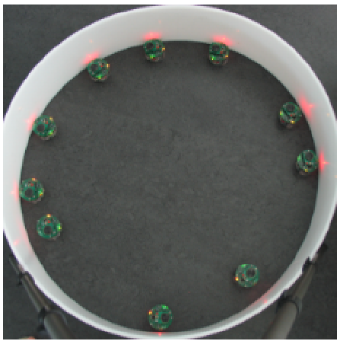
We can find three main domains of use of swarm robots in education:

Robotics Education

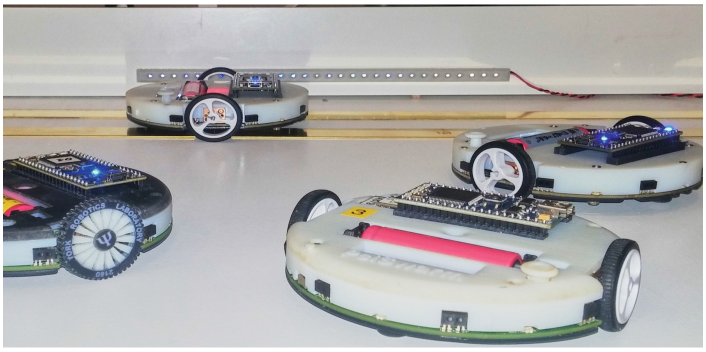
Centered on the core principles of simplicity, affordability, and replaceability, swarm robotics has emerged as a valuable teaching tool in K-12 and undergraduate education for robotics. Some platforms have primarily been developed at universities (and later commercialized) for teaching engineering courses following the line of “robots to teach robotics.” One example is the e-Puck, which was successfully designed at EPFL and adopted in numerous other robotics laboratories worldwide. The e-Puck (Mondada et al., 2009), a low-cost, small-scale mobile robotic platform designed for education, allows for the incorporation of real robotic hardware into the classroom to teach swarm-robotic concepts. In their class “Swarm Intelligence” at EPFL, Cianci et al. (2007) developed a custom module for local radio communication as a stackable extension board for the e-Puck, enabling information exchange between robots. This helped students not only use the simulation software in their class but also apply it on real robots, such as in a project of collective decision-making among a group of 10 robots (Figure 2.1a). Another example is Pi-swarm (Hilder et al., 2016), designed to reduce costs and construction complexity; this trackable, sensor-rich, and expandable platform requires only an internet browser to program and includes a tablet-computer-based programming environment aimed at teaching programming concepts to primary school-aged children (Figure 2.1b).

Other platforms include open-source miniature swarm robots, such as Kilobots (Rubenstein et al., 2014) in Figure 2.1d and Alice (Caprari et al., 2000) in Figure 2.1c to name a few, which were primarily designed for swarm-related research. However, their size and production costs increase the potential of making them well suited for K-12 STEM education (Karim et al., 2015). In this context, each student can have their own robot, encouraging collaboration with classmates and educators during interactive learning activities. Their hardware design is open-source, and the necessary COTS (Commercial Off-The-Shelf) components are easily obtainable (Karim et al., 2015).

MicroMVP (Yu et al., 2017), standing for micro-scale Multi-Vehicle Platform, is a testbed for multi-robot planning and coordination algorithms, as well as an educational tool for teaching mobile robotics and multi-robot systems. Enabled by 3D-printing and the maker culture, it is

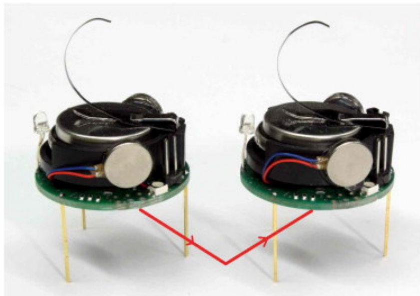


(a) E-puck robots from (Cianci et al., 2007)



(b) A swarm of Psi Swarm robots from (Hilder et al., 2016)

(c) Alice miniature robots from (Caprari et al., 2000)

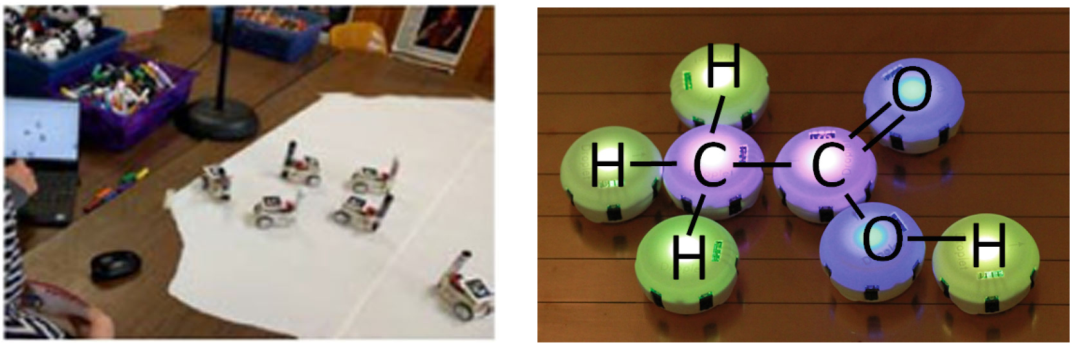


(d) Kilobots from (Rubenstein et al., 2014)

Figure 2.1: Examples of some of the swarm robots used in engineering education.

portable, affordable, low-maintenance, and an open-source multi-vehicle platform. The open-source effort aims to engage the community in promoting robotics research and education.

Lastly, MOSAIX (Alhafnawi et al., 2022b) is a novel robot swarm platform designed for social settings. Comprising up to 100 individual robot tiles, MOSAIX is specifically developed as a social swarm system, aiming to assist humans in social tasks like opinion-mixing and brainstorming. It has already been utilized outside laboratory environments, gathering 154 opinions on climate change in a local shopping mall, enabling collaborative art creation, and serving as an educational tool for school children. Furthermore, in educational settings, MOSAIX has been used as an engaging and innovative educational tool for teaching children about robotics. By dividing an image into equal pieces and using a swarm of 9 robots, children were tasked with assembling the puzzle by stopping the moving robot tiles and aligning them correctly. This activity facilitated discussions about swarm robotics, sensors, motors, and sparked curiosity about space robotics and STEM careers. The younger children demonstrated enthusiasm and amusement while solving the “robotic puzzle”, highlighting the potential of early exposure to robots in shaping their interest in STEM fields from an early age (Alhafnawi et al., 2022b).



(a) MicroMVP robots used in a theater activity from (Barnes et al., 2020)

(b) Chemistry tangible simulation with droplets robots from (Randall et al., 2016)

Figure 2.2: Examples of some of the swarm robots used in non-engineering subjects.

Other STEM disciplines

Non-robotics subjects received significantly less attention within educational swarm robotics.

In an artistic context, robot swarms are used in a child-robot theater (Barnes et al., 2020). By incorporating swarm robots in an afterschool module, the program introduced children to the concept of computer-aided design and allowed them to actively engage in creating artwork collaboratively with the MicroMVP robots (Figure 2.2a). In successive modules focused on the story of *Beauty and the Beast*, children were first introduced to the story and robots, cast themselves and the robots in roles, and wrote dialogues for scenes. Children focused on character and set design, drawing robots as story characters and using geometric shapes to draw the Beast's castle. They also used swarm robots to create background scenery.

Randall et al. (2016) present a tangible simulation for teaching basic rules of chemistry using a swarm of miniature robots, called Droplet, mimicking atoms and forming molecules (Figure 2.2b). The swarm robotics platform and augmented reality (AR) interface provide a tangible and interactive learning experience for high school chemistry students (Batra et al., 2022). Each Droplet in the swarm represents an atom and communicates relevant data about itself using electronegativity values. When Droplets come into proximity with each other, they exchange information and determine if bonds can form based on their electronegativity values. The educational value of this simulation is enhanced through AR, which allows users to visualize and differentiate individual molecules, eliminating the need to remember colors and look up statistical data for each atom.

Özgür (2018) presents an educational activity about the "particles of matter", demonstrating the behaviors of particles in different states of matter using a swarm of tabletop robots (Figure 2.3). This study was one of the first, among the studies about tangible swarm robots in a non-engineering subject, to provide a complete experiment studying learning gains of students engaging in the activity, and revealing weaknesses, strengths, and added value of the robots. The study analyzed the learning gains of students regarding the macroscopic and microscopic views of matter. It found that the students improved their understanding of particle motion and spacing. The study also found that students successfully abstracted particles in matter using robots, referring to them as "particles," "atoms," or "molecules" without mentioning the robots.

Complex systems

More recently, some studies have started to tackle the idea of using a swarm of robots to help students understand complex dynamic systems. One such example is a demo (Gourlet et al., 2017) using the Zooids robots to develop an activity inspired by the “Lemmings" game[1], where each robot can be assigned an action that will affect the system behavior. The demo thus invites children to resolve situations by changing individual actions in a dynamic system. Another example is presented by Vitanza et al. (2019), where two case studies are discussed: (i) the collective decision-making problem inspired by house-hunting honeybees (Seeley et al., 2012), and (ii) a collaborative task inspired by the classic stick-pulling experiment (Ijspeert et al., 2001). Preliminary results from a demo setting during a robotics festival are reported to be "very promising", with participating kids understanding "that cooperation and consensus are not easy to achieve, and that some flexibility in one’s opinion is a necessary condition". Although these events typically attract a large number of participants, learning evaluation is often not feasible, and researchers rely on acceptability questionnaires. The studies reviewed have reported positive outcomes in terms of acceptability but do not demonstrate a clear improvement in learning compared to other instructional methods.

Conclusion

Overall, while swarm robotics research has primarily focused on designing robot systems, its potential in education is now being increasingly explored. Integrating swarm robotics into educational settings opens up new avenues for interactive and engaging learning experiences, empowering students to develop essential skills for the future. In the realm of education, research is conducted to explore opportunities for using swarm robots to teach a wide range of subjects, including engineering disciplines such as robotics and programming. Additionally, swarm robotics has the potential to enhance learning in science subjects like chemistry and physics, as well as in the arts. Lastly, it offers a pathway to develop a deeper understanding of complex systems.

In their review, Johal et al. (2020) provide an overview of recent research in swarm robots for educational purposes by examining papers that offer empirical evaluations. One of their main findings, after collecting and filtering the papers that empirically evaluated learning scenarios, is that despite the extensive literature describing the hardware and software of novel swarm robotics systems designed for education, very few studies have actually conducted user experiments with the specific goal of testing their platforms in educational settings

[1] https://en.wikipedia.org/wiki/Lemmings_(video_game)

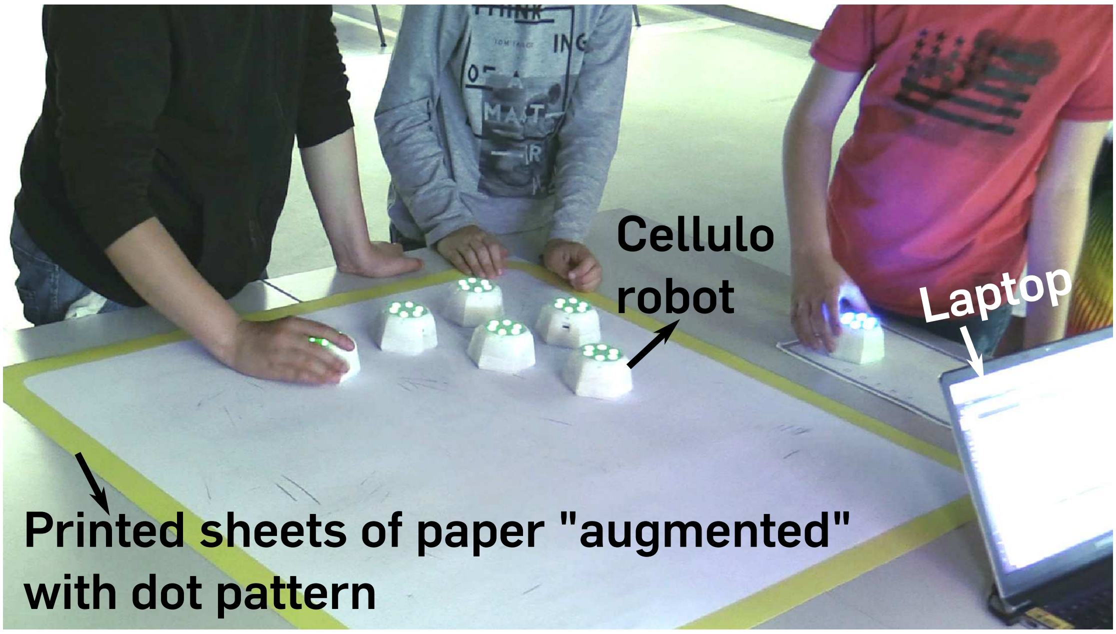


Figure 2.3: The Cellulo robotic platform in action, demonstrating its components: the robots, the printed map for localization, and the PC/tablet running the interaction activity. Here, the platform is facilitating a learning activity on particles in matter from (Özgür, 2018)

and demonstrating learning outcomes. In many cases, education is discussed as a potential application, or only suggestions are provided without concrete empirical evidence.

In this thesis, we adopt a more comprehensive approach rooted in the learning sciences. We not only focus on developing the technological platform and providing insights and suggestions for swarm robotics, but we also close the loop by adopting a learner-centric evaluation approach. To demonstrate the added value and the effectiveness of our activities, we conduct experiments that specifically evaluate their impact on learning outcomes. By taking this approach, we aim to contribute to the existing body of knowledge by providing empirical evidence of the educational benefits of swarm robotics identified from our research.

2.2 Cellulo Robotic Platform

Within the context of this thesis, we utilize the Cellulo robotic platform as the central swarm robotic framework. This platform was developed at the Computer Human Interaction for Learning and Instruction (CHILI) lab at EPFL (Özgür, 2018).

Cellulo robots serve as versatile agents capable of representing both abstract and concrete objects. By manifesting various properties through kinematic motions, such as velocities and accelerations, as well as behaviors in collective groups, these robots provide enriched learning experiences for students. The tangible presence of Cellulo robots enables direct, hands-on, and haptic interaction with learning activities, rendering intangible concepts into accessible, tangible experiences.

Since their inception, Cellulo robots have been utilized in numerous studies spanning diverse educational domains, such as handwriting (Guneysu Ozgur et al., 2020), mathematical concepts like linear functions (Khodr, Kianzad, et al., 2020), symmetry (Johal, Andersen, et al., 2019), chemistry topics related to particles in matter (Özgür, 2018), physics concepts such as wind fields (Özgür, Johal, et al., 2017), and computational thinking (Nasir et al., 2019).

Designed with practicality and versatility in mind, the Cellulo platform provides the necessary affordances for our application in this thesis. The robots are easy to deploy and operate, requiring only the robots themselves, a printed map for localization, and a PC/tablet to run the application or interaction activity (Figure 2.3). Cellulo robots are handheld, haptic-enabled mobile devices capable of holonomic motion and absolute global localization (position and orientation (*x,y,theta*)). When placed on an "augmented" paper sheet, each robot can self-localize with sub-millimeter accuracy (<1 mm). Additionally, each unit is equipped with six capacitive touch sensors and six full RGB LEDs on its top surface, allowing for straightforward visual and tactile interaction.

This thesis significantly contributes to the evolution of the Cellulo platform by providing a Software Development Kit (SDK) for application development, which is implemented in the creation of all games and learning activities presented within this work. Furthermore, it broadens the range of interactions and stregthens the platform's versatility and practicality by incorporating modularity into its structure. The details of these enhancements will be further discussed in Part III of this thesis.

2.3 Tools for Learning Complex Systems

Although the use of swarm robots in education is still in its infancy, a more active research area within the learning sciences for teaching and learning complex systems involves the development of computational tools and resources. These tools can be broadly classified into two categories: agent-based modeling tools and participatory simulations.

2.3.1 Agent-Based Modeling Tools

Agent-based modeling (ABM) tools allow users to create, explore, and analyze complex systems through the simulation of individual agents and their interactions. These tools provide a powerful way to understand emergent properties and phenomena within various systems, including biological, social, and physical systems. One influential application of ABM in the field of social sciences is Thomas Schelling's work on residential segregation (Schelling, 2006). Schelling's research explored how individual preferences for living near others of the same race, even when relatively mild, can lead to significant patterns of segregation at the macro level. By incorporating Schelling's insights into ABM, researchers can create sophisticated simulations that capture the dynamics of residential segregation, allowing for a deeper understanding of the emergence and persistence of such phenomena within societies.

ABM tools thus provide a valuable means of studying and exploring the complex interplay between individual behavior and collective outcomes in various domains. Examples of such tools used in educational contexts include NetLogo, StarLogo, and AgentSheets.

AgentSheets by Repenning and Sumner (1995) developed in 1995 and later upgraded to 3D with AgentCubes (Repenning & Ioannidou, 2006; Repenning, 2016), is an agent-based simulation-authoring tool that uses visual programming approaches, enabling users without programming experience to create and publish interactive simulations as web apps.

Designed more specifically to model and explore the workings of decentralized systems, StarLogo (Resnick, 1996) began at the MIT Media Lab in 1989. StarLogo is an extension of the original Logo programming language (Papert, 1980). In traditional versions of Logo, students create pictures and animations by giving commands to a graphic “turtle.” StarLogo extends Logo in three major ways: it has many turtles, the turtles have behaviors, and it reifies the turtles’ world. StarLogo builds on the idea of accessible agent-based modeling (Resnick, 1996). StarLogo TNG (Begel & Klopfer, 2007) uses a block-based programming language that simplifies the coding process for users with varying levels of experience

A descendant of StarLogo, NetLogo has become the most widely used agent-based modeling environment in education and research. It features a robust programming language and a user-friendly graphical interface, making it suitable for a wider range of applications and user expertise levels. NetLogo provides an extensive library of pre-built models that can be customized to fit specific needs, and it has been widely adopted for research, teaching, and learning purposes across various disciplines (Wilensky, 1999).

Such environments help students construct a connection between the interactions of micro-level particles/agents and the emergent complex macro behaviors. These learning environments are based on the assumption that whenever students have the opportunity to take the point of view of an atom, an electron, a sheep, a fish, an ant, or a trader, they are more ready to conceive complex systems (Wilensky & Rand, 2015). Learners investigate various agent types and their corresponding action rules to observe emerging phenomena. To create an ABM for a specific phenomenon, the learner breaks down the phenomenon into its micro-level elements, or agents, and establishes rules governing agent actions and interactions to produce the macro-level emergent phenomenon (Jacobson & Wilensky, 2022). These models serve as constructionist microworlds (Papert, 1980), enabling students to manipulate computational agents’ behaviors to implement operations and simulate microworlds. Students can witness the consequences of their manipulations on the emergent patterns at the system level through several interconnected computational representations, such as direct visual representations of a system, graphs, and numerical values. An example for a bee hive finding activity in Netlogo is show in Figure 2.4a. Multiple case studies have been conducted with students learning from microworlds. Examples include middle and high school students deducing the ideal gas law from micro-level interactions of gas particles within a container, investigating and modeling evolutionary processes, and simulating firefly synchronization (Jacobson & Wilensky, 2022).

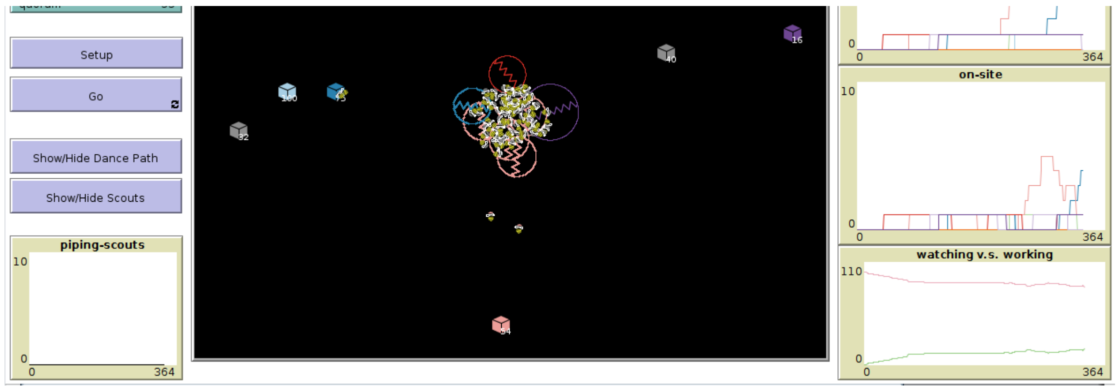
In conclusion, the agent-based modeling approach allows learners to assume an agent's perspective and strive to construct emergent explanations for macro-level phenomena. In this approach, users are observers who explore and analyze a simulation where each agent adheres to specific rules of action and interaction. However, a considerable challenge persists: the task of establishing a clear connection between the perspective of individual agents and the emergence of broader, macro-level phenomena. This difficulty is indicative of an enduring struggle to overcome a "deterministic-centralized mindset" (Jacobson & Wilensky, 2022; Resnick, 1997).

2.3.2 Participatory Simulations

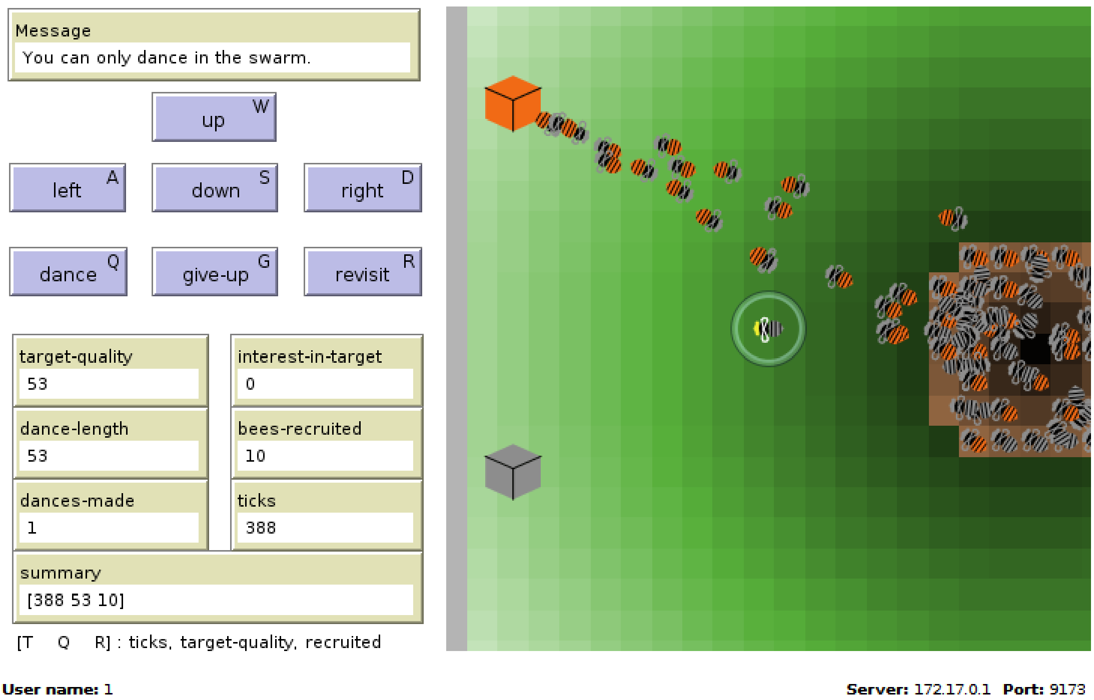
One approach studied in the literature to help novices move beyond the centralized mindset is to contextualize the discussion of complex phenomena within everyday scenarios, enabling students to go beyond being mere observers and become active participants. A participatory simulation is a type of experiential learning activity where participants take on roles as agents and engage in a simulated environment. In what follows, various examples of participatory simulations found in the literature to be used in the context of complex systems learning are reviewed.

Repenning et al. (2010) developed "Collective Simulations" that help students learn about the intricacies of interdependent complex systems by engaging them through social learning techniques in the classroom along with networked computers. A prototype of Collective Simulations called Mr. Vetro was implemented and evaluated in local high schools to teach students about physiology through technology-enhanced role-play. In this setup, human systems or organs are simulated on wirelessly connected computers, and students control different organs such as the heart, lungs, or activity levels. As students change parameters, the central simulation updates visually and audibly, encouraging collaboration and scientific discourse between student "organs" to maintain homeostasis. The curriculum was designed in collaboration with medical and educational experts, and the evaluation study included five science teachers and 400 high school students. Results showed that Mr. Vetro improved teacher performance, student learning, and engagement, especially for less advanced students. The simulation also promoted self-awareness and healthier lifestyles among participants.

Similarly, Hubnet (Wilensky & Stroup, 1999) was developed as a type of role-playing activity used in classrooms to explore how complex dynamic systems evolve over time through individual elements acting. HubNet is a technology that employs NetLogo to conduct participatory simulations in the classroom. In these simulations, an entire class participates in representing a system's behavior, with each student controlling a portion of the system through an individual device, such as a networked computer (Wilensky & Stroup, 1999). BeeSmart (Guo & Wilensky, 2016) is one example of an activity using Hubnet designed to teach honeybees'



(a) NetLogo Version: Students observe the simulation and can use the sliders to manipulate the different variables such as the initial scout percentage, or the initial explore time. Different plots show the convergence over time. For more advanced students, the NetLogo version even allows to code/modify some of the agent's behavioral rules.



(b) Hubnet Version: Students embody the role of bee scouts. In addition to some simulated bees, each student can have control over one bee, and can perform dances by pressing a button. They can also take the decision to follow others by clicking a nearby dancing bee.

Figure 2.4: An example of the bee hive finding activity in (a) an ABM approach (Netlogo), and in (b) participatory simulation approach (Hubnet). Initial scouts fly in different directions in an exploration phase. If one scout finds a potential hive site, it flies back to the swarm to advertise its location and quality by performing a waggle dance. The better the quality, the longer is the dance, the more likely idle bees join the inspection of the newly advertised sites and express their opinion similarly through waggle dances. Therefore, high quality sites will attract more and more bees until the winning hive site is identified by unanimous decision.

hive-finding behavior (Figure 2.4). It allows students to act as individual bees seeking hive sites in a virtual landscape. The activity begins with a driving question: How can a swarm of bees pick the best hive site from many choices? As the simulation progresses, the teacher reduces the bees' vision radius, forcing students to consider the limited abilities and rules followed by real bees. Discussions and aggregated data from the simulation help students understand the randomness and patterns involved in complex systems.

Other studies used mixed reality technologies such as in (DeLiema et al., 2019), where learners use their bodies to imaginatively become microscopic atoms inside a collaborative simulation of state change (solid-liquid-gas). Kinect cameras are installed in the classroom and particles are overlaid on the corresponding positions when displayed on a projected overview.

Another type of participatory simulations involves the uses of multiplayer games. For example, City Settlers (Kumar et al., 2021) is a city management game that supports immersive and participatory learning, fostering various types of collaboration and promoting a complex systems understanding of sustainable development. City Settlers is a whole-room immersive simulation that transforms a physical space into a shared planet, where participants develop their cities using mixed reality and face-to-face learning. A study by Lux and Budke (2020) highlights the potential of digital games as a tool to foster systems thinking in the context of complex socio-ecological challenges, such as city development, migration, climate change and/or resource usage. A framework was developed to assess the complexity of in-game systems, and a selection of commercial strategy and simulation games were analyzed to determine how system complexity is designed differently in each game. Based on the analysis, recommendations were made for selecting and using various games in both formal and informal learning contexts to promote networked thinking and systemic competence. The study analyzed 18 commercial games, identifying politics-centered games like Civilization VI, Democracy 3, and Fate of the World, and economy-centered city builders like Anno 2070 as the most complex games in the selection. These games require players to manage numerous parameters, topics, actors, and scale levels while considering complex chains of effects. The model can serve as a guideline for teachers to incorporate games into the classroom or build upon students' gaming experiences.

In conclusion, participatory simulations offer an engaging approach to complex systems learning by allowing students to actively engage with the subject matter, collaborate, and develop a deeper understanding of the systems they interact with. These activities are run at the classroom level, requiring many participants and a teacher to orchestrate the activity and facilitate debriefing and discussions, which may limit their applicability in smaller or less resource-rich setting.

In this thesis, we aim to incorporate swarm robots into novel learning activities designed to foster complex systems understanding. In our design process for these learning activities, we utilize the principles of agent-based modeling, which allow learners to apply rules that determine the actions of individual agents at a micro-level. Additionally, we draw inspiration from participatory simulations to create immersive scenarios where learners are not just observers but active contributors. This approach fosters an engaging and interactive learning experience, enhancing the comprehension of complex systems.

2.4 Learning Strategies

The creation of learning tools is driven by and rooted in one or more evidence-based learning strategies, with the aim of incorporating these methods into the design to enhance learning outcomes. In the following sections, we will elaborate on two primary strategies:

2.4.1 Game-Based Learning

The significance of play in cognitive development and learning has been recognized by psychologists for a long time. Piaget (1962) viewed play as an integral part of children's cognitive development, which evolved as they progressed through various developmental stages. Game-based learning involves using games to achieve specific learning outcomes, often in a digital format, while balancing subject matter with gameplay (Plass et al., 2015). This differs from gamification, which uses game elements to motivate learners to complete tasks they may not find engaging. Scholars debate on the definition of games, but generally agree that games involve a conflict with quantifiable outcomes. For example, gamification of math homework may involve giving points for completing activities, while game-based learning would imply redesigning activities to ground the mathematical tasks in gameplay through appropriately designed rules of play and conflict (Plass et al., 2015).

Video games play four key roles in the learning process (Steinkuehler & Squire, 2022). Firstly, as content providers, they impart specific subject knowledge in areas like history, math, language learning, physics, and medicine. Secondly, they serve as an enticing tool that leverages the engaging nature of video games to draw students into learning, even when it isn't immediately apparent. Thirdly, video games function as assessment instruments, using features such as "leveling up" to gauge a player's progress and growing expertise. Finally, video games act as engagement frameworks, analyzed by exploring the reasons behind gameplay and designing games that foster involvement, immersion, and commitment—often through the use of narrative structures (Steinkuehler & Squire, 2022). A broad range of studies has explored the impact of video games on learning in various school subjects. As an argument for why games are effective learning environments, the most frequently cited characteristic is their motivational function, achieved through incentive structures and interesting game mechanics (Connolly et al., 2012; Rotgans & Schmidt, 2011). Another important argument is that games allow for a wide range of ways to engage learners, including cognitive, affective, behavioral, and sociocultural engagement (Plass et al., 2015). However, the effectiveness of games varies based on factors such as gender and whether the content navigation is self-driven or guided by a teacher or computer (Vogel et al., 2006). Some studies have found that video games outperform traditional classroom instruction (Sitzmann, 2011), while others have reported mixed or contradictory results (D. B. Clark et al., 2016).

Two meta-analyses provide further insights into the contradictory conclusions found in the literature. The first, by Young et al. (2012), evaluated the connection between playing educational video games and academic performance, finding evidence for positive effects of video games on language learning, history, and physical education, but limited evidence for the academic value of video games in science and math. The second, by Wouters et al. (2013), analyzed 39 studies comparing games to more conventional instruction methods and found that games were more effective for acquiring knowledge and cognitive skills, especially when supplemented with other instructional methods and played in groups. However, they also found that games were not more motivating than conventional methods. More recent analyses, such as (D. B. Clark et al., 2016), have reported that digital games were associated with a 0.33 standard deviation improvement relative to non-game comparison conditions.

To summarize, research indicates that video games can improve engagement and learning in specific subjects, with their learning effectiveness subject to factors such as game mechanics, context, and individual play patterns. Although the existing literature presents mixed findings and limitations, it is clear that video games hold the potential to positively influence learners' motivation and engagement. Drawing on this motivation, we have designed the learning activities for this thesis around game-based interactions.

2.4.2 Tangibility in Learning

One of the most fitting arguments for tangible computing in learning sciences lies in the theory of embodied cognition. Embodied cognition theories suggest that our ability to comprehend the world stems from having a body that can interact with it (Dourish, 2001).

Engaging with physical materials can serve as an embodied basis for understanding concepts across various subjects. For example, paper represents a close approximation to a two-dimensional surface, a string demonstrates tension without pushing force, and billiard balls mimic the behavior of molecules in an ideal gas. In all these cases and many others, hands-on experience with physical materials and objects helps to develop intuition and offers a source of metaphorical insights and imagery (Pares & Eisenberg, 2022). Nonetheless, it is crucial to keep in mind that depending solely on physical experiences has its drawbacks. While drawing parallels between molecules and billiard balls can be beneficial in certain situations, such as in early statistical mechanics, it might prove to be unhelpful in other cases, including the study of hydrogen bonding or electrical currents (Pares & Eisenberg, 2022). It is important for the embodied learning to be accompanied by a phase of "reconciling" (Abrahamson et al., 2020) in which students reflect on their actions and integrate their embodied ("firsthand") knowledge with formal notions of the concept which use words and symbols ("secondhand" knowledge) (Schwartz et al., 2005). This is because both firsthand knowledge, which is based on direct experience, and secondhand knowledge, which is based on removed interpretations of descriptions of experience, are necessary for reasoning and problem solving (Schwartz et al., 2005).

As Tangible User Interfaces (TUI) gained popularity in HCI communities, researchers began investigating the added value of tangibility over pure Graphical User Interfaces (GUI) in educational settings. Marshall (2007) suggested four potential benefits of using TUI in learning: collaboration, accessibility, novelty of links (coupling physical with digital), and playful learning. Schneider et al. (2011) investigated the role of tangibility in a problem-solving task in pairs and concluded that the main impact was on promoting constructive behavior (exploration, collaboration, and playfulness of the task). Tangible programming with educational robots, such as Thymio[2], has also gained popularity and is widely used in primary and secondary schools to teach programming. Promising results have shown a higher degree of collaboration within groups (Mussati et al., 2019).

A meta-analaysis (Li et al., 2022) provides a comprehensive review of tangible learning studies presented at TEI (Tangible, Embedded and Embodied Interaction) conferences from 2007 to 2022, with a focus on collaborative learning, TUIs' impacts on learning, and comparative studies between TUI and other interfaces. On the topic of impact on learning, TUIs are found to support learning by providing scaffolding (for instance, facilitating concept comprehension and decreasing cognitive burden), modifying learning behavior (for example, enhancing attention, management, and expression), and enhancing emotional involvement in the learning process (such as increasing engagement, immersion, and enjoyment).

It is worth noting that the effectiveness of TUIs in comparison to other interfaces is not universally applicable as some studies have found no difference when using TUIs (Li et al., 2022). Moreover, some teachers have raised concerns about the technology's novelty, difficulty assessing learning outcomes, and the potential for TUIs to be shoehorned into situations where they may not be helpful (Li et al., 2022). However, when used judiciously and thoughtfully employed, physical materials would rather serve as "objects to think with," a term coined by Seymour Papert and echoed by other researchers in the field. In this thesis, we employ physical robots as part of a tangible learning strategy.

2.5 Conceptual Change

Certain subjects systematically prove difficult for students, and traditional teaching methods tend to be insufficient, as learning these topics necessitates a process of *conceptual change*. The notion of "conceptual change" highlights the primary challenge students encounter in constructing new ideas while considering their existing knowledge, underlining the importance of transformation rather than mere accumulation or "blank slate" acquisition of knowledge. Strong evidence suggests that pre-existing ideas can support or hinder learning in a wide range of areas. The term "conceptual" within this context should not be taken too literally, as various learning theories also emphasize beliefs, or ontologies, alongside concepts (diSessa, 2014).

[2] https://www.thymio.org/

Certain learning tasks necessitate a process of conceptual change, which involves not only the acquisition of new, accurate knowledge but also the revision and modification of pre-existing alternate conceptions. When learners encounter situations that induce cognitive dissonance, they are presented with a choice: they can either adapt their mental models to accommodate the new information (accommodation) or reject the conflicting evidence by considering it an exception, a measurement error, an anomaly, or by simply ignoring it (assimilation) (Piaget, 1962). Erroneous conceptions can hinder deep learning unless they are supplanted with accurate ones. These early, intuitive, and incorrect beliefs are sometimes referred to as naïve theories (diSessa, 2022).

The concept of force serves as an excellent example of conceptual change. In a simple event, such as tossing a ball into the air, after the ball has left the hand, only one force (gravity) acts on it. Gravity affects the ball's speed, reducing it until the object reaches zero speed at the peak of the toss. Gravity then continues to pull the speed of the ball downward, accelerating the ball in a downward direction. Prior to conceptual change research, instructors might have attributed student difficulties in understanding this phenomenon to the abstract nature of physics or its complexity. They might have tried re-describing the situation or simplifying the language. The "blank slate" cognitive model assumes a simple acquisition model of learning. However, by closely listening to student explanations, a surprising discovery emerged: students do not merely lack knowledge; they think differently. A common novice explanation involves two competing forces. Students believe that the hand imparts an upward force that overcomes gravity, driving the ball upward. This upward force gradually diminishes until it balances gravity at the peak, after which gravity takes over and pulls the ball downward. This two-force explanation reveals that students have a prior concept of force, different from what physics prescribes. Instruction must address and change these prior ideas, hence the notion of conceptual change (diSessa, 2022).

Conceptual change stands in contrast to more straightforward types of learning, such as acquiring skills (for instance, mastering a physical skill or an algorithm for long division) or learning facts (like arithmetic number facts). The difficulties in these areas are more evident, stemming from the extensive amount of material to learn and the need for consistent practice to attain proficiency. Conversely, conceptual change refers to a more profound form of learning that is not easily addressed by traditional teaching methods.

Conceptual change not only plays a central role in the learning sciences but also has implications for developmental psychology, epistemology, and the history and philosophy of science. Researchers continue to investigate the nature of knowledge, the challenges of acquiring it, and methods for fostering deep understanding. Conceptual change research is an evolving field that incorporates various perspectives and ideas from multiple academic disciplines, resulting in a diverse body of work that has yet to reach a consensus on well-established theories (diSessa, 2022). Two prominent broad theoretical perspectives regarding knowledge structure coherence exist namely, (1) knowledge-as-theory perspectives and (2) knowledge-as-elements perspectives (Özdemir & Clark, 2007).

Essentially, the knowledge-as-theory perspective posits that learners have coherent, organized mental structures or theories that guide their understanding of the world. When students encounter new information that conflicts with their existing theories, they experience cognitive dissonance, which prompts them to reorganize their mental structures and incorporate the new knowledge. According to this perspective, alternate conceptions arise due to learners' naive theories, which are often deeply rooted and resistant to change. To facilitate conceptual change, educators must help students recognize the limitations of their current theories and present them with a more scientifically accurate alternative that is intelligible, plausible, and fruitful. A typical example given by Prof. Dillenbourg is the "wool adds warmth" misconception. A trap to address it is an experiment comparing the melting time of an ice cube covered with wool vs one not covered with anything: contrary to the misconception, the second ice cube will melt first. Faced with a contradiction, learners revise their conceptions towards the scientifically accurate understanding.

According to Posner et al. (1982), four conditions must be met for successful conceptual change: (1) dissatisfaction with existing conceptions, (2) intelligibility of the new concept, (3) plausibility of the new concept, and (4) the fruitfulness of the new concept in solving problems and explaining phenomena. By addressing these conditions, educators can facilitate a deeper understanding of complex scientific ideas and promote meaningful learning experiences. This model highlights the importance of recognizing and addressing students' alternate conceptions in the learning process and underscores the dynamic nature of knowledge acquisition and development. The process of conceptual change involves identifying and challenging these alternative conceptions through targeted instruction and activities that encourage learners to confront and reconcile inconsistencies between their existing knowledge and new information (Vosniadou & Skopeliti, 2014).

In contrast, the knowledge-in-pieces perspective argues that learners' knowledge is not organized into coherent theories but rather consists of numerous, fragmented pieces of information called "knowledge elements"(DiSessa et al., 2013). These elements can be context-dependent and may sometimes contradict each other. Conceptual change, according to this perspective, occurs when learners reorganize and restructure these knowledge elements in response to new experiences or information. The knowledge-in-pieces perspective emphasizes the role of learners' active engagement in the process of knowledge construction and the importance of providing opportunities for students to apply and refine their understanding in various contexts. In this view, misconceptions are seen as the result of the activation of inappropriate knowledge elements in specific situations, and addressing these misconceptions involves helping students activate the correct knowledge elements and integrate them coherently (DiSessa et al., 2013).

Ultimately, both knowledge-as-theory and knowledge-as-elements perspectives agree on several aspects of learning. They concur that learners acquire knowledge from their daily experiences, which heavily influences their conceptual understanding, and that this naïve knowledge impacts their formal learning. Students come to science instruction with a variety of alternative conceptions that often conflict with scientifically normative ideas. Both perspectives acknowledge that much of naïve knowledge is resistant to change due to its entrenchment in everyday experiences (Özdemir & Clark, 2007).

As discussed in Chapter 1, alternate conceptions are frequently observed when it comes to understanding complex systems. Complex systems are inherently difficult to learn due to their dynamic, interconnected, and often non-linear nature, which challenges students' pre-existing mental models and intuitive reasoning. Consequently, it is crucial to adopt a pedagogical approach that effectively addresses these misconceptions and promotes a more accurate understanding of complex systems. This is where the conceptual change theory becomes particularly relevant and applicable. Therefore, conceptual change theory can guide instruction and activity design for complex systems understanding.

2.6 Conclusion

In this thesis, we aim to contribute to the fundamental societal challenge of complex systems understanding, and we envision to do so by leveraging the properties of the under-explored **swarm robots**. We propose a novel learning framework to engage students in **tangible** and interactive **game-based** learning experiences that addresses complex-systems related misconceptions, and thus stands on the theory of **conceptual change**. This approach complements existing tools. Drawing inspiration from the agent-based modeling approach, a learner would be able to set and change actions of agents at the micro-level. Inspired by participatory simulations, we develop engaging scenarios where learners are active participants to help transfer the link between the micro/agent level and the macro-level phenomena. By doing so, we leverage the strengths of existing tools, hope to support and facilitate the understanding of complex systems and enable learners to have an engaging and enjoyable experience.

Part I

Individual Learning with Swarms: Main Studies

Preface

Part I forms the core of this thesis. It focuses on the design, development and experimental evaluation of an individual learning activity that involves interaction with a swarm to understand complex systems, and related resources.

Research Questions

The research questions guiding this part are fourfold:

- How can we measure learners' understanding of complex systems?
- How can we foster a cognitive shift from a clockwork, centralized-deterministic mindset to a more decentralized, complex systems thinking mindset?
- Are there benefits to physical swarm interactions compared to virtual ones?
- Can we influence complex systems understanding across age groups?

Methodology

To address these research questions, we employ the following approach:

- Develop an assessment instrument to measure learners' understanding of complex systems.
- Iteratively design a game-based learning activity for complex systems understanding.
- Conduct empirical studies to evaluate the effectiveness of our design and test it across different age groups.

The first step involves developing an assessment instrument to measure learners' understanding of complex systems. This instrument will help us evaluate the effectiveness of our learning activity and quantify the changes in learners' understanding of complex systems after participating in the activity. We achieve this in Chapter 3.

The second step involves iteratively designing the learning activity for complex systems understanding. We aim to create a learning activity that is engaging, effective, and suitable for learners of different ages and backgrounds. We will refine the activity based on feedback from learners and instructors, as well as our own empirical observations. This is done is Chapter 4.

The third step involves conducting empirical studies to evaluate the effectiveness of our design and test it across different age groups. We will collect data on learners' performance on the assessment instrument, as well as their engagement and satisfaction with the learning activity. We will also analyze the data to determine whether the learning activity is effective in facilitating cognitive change from a centralized to a decentralized mindset, and whether there are benefits to physical swarm interactions compared to virtual ones. Chapter 5 reports on the first study in this direction. Chapter 6 presents a second study that investigates the efficacy of the activity with participants across different age groups. The study extends the previous one by including participants of different ages, from elementary school students to university students to older people from the general population.

In summary, this part of the thesis focuses on an individual learning activity that involves interaction with a swarm to understand complex systems. The research questions guiding this part are centered around measuring complex systems understanding, facilitating cognitive change, exploring the benefits of physical swarm interactions, and investigating the activity's efficacy across age groups. To address these questions, we employ the methods of developing an assessment instrument, iteratively designing the intervention, and conducting empirical studies. The four chapters in this part present the results of these studies and provide insights into the effectiveness of the designed activity in enhancing learners' understanding of complex systems.

3 Assessment Instrument: Development and Validation

Most of the material of this chapter is based on the paper: Khodr, H., Kothiyal, A., Bruno, B., & Dillenbourg, P. (2022). An assessment framework for complex systems understanding. In C. Chinn, E. Tan, C. Chan, & Y. Kali (Eds.), Proceedings of the 16th international conference on learning sciences - ICLS *(pp. 99–106)*

3.1 Introduction

Assessment instruments are crucial tools for evaluating learning outcomes and gauging the effectiveness of educational resources and tools. As complex systems play an increasingly important role in many areas, it is essential to have an assessment instrument that can measure learners' understanding of these systems. In this chapter, we will discuss the development of a general assessment instrument for assessing essential complex systems ontological concepts. We will describe the validation process used to ensure the instrument's reliability and validity. Furthermore, we outline a clear coding scheme, complemented by an automated methodology for question coding. This chapter will provide educators and researchers with a valuable resource for assessing learners' understanding of complex systems and evaluating the effectiveness of related educational interventions.

3.2 Theoretical Background

As discussed in Chapter 1, complex systems are present everywhere in nature, and many are commonly observed or experienced by all of us during our daily life. However, experts and novices think and build their knowledge about complexity in significantly different ways. Hmelo-Silver et al. (2007) found that novices focus on the visible and static structures of the subsystems involved, while experts incorporate structures, causal behaviors, and functions. Learning about complex systems is challenging (Hmelo-Silver & Azevedo, 2006), and many studies investigate the difficulty that humans have in understanding complex systems, alongside common misconceptions about them. Recapping from Chapter 1, Resnick (1997) discusses students' bias to generally assume that the behavior of a system is controlled by lead - a leader orchestrating the system's patterns, or by seed - a single pre-existing entity in the environment, rather than resulting from decentralized interactions. Moreover, probabilistic non-deterministic behavior in systems is also found to be a challenging concept to comprehend (Wilkerson-Jerde & Wilensky, 2015). Similarly, people tend to assume immediate effects instead of indirect cascading effects (Grotzer et al., 2015). Chi (2005) have also shown that a process is generally thought of as a direct static phenomenon rather than emergent.

An assessment instrument for complex systems understanding should therefore be designed to detect the presence of these misconceptions.

Yoon et al. (2019) conducted a study where they analyzed students' responses to one open-ended question in an ecology context concerning the effects of the arrival of geese on a park ecosystem. Results showed that the most difficult ideas to grasp were those related to the decentralized organization of the system and the unpredictable or non-deterministic nature of effects. They then proposed a learning progression to systematize the pathways students undertake to enhance their conceptual competence in complex systems. In another context, Wilensky and Abrahamson (2006) studied the student learning of the spread of diseases in a participatory simulation activity. Their findings identified several reasons for students' incorrect interpretation when explaining the agent-to-aggregate (i.e. micro to macro) relationship. These include proportional and linear reasoning, randomness-determinism confusions, disregarding feedback loop effects, and anticipating emergence from agents' rule-based interactions. Their assessment was composed of questions where they can select and draw graphs, and interviews where the situations, modeling, and representation tools relate thematically to the disease spread context. The students were provided with a modeling kit (stationery, lego, strings, etc.) and data-charting tools to give them a choice between either modeling the process itself and concluding their response based on this or directly charting quantitative data without thinking through the process. The key items in the questions were the spread of a mold on a moldy piece of bread, the scattering of people in a gym class, the spread of a rumor in a class, and the disease spread.

Hmelo-Silver et al. (2007) suggested a framework based on structure-behavior-function theory to represent how people think about complex systems, in particular, in a biological context. Structures refer to the elements, functions to the role of an element, behaviors to mechanisms of how the structures achieve their function. The case study included two biological complex systems: the human respiratory system and an aquarium ecosystem. To assess complex systems understanding, interviews were conducted, including drawing and open-ended questions. These were then coded by the presence and absence of the target concepts (structures, behaviors, functions). Jeong et al. (2022) used concept maps as the primary instrument for assessing students frameworks for understanding the watershed ecosystem. Concept maps serve as external depictions of mental models through drawing, which are composed of nodes (concepts) and links (relationships between them).

Jacobson (2001) developed another framework depicting multiple categories of complex systems concepts based on ontological and epistemological "component beliefs". Multiple studies then followed and adapted this framework (Jacobson et al., 2011; Yoon et al., 2019). Jacobson et al. (2011) then refined this framework to focus on five main concepts - control, causes, agent effect, action effect, and process - which we also draw upon in this work (more details later). They developed a hypermedia learning environment consisting of five NetLogo agent-based models: foraging of ants, traffic jams, self-organization of slime molds, social segregation, and wolf-sheep predation. To support learners' understanding of complex ontologies, they introduced interactive scaffolding and problem-solving exercises that encouraged learners to compare and contrast different aspects of complex systems. The observation of ontological shifts further confirmed and validated previous findings on the main concepts that make up complex systems.

In the literature discussed above, numerous studies have explored the importance of learning about complex systems, as well as the difficulties encountered in the process. They also aimed to define and validate the main concepts of complex systems. These studies typically relied on context-specific questions, interviews, or thematic analyses, which require expert judgment, or they provided only representative items of questions used, rather than the full set. However, there is no openly available quantitative instrument to objectively score a learner's understanding across different concepts of complex systems, as well as various contexts. Our work in this chapter is motivated by this limitation. Based on and inspired by the findings above, we compose a general assessment instrument and its associated scoring rubric that can be used to assess expertise within complex systems. The instrument requires participants to answer questions related to five different scenarios which are then scored to extract their competence in five ontological concepts required for complex systems understanding. More specifically, our research question is:
Can we quantify the expertise in complex systems based on answering questions targeting related ontological concepts?

3.3 Method

3.3.1 Materials

The instrument is designed as a questionnaire consisting of five scenarios, each related to a subset of the key concepts of complex systems understanding. These concepts, identified as core complex systems characteristics, used and validated in prior research (Jacobson, 2001; Jacobson et al., 2011; Yoon et al., 2019), are the five ontological categories of complex systems understanding — Control (Centralized, Decentralized), Causes (Single, Multiple), Actions effect (Linear, Nonlinear), Agents effect (Predictable, Random), and Processes (Static, Emergent). The scenarios draw inspiration from literature and are: Traffic Jam (Resnick, 1997), Scatter (Wilensky & Abrahamson, 2006), Flock of Birds (Jacobson, 2001), Butterfly Effect (Jacobson, 2001), and Robots and Gold (Resnick, 1997). For each scenario, the questions were formulated as a mix of Multiple Choice Questions (MCQs) and open-ended questions. They were written in an appropriate way to be answered by both experts and novices alike. No formal declarative concepts were used. The detailed questions can be found in Appendix A. Whereas the last two scenarios (butterfly effect and robots and gold) are thought experiments, the first three scenarios include familiar and man-made (scattering, traffic jam) or naturally observed behaviors (flocks of birds). Moreover, we distinguish two types of questions: questions requiring explanation of the behavior of a system (e.g. How is it that the birds fly in a flock? in the Flock of Birds scenario) vs. those asking to define rules to enact the behavior of a system (e.g. How should the robots move in order to find the gold? in the Robots and Gold scenario). Finally, the questions, even in the same scenario, can focus on different perspectives: local (agent perspective, e.g. What does each person need to do in order to scatter? in the Scatter scenario) and global (external observers, e.g. If you were up really high, and could watch your class scatter, what would you see from above when they're scattering? in the Scatter scenario).

3.3.2 Procedure

An iterative approach was followed to refine the questions and validate the instrument. A first version of the questionnaire was sent out by email to 11 participants of varying ages and educational backgrounds. The goal of this first validation was to evaluate the understandability of the scenarios and questions, and to obtain a range of the possible answers from our target population in order to develop a coding scheme. Based on this first feedback, the questions were refined, and confusing statements were clarified. The first draft of the coding scheme was also developed.

Our goal in the second iteration was to validate that the chosen questions assessed the set of complex systems concepts, further validate the understandability of the questions and refine the coding scheme. The second validation phase included two steps. First, we discussed and validated the questions and coding scheme with an expert in complex systems education to ensure that the five scenarios and accompanying questions were representative of all the complex systems concepts. Second, we followed a verbal cognitive methodology (Crutcher, 1994) in which three novices (undergraduate students) and three experts (graduates/post graduates in swarm robotics) were asked to read the questions and describe verbally what they understood from the questions and how they would answer them. Then they wrote the answers in the questionnaire. Based on this second validation, final refinements were done to the questions and coding scheme.

3.3.3 Participants

The final version of the questionnaire was administered to two groups of participants:

- An **expert** group including eight subjects who responded to a personal request for participation sent by email. They were purposefully sampled to include PhD students

Concept	Focus	Clockwork Mental Model	Complex Systems Mental Model
Control	The control is centralized or decentralized	Response indicates that the system is controlled by a leader (order is top-down)	Response indicates that the system is decentralized (order is bottom-up)
Causes	The number of causes that contribute to the outcome	Response attributes the outcome to one primary cause/factor	Response attributes the outcome to multiple causes/factors
Actions	Interaction between agents are linear/proportional or not	Small action → small effect; Large action → large effect	Small action → large effects; Effects of actions may not be repeatable
Agents	Predictability of actions of agents	Actions of agents are predetermined	Actions may not predictable, random
Process	How the system works	System is an event defined by its global behavior	System is a dynamic process, self-organizes through agent and environment interactions

Table 3.1: General coding scheme: concepts, focus, and the difference between clockwork and complex responses.

or PhD holders in a related domain (mainly in swarm robotics, physics, and computer science). They self-rated their expertise in complex systems with an average of 5/7 and confidence in answers with an average of 5.625/7. The mean age of the group was 33.88 years (SD = 3.59), with six males and two females. Classified by ISCED[1] Majors, 3 are in Engineering and 5 are in Science.

- A **non-expert** group including 29 subjects. They were recruited by convenience sampling with requests diffused online and among students in university and high school. The mean age of the group was 24.21 years old (SD = 6.31) with 16 females (55%) and 13 males (45%). These included 4 students from high school, 16 in Bachelor, 7 in Masters, and 2 PhD students. Classified by ISCED Majors, 15 are in Science, 3 in Health, 4 in Engineering, 1 in Education, 2 in Social Science, 4 in General Knowledge (high school students).

[1] http://uis.unesco.org/en/topic/international-standard-classification-education-isced

3.3.4 Scoring and coding of responses

The responses were scored based on the five key concepts mentioned previously. We consider 3 levels of understanding: a clockwork (corresponding to a centralized mindset) mental model, a complex systems mental model, and an intermediate or in-between level. Table 3.1 provides the concepts, focus, and the difference between clockwork and complex responses (Jacobson et al., 2011; Yoon et al., 2019). Based on these concepts a detailed coding manual was developed for each of the questions (available in Appendix A).

The inter-rater reliability of the coding scheme was assessed with three raters all coding 25% of the responses. The inter-rater agreement was measured using Fleiss's Kappa and was found to be 0.66 which denotes a substantial agreement. The discrepancies were discussed, and related details were added to the coding scheme to make it clearer. The remainder of the responses were subsequently split among the three raters to code. We note that another inter-rater reliability assessment was conducted a year later, with the three raters coding five new responses from a new set of data from our study described later in Chapter 5. This last reliability assessment was done to ensure that the established coding scheme remained valid even after some time had passed. The agreement was found to be 0.81 which denotes a strong agreement.

Below we describe the traffic jam scenario as an example. The instruction is:

On the road, each car followed only three rules:

1. If there is a car close ahead of you, slow down.
2. If there aren't any cars close ahead of you, speed up (unless you are already moving at the speed limit)
3. Comply with driving regulations concerning all other elements on the road, such as: slow down if you detect a radar trap, traffic lights, accidents, entry ramps… (we attach a gif)

Q.1: What is causing this traffic jam? Select all applicable choices (radar trap, broken bridge, entry ramp with merging traffic, accident, nothing, other)

Q.2: Explain in your own words how your choice(s) is(are) causing the traffic jam.

Q.3: The radar trap is removed, and there are no accidents on the road, no broken bridge, no entry ramp or any other external events. Can a traffic jam still form? (Yes, No)

Q.4: If yes, why do you think a traffic jam can still form? If No, explain in your own words, how the cars will behave on the road.

We target three main complexity concepts with this scenario: causes, control, and action effects. Examples of responses and scores are given in Table 3.2.

ID	Q.1	Q.2	Q.3	Q.4	Scores
P1	Radar trap, Broken Bridge, Entry Ramp with merging traffic, Accident, Nothing	Slow down compounds resulting in a complete halt. Netlogo has this model and explains it well too.	Yes	As one car's behavior is dependent on others, the speeds and behaviors become nondeterministic. Even a slight slowdown to avoid bumping into the car in front can compound as the reaction time may have a slight delay as well.	Control(1) Causes(1) Action(1)
P2	Radar trap	Cars are driving slower after they stopped suddenly	No	They will behave freely with no stops or jams	Control(0) Causes(0) Action(0)
P3	Broken Bridge, Accident	They basically cause a blockage to traffic, so that vehicles 'pile up'	Yes	It also depends on the amount of cars on the street	Control(1) Causes(0.5) Action(0)

Table 3.2: Scores of 3 participants for the first scenario: Traffic Jam.

3.3.5 Results

For each participant, we summed up all the scores per concept, as well as per scenario. Each raw score was divided by the total possible score, to have standardized scores for the different concepts and scenarios. The total score was calculated as the average over all the concepts. A significant difference was found between the scores of experts and non-experts over all concepts, except for the concept "Causes" where the general population also scored high (see Figure 3.1). Considering scenarios, significant differences were found between the experts and general population participants in all scenarios except for the second (Scatter) scenario (see Figure 3.2). In the Robots and Gold scenario, the difference was larger with a huge effect size (Cohen's d = 2.2). These findings support the claim that our assessment is capable of distinguishing between experts and novices and effectively measures what it intends to measure.

A strong significant correlation exists between total score and each of the respective concept scores (Spearman correlation with control = 0.75, action effects = 0.72, agent effects = 0.68, process = 0.66) and a moderate correlation with the causes concept (0.46). This lets us verify that no specific concept had low correlations with the overall scale. We also check the correlation between concepts. A moderate significant correlation exists between the control concept and each of action effects (0.44) and agent effects (0.62) concepts and between action effects

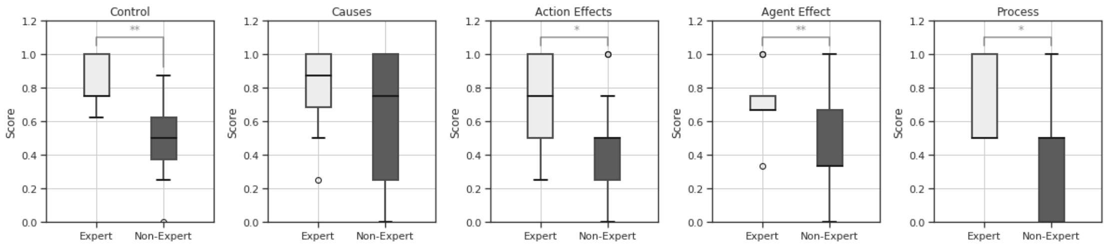


Figure 3.1: Comparison of experts and non-experts over the concepts. * (**) indicates a significant difference with p-value<0.05 (<0.005) with a Mann-Whitney Test.

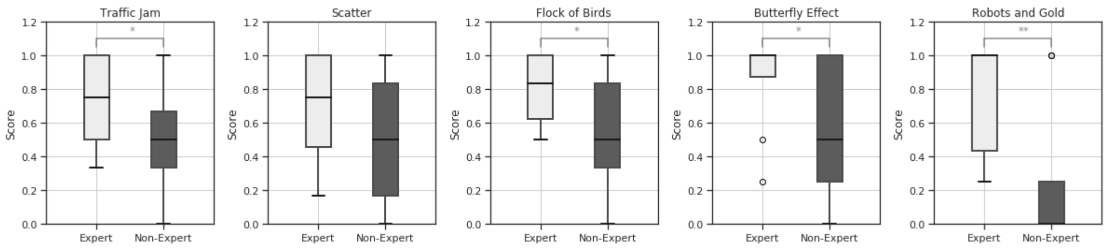


Figure 3.2: Differences per scenario among the experts versus the non-experts.

and process (0.48) concepts. A weak significant correlation (0.34) also exists between the concept action effects and agent effects. Furthermore, we investigate the correlation between scenarios. We only found weak significant correlation between scenarios 1(Traffic Jam) and 3(Flock of Birds) (0.35), as well as scenarios 2(Scatter) and 4(Butterfly Effect) (0.37).

While experts are expected to be knowledgeable on the concepts (and this assumption is supported by their high scores), participants belonging to the non-expert group can have varying degrees of understanding of complex systems, lying at different points on the continuum from novices to experts. Therefore, to verify whether the population can be divided into two groups, namely the novices and the experts, we conduct a clustering on the whole dataset involving the 37 participants on the basis of 10 features taken into account: the scores of each concept, as well as the scores of each scenario. A Principal Component Analysis (PCA) is conducted to extract the 6 most significant principal components, explaining 95% of the variance. The purpose is to make groups which are not necessarily easily separable in the space by the original features, but better separable in the new space spanned by the Principal Components (PCs). To verify the separability of our population, we perform a C-means clustering for the association of each participant to a cluster. The number of clusters is fixed to 2 to verify whether the expert participants indeed cluster together when all features are considered together. The results, shown in Figure 3.3, reveal that all the experts belong to the same cluster, as well as some of the participants from the non-experts. An extended discussion of this finding is reported in the Discussion section.

So far, we have analyzed our participants' responses either over single concepts/scenarios, or over the dimensions returned by PCA, which are a combination of these scores. What would

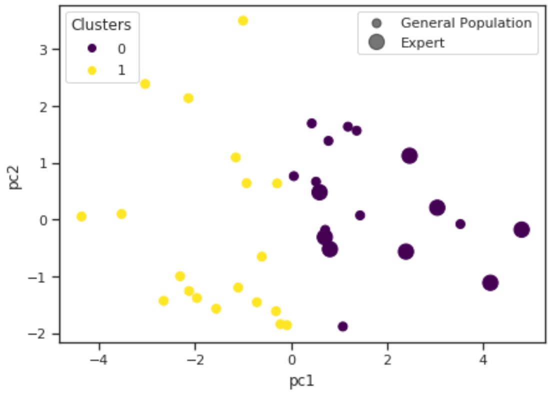


Figure 3.3: Clustering results

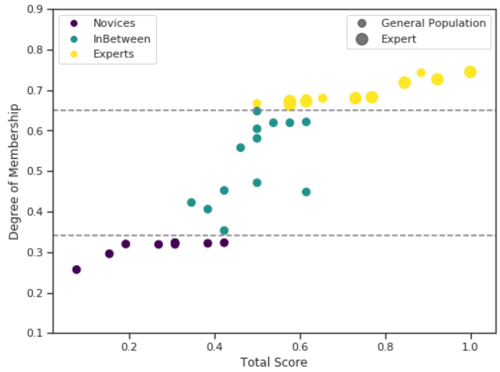


Figure 3.4: Fuzzy system output vs total score and division into 3 classes

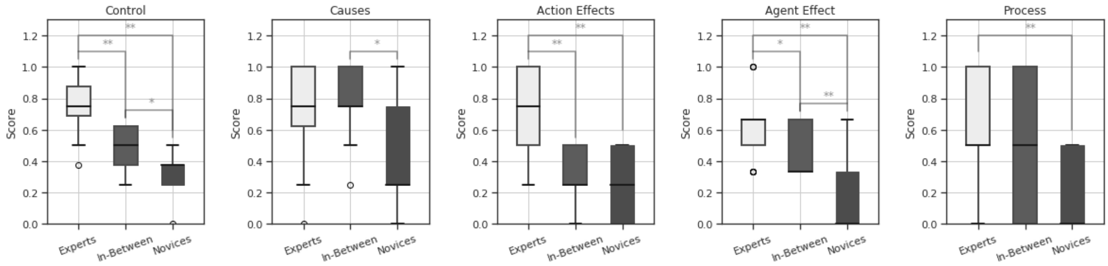


Figure 3.5: Differences per concept among the three groups.

be interesting is to have a global complex systems expertise score, which makes it immediate to evaluate a person's complex systems understanding. More concretely, our goal is to define a mapping between the concepts/scenarios and a global score of expertise. We investigate two approaches: the first is the traditional approach to sum and average over all the concepts. The second uses techniques from fuzzy logic systems. Fuzzy logic allows us to encode qualitative relations among multiple variables (fuzzy rules) in a unifying formalism (Driankov, 2001). In our case, the relation to describe is the one between a given vector of values for the complex systems' scores (concepts and scenarios), and the corresponding global level of expertise. The fuzzy rules are encoded as if-then rules underlying the qualitative fact that an expert level should have a high level in all the input variables; conversely, a novice level should have a

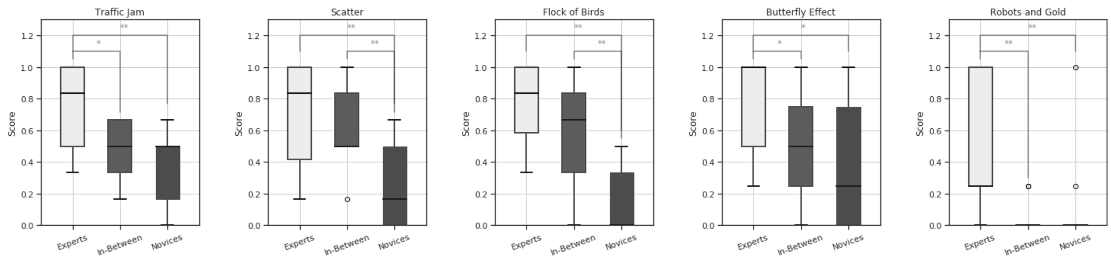


Figure 3.6: Differences per scenario among the three groups.

low level in all the input variables. Each input membership function is modeled as a logistic sigmoidal function. The parameters of a sigmoid are fitted after applying a 1D clustering on each of the dimensions. The output membership function is modeled as a trapezoidal function with cut-off limits set between 0.35 and 0.65. Figure 3.4 shows the relation between the total score and the output fuzzy logic score. We notice that the relation is sigmoidal rather than linear. Again, we verify that the group that we already know to be experts actually score high on the degree of membership of the expert level. We can also identify a separation of three classes: experts scoring higher than 0.65, novices scoring lower than 0.35 and an average in-between group. We recomputed the differences per concepts and scenarios between the populations of the three identified classes. Results are shown in Figures 3.5 and 3.6.

3.4 Discussions and Implications

In our first analysis, we found a significant difference between the scores of experts and the general population over different concepts and scenarios, which is in line with previous literature Jacobson (2001). One exception is the concept of "Causes" where the general population also scored high (see Figure 3.1). While it may be possible that this concept is easier to grasp than others, an alternative explanation is that in our instrument the questions measuring causes were mainly MCQs, which possibly led respondents to select more than one option. Conversely, in the last scenario (robots and gold), the difference in scores was much larger: a possible explanation is that in this scenario, participants were asked to enact/design rules in a novel scenario, rather than explaining an observed phenomenon. Given the ubiquitousness of centrally controlled systems in the world around us and novices' inclination towards the centralized mindset (Resnick, 1997), it is likely that novices believe in the superiority of centralized over decentralized systems and so chose to enact centralized rules.

In our total score analysis, the results indicate a separation between 3 groups of expertise. The already known experts all belong to the expert cluster. Few participants from the general population also belonged to the expert group. By examining their major, we found that they were a group of students doing their masters in biology, so they perhaps have a background knowledge in biological complex systems, which may explain their performance. The other two groups included participants of mixed educational backgrounds.

To better understand the differences between the three groups, we look again at the differences between the concepts. Significant differences were found for the concepts: control, causes, and agent effects between the novices and in-between groups; and for the concepts: control, action effects and agent effects between the expert and in-between groups. For the scenarios, significant differences were found in Scatter and Flock of Birds scenarios between the novices and the in-between groups, and Traffic Jam, Butterfly Effect and Robots and Gold between the expert and the in-between groups. Looking back in literature, we observe similar trends in difficulties related to complex ideas. Yoon et al. (2019) reported that order and deterministic effects (which link to control and agent and action effects in our definitions) constitute the hardest concept to grasp. Similarly, Resnick (1997) suggested that the commitment to centralized mindset constitutes a strong misconception for learners. Lastly, Grotzer et al. (2015) focused on the difficulty of understanding non-linear effects in a system. Finally, although through our questionnaire we only target the implicit knowledge of complex systems, the use of declarative knowledge (emphasis in italics) was observed in the experts' responses (Jacobson et al., 2011). An example is the explanation of the butterfly effect given by an expert:

> In a system with *nonlinear dynamics* such as with *positive feedback* loops, small changes in certain periods or locations can have large effects. Collective decision-making for a new nest in ant groups is one example, where the number of insects recruited to go to a specific new nest depends on how many are recruiting - the more that are already going, the more that are recruited.

On the other hand, an answer to the same question, given by a participant from the general population and which also displays implicit knowledge but lacks declarative knowledge is:

> I think those effects are the norm rather than an exception. For example, teachers influence the future life of students in small ways every day, and therefore also influence the lives of those they interact with.

3.5 Automatic Grading

During the course of conducting research and executing experiments, we collected a substantial volume of responses to the developed questionnaire. These responses were subsequently evaluated among three raters by following the coding scheme discussed above, culminating in the development of a dataset comprising 264 graded responses. To expedite the grading process in the future, we approached this task as a classification problem using the existing data, in which each concept within a given scenario is treated as a distinct classification output, while the responses to the questions within that scenario serve as input vectors. Consequently, we identified 13 classification tasks corresponding to all the concepts within their respective scenarios (ref. Table 3.3).

To address these classification tasks, we evaluated the performance of three distinct algorithmic approaches:

- Random Forest Classifier: This machine learning technique constructs multiple decision trees and aggregates their predictions for a more accurate and robust classification result (Breiman, 2001). As a preprocessing step for handling text data, one-hot encoding is employed to convert words into a binary format. This sparse representation allows the Random Forest Classifier to efficiently handle the dataset's textual nature and improves the model's interpretability and performance.

Task Number	Concept	Scenario	Best Performing Model
1	Control	Traffic Jam	Random Forest Classifier
2	Causes	Traffic Jam	Random Forest Classifier
3	Action	Traffic Jam	Fine-tuned BERT
4	Control	Scatter	Fine-tuned BERT
5	Agent	Scatter	Fine-tuned BERT
6	Process	Scatter	Fine-tuned BERT
7	Control	Flock	BERT-Tokenized Random Forest Classifier
8	Causes	Flock	Fine-tuned BERT
9	Agent	Flock	BERT-Tokenized Random Forest Classifier
10	Action	Butterfly	Fine-tuned BERT
11	Process	Butterfly	Fine-tuned BERT
12	Control	Robots and Gold	RandomForestClassifier
13	Agent	Robots and Gold	BERT-Tokenized Random Forest Classifier

Table 3.3: Classification tasks, their corresponding concept and scenario, and the corresponding best performing classification model

- Fine-tuned BERT model: BERT (or Bidirectional Encoder Representations from Transformers) is a state-of-the-art language model demonstrating exceptional performance in various natural language processing tasks, including text classification (Devlin et al., 2018). By augmenting a pre-trained BERT model, we capitalize on its advanced language understanding capabilities and contextual representation. This approach is highly beneficial when analyzing complex text data, such as questionnaire responses, potentially leading to improved classification performance.
- BERT-Tokenized Random Forest Classifier: This model represents the fusion of BERT's tokenizer for preprocessing and the RandomForestClassifier for classification. This approach harnesses the power of transfer learning by utilizing a pretrained model to improve the performance and generalization capabilities of the RandomForestClassifier, especially when dealing with limited labeled data or aiming to reduce training time and computational resources.

These three algorithmic choices were selected based on their ability to effectively handle categorical data, leverage transfer learning, and incorporate advanced language understanding capabilities. Each method offers unique advantages, and by comparing their performance, we can determine the most suitable approach for addressing the specific classification tasks within our study.

For each of the three algorithms, a consistent methodology was employed to ensure a fair comparison of their performance. This methodology involved the following steps:

1. Data Splitting: The dataset was divided into a training set and a testing set. Of the 264 responses, 40 responses (corresponding to 15% of the total) were reserved as unseen

test data, while the remaining responses were used for model training.

2. Cross-Validation: To further enhance the model evaluation process and minimize the risk of overfitting, a 5-fold cross-validation technique was applied to the training dataset. This approach partitions the training data into five equal-sized subsets, utilizing four of these subsets for training and the remaining subset for validation. This process is iterated five times, with each subset serving as the validation set exactly once, enabling the computation of an average performance metric across all folds.

3. Hyperparameter Tuning:
   a. For the Random Forest classifiers, a grid search was conducted to identify the optimal combination of the hyperparameters: maximum depth and number of estimators.
   b. For the BERT model, we used the following set of hyperparameters: The batch size was set to 16, which is a common choice to balance memory usage and model performance. The number of training epochs was set to 10, which provided a reasonable trade-off between training time and the potential risk of overfitting. A warmup percentage of 0.1 was used to gradually increase the learning rate during the initial phase of training, thereby avoiding large gradients and ensuring a more stable optimization process. A weight decay of 0.01 was applied to prevent the model from becoming overly complex and to improve generalization.

4. Model Evaluation: After training each algorithm with the optimal hyperparameters and completing the cross-validation process, the models were assessed on the unseen test dataset. This step allows for an unbiased comparison of the performance of each algorithm, providing valuable insights into their suitability.

We considered three metrics for evaluation: accuracy, F1 score, and Cohen's Kappa. These evaluation metrics provide a comprehensive assessment of the classification models:

- Accuracy: measures the proportion of correctly classified instances out of the total instances in the dataset. It is a widely used metric for classification problems.
- F1 Score: is the harmonic mean of precision and recall. It provides a balanced measure of a model's performance by considering both false positives and false negatives.
- Cohen's Kappa: measures the agreement between two raters, taking into account the possibility of agreement occurring by chance (Landis & Koch, 1977). In this context, it evaluates the agreement between the model's predictions (AI rater) and the true labels (human rater).

Results for the metrics are illustrated in Figure 3.7. The best model for each task is presented in Table 3.3. Overall, the Random Forest Classifier performs best or equally well in tasks 1, 2, and 12. In these tasks, the coding could be mostly done based on the answers to the multiple-choice questions, so one-hot encoding combined with random forest classification was sufficient.

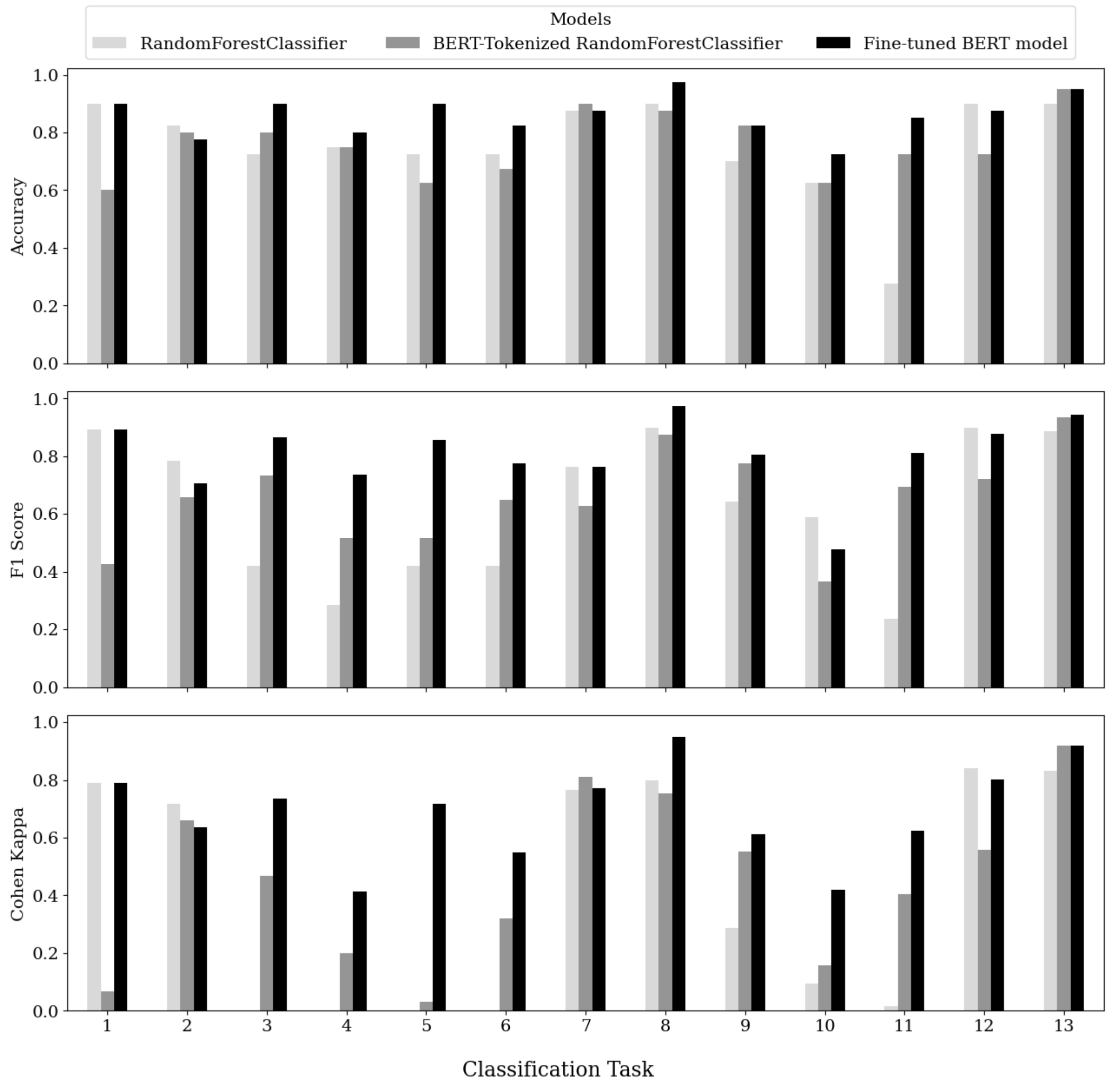


Figure 3.7: Comparison of the classification results between the three considered algorithms. The three metrics are accuracy, F1 score and Cohen's kappa computed on the unseen test dataset.

Additionally, this approach is computationally efficient and fast. The BERT-Tokenized Random Forest Classifier outperforms or achieves similar results to the other models in tasks 7, 9, and 13. By leveraging the power of transfer learning through the use of a pretrained BERT tokenizer, it enhances the performance and generalization capabilities of the Random Forest Classifier, while needing much less compuational power compareed to a full fine-tuned BERT model. Finally, the Fine-tuned BERT model excels in tasks 3, 4, 5, 6, 8, 10, and 11. In these tasks, coding the answers required the ability to capture the intricate relationships between words and their meanings within the given context. This capability is provided by the advanced language understanding and contextual representation offered by the pre-trained BERT model.

Finally, considering the best model for each task and combining the results over all tasks, we achieve an overall accuracy of 0.87, an F1-score of 0.81, and a Cohen Kappa of 0.77. These results indicate that the chosen models provide a relatively high degree of agreement between the human raters and the automatic classification, as well as a good balance between precision and recall. In particular, the substantial Cohen Kappa score suggests that the agreement between the models and the raters is significantly better than what would be expected by chance alone. This demonstrates that the selected models effectively address the classification tasks within this study and can be considered reliable tools for the automatic coding of the questionnaire responses in future studies. The use of these models offers several benefits, including significant time savings and improved efficiency for educators, researchers, and evaluators.

3.6 Conclusion

In this chapter, we proposed an instrument for assessing a person's expertise in complex systems, which requires solving five different scenarios and relies on a coding scheme to grade a person's competence in five ontological concepts essential for complex systems understanding. A study with 37 participants including experts and non-experts was conducted to validate our instrument after it had undergone two iterations of iterative refinement. Significant differences were found in the scores of the two groups of participants. Finally, an automatic approach for grading was developed. The best combination of models over the concepts per scenario achieved an overall accuracy of 0.87, an F1-score of 0.81, and a Cohen Kappa of 0.77.

In conclusion, the instrument, coding scheme and automatic scoring approach proposed in this paper advance our knowledge of assessing people's understanding of complex systems. Moreover, our instrument contributes to the design of learning activities to engage students about complex systems and helps investigate the effectiveness of such activities. In subsequent chapters, this assessment instrument will serve as a pre- and post- test to investigate the effectiveness of the designed activities.

4 The design of *Cellulan World*

Some of the material of this chapter is based on the paper: Khodr, H., Bruno, B., Kothiyal, A., & Dillenbourg, P. (2022). Cellulan world: interactive platform to learn swarm behaviors. Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, 1908–1910*

4.1 Introduction

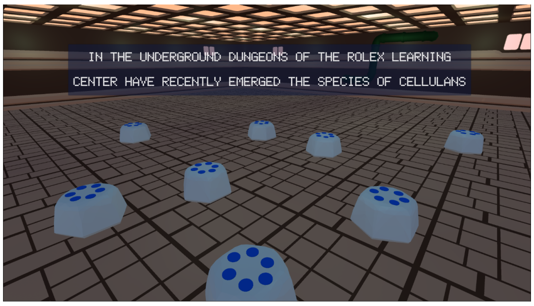
In this chapter, we present *Cellulan World*, a novel design for a learning activity in the context of complex systems understanding. *Cellulan world* is an individual educational game featuring Cellulans, a new species embodied by the Cellulo robots (introduced in Chapter 2), as the main protagonists. The learning goal of the activity is to foster complex systems understanding, focusing on the main ontological concepts of control, causes, process, action and agent effects discussed earlier in Chapter 3.

4.2 Cellulan World - The Story Line

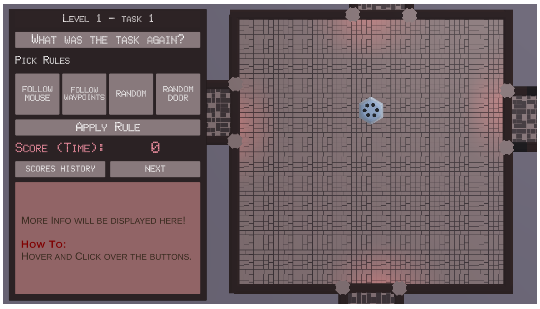
The activity starts with an animation[1] to introduce and engage the user in the story line .

Introduction (Figure 4.1a)
Hello [Name]! Welcome to the EPFL Rolex Learning Center. In the underground dungeons of the Rolex Learning Center have recently emerged the species of Cellulans. They are friendly creatures who buzz around on three wheels, with very few needs. They live in an underground maze, trying to find ways of acting that help them survive and grow. They have to find a way to work together and address all their challenges. Let's explore their daily life, and help them navigate it.

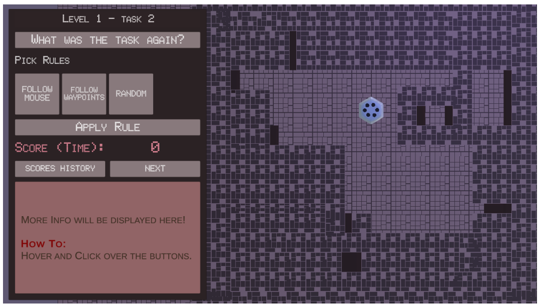
[1] https://youtu.be/k7OJPDWnv58



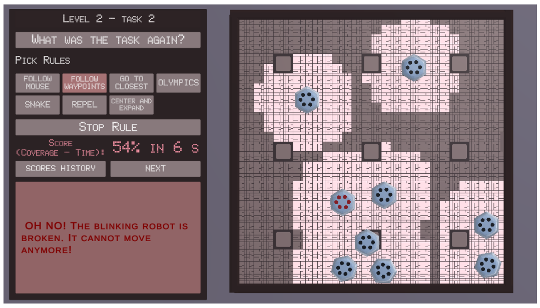
(a) Introduction



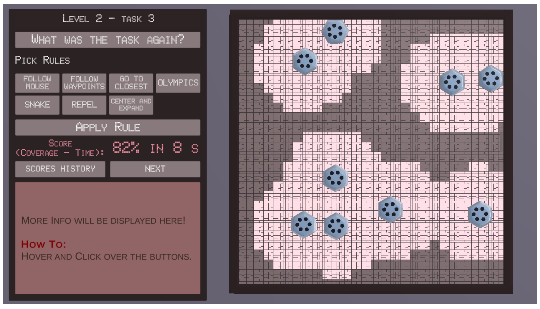
(b) Task 1 - Level 1



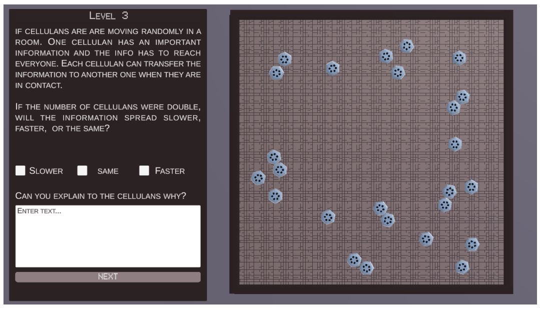
(c) Task 1 - Level 2



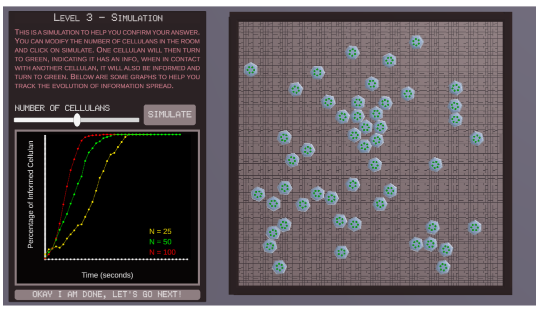
(d) Task 2 - Level 2



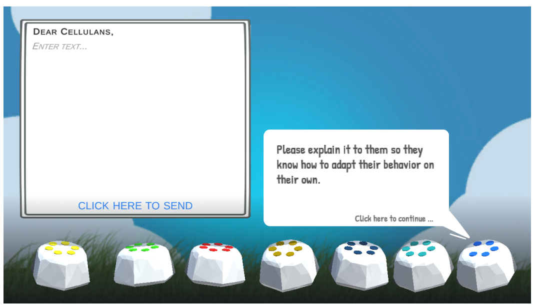
(e) Task 2 - Level 3



(f) Task 3 - Predict/Confirm



(g) Task 3 - Simulate



(h) Task 4

Figure 4.1: Screenshots from the game

The game includes a total of 4 main tasks:
Task 1 - Level 1 (Figure 4.1b)
Oh no! It seems like one Cellulan is stuck alone in a room while fellow Cellulans are behind one of the doors. Help the lonely Cellulan find the other Cellulans in the fastest time possible.

In this scene, only one robot is present. This level primarily serves as an introduction, enabling users to familiarize themselves with the system, the user interface, and the in-game interactions. A selection of rules is made available to users. The goal is to identify and utilize the most effective rule to accomplish a given task. Users can apply a rule, observe the resulting behavior, and assess the score it yields. When they feel ready to advance, they can proceed to the next level after determining what they believe to be the most effective rule. The available rules are detailed in Table 4.1 and further elaborated upon in Section 4.4. The game mechanics outlined here are consistently applied throughout all levels of tasks 1 and 2.

Task 1 - Level 2 (Figure 4.1c)
Things just keep getting harder for the Cellulans! It seems like another Cellulan is stuck alone in a room. To make things harder, your global view is disrupted now. Fellow Cellulans are still behind one of the doors; however, the location of the doors is now unknown. Re-explore the search rules: does your previous best strategy still work? If not, which one is the best now?

Task 2 - Level 1
The Cellulans have reached this big room. To explore what's in there, one approach is to spread out as much as possible so they don't miss anything. Explore which rules they can use to achieve this. The more surface they cover in a short period of time the better, since they have limited energy. If the coverage remains constant (i.e, does not change) for 7 seconds, the task is marked as done!

In this setting, multiple robots are introduced. The task requires a coverage behavior (i.e, to cover the given area of interest). There are predefined spots placed equidistantly in the room in a way to attain a high coverage percentage. The available rules are described in Table 4.1 and are discussed in more details in Section 4.4.

Task 2 - Level 2 (Figure 4.1d)
With all these activities, the Cellulans are getting hungry! Food is scarce and here they are in the room searching for food again. They again need to maximize their spread across the room. However, not all Cellulans are feeling well, their energy is low because they haven't been getting enough food. Few of them might stop at times to rest. Re-explore the rules: is the previous best strategy still the best?

This level is designed such that a randomly chosen robot will always break after a few seconds from the start of a rule execution. To notify the user of the event, the LEDs of the robot blink, a sound is displayed and a pop-up message is displayed.

Task 2 - Level 3 (Figure 4.1e)

The Cellulans can't seem to catch a break! They ate, slept for a while and now are ready to harvest more food to continue their journey. But surprise! The markings on the floor are gone this time! Re-explore the rules: is the previous best strategy still the best?

Task 3 - Predict (Figure 4.1f)
Thank you,[Name], for helping the Cellulans! They are now able to efficiently explore their environment. In their exploration journey, while the Cellulans were moving randomly around the place, one made an important discovery. The info has to reach everyone! If the number of Cellulans were doubled, would the information spread faster, slower or the same? Let them know by answering below.

Task 3 - Simulate (Figure 4.1g)
This answer is very important for the future of Cellulans. You have to be sure. To verify your answer, play with the next simulation.

In this level, the user can modify the number of Cellulans in the room. When clicking on the simulate button, one Cellulan will turn green indicating it has an information. Upon getting in contact with another Cellulan, the latter will also be informed and turn green. Graphs are plotted to track the evolution of information spread in each case.

Task 3 - Confirm (Figure 4.1f)
After you played the simulation, you can now confirm or change your answer!

Task 4 (Figure 4.1h)
Thank you, [Name], for helping the Cellulans in their discovery journey. You can now set them free and let them live on their own. Before you go, the Cellulans are asking for one last piece of advice on how to adapt their behavior on their own. Think back at all the tasks you've solved with them. What was your reasoning when you chose the best rule to be applied? Please explain it to them so they know how to adapt their behavior on their own.

This serves as a closing task. From a pedagogical perspective, it also serves as an overall reflection task for the users and gives us insights into their ways of thinking.

4.3 Pedagogical Design Strategies

Our activity is designed as a game that immerses the user in solving a variety of tasks and missions involving the Cellulans. By utilizing game features, we aim to leverage the benefits of **game-based learning** to increase motivational and cognitive engagement, thereby facilitating effective learning (as discussed in Chapter 2). In addition to this, we also design our learning activity to allow for the seamless integration of tangible robots, thus making it possible to leverage and investigate the added value of **tangibility** (as discussed in Chapter 2) in the context of our learning goals.

Borrowing from the concept of **agent-based modeling**, our game design allows the user to tackle tasks by initiating rules and observing the subsequent behavior and the scores achieved based on the specific task at hand. Further, taking inspiration from **participatory simulations**, our game encourages active participation of the learner as an agent. Specifically, similar to participatory simulations, the user can actively participate in the game by selecting the rules and further having direct control over the individual Cellulans in certain rules (as referenced in Table 4.1).

Within *Cellulan World*, the activity is organized into levels, with each level presenting the user with a number of challenges that require them to complete a specific mission. The game leverages the learning theory of **conceptual change**, helping learners to revise and restructure the concepts that build their knowledge by "trapping" and correcting alternate conceptions they may have through a series of traps built around the robustness of decentralized control when a dynamic change occurs. These traps belong to one of the following three categories:

1. Perspective change: global view vs local view; i.e, restricting or granting access to the global environment.
2. Environment change: a static vs a dynamic environment; or a well-known versus unknown environment.
3. Capabilities/Functionalities change: expanding or limiting the capabilities of an agent (e.g. field of view, communication range, etc) or even breaking it (complete failure).

An example of the first category, perspective change, is illustrated in Task1-Level2. In this level, the global view is restricted, causing the user to lose sight of the door placements, which are now randomly changing. As a result, the centralized rule of the Cellulan controlled by the user performs worse than the local rule (random).

The second category of traps, environment change, is exemplified in Task2-Level3. In this level, the predefined spots of the environment are removed, causing all the rules related to them to break and making it more difficult for the user to determine the positions of the robots, which negatively affects the final score. The best rule to use in this scenario is the Repel rule, pushing players to adapt to a more dynamic and unknown environment.

Lastly, Task2-Level2 presents an example of the third category of traps, capabilities/functionalities change. In this level, a random Cellulan can break at any point, causing the rules requiring global communication (such as Center and Expand) or any of the leader-follower rules (Snake, Olympics), to break. The distributed rule of Repel emerges as the most robust solution against this failure, as the Cellulans can still repel away from a non-moving agent.

By exposing the limitations of centralized strategies through these different traps, the game encourages learners to undergo conceptual change and explore more effective decentralized solutions. Through this process, they can revise their mental models and better understand the advantages of decentralized control in various challenging situations.

Another trapping strategy is using the **Prediction–Observation–Explanation (POE)** model. In this model, students are first asked to make predictions about a scientific phenomenon based on their prior knowledge and experiences. They are then given an opportunity to observe the phenomenon and collect data to support or refute their predictions. Finally, students are asked to explain their observations and reconcile any inconsistencies between their predictions and the actual results. In Task 3 of our game, users begin by making their prediction (Task 3 - Predict). The observation phase in POE is mapped in our case to a simulation where the user can change the number of Cellulans and observe the speed of information spread. Finally, this stage is followed by Task 3 - Confirm, which maps to the explanation phase in POE. Thus, users have the opportunity to modify or confirm their prediction as well as provide an explanation.

4.4 System Description

We design a framework to easily develop learning activities applying the learning strategies discussed in Section 4.3. Our system is composed of four main components: Agents, Environment, Learner(s) and Interactions. Furthermore, we seamlessly support two modalities of the activity, a virtual and a physical one (Figure 4.2 shows the physical modality setup).

1. **Agents:** Our agents are the *Cellulans* from our learning activity (ref. Section 4.2). These can be either virtual agents on the screen, or physical agents embodied by the Cellulo robots.

2. **Environment:** The environment is the space where the agents are placed and are interacting. In the virtual modality, it consists of a virtual "room" possibly including real-time changes and animations depending on the (sub-)level objectives. In the physical modality, it consists of the printed sheets of paper "augmented" with a dot pattern. The graphics on the paper can be designed according to the specification of the activity. While the environment in this case is thus, by default, static, we make it dynamic by projecting visual elements on it from a small overhead projector (Figure 4.2).

3. **Learner:** The learner is the user interacting with the virtual/physical agents and the environment.

4. **Interactions:** Our system enables a number of interactions occurring between the first three components.

   (i) *Agent-agent interactions:* These interactions occur at the "micro" scale upon the encounter of two agents. We have implemented several agent-agent interaction rules, including centralized rules such as *follow a leader*, and decentralized ones such as *repel* (avoid collisions with nearby agents) or *cohere* (attempt to stay close to neighbors) or *align*(attempt to match the velocity of neighbors). Multiple rules were implemented in the overall framework but the ones used in the final version of the game are explained in Table 4.1.

Rule	Description	Interaction – How To	Tasks
Follow Mouse	The Cellulan will follow the mouse dragged by the user.	Learner-Agent Interaction – The user clicks and drags the Cellulan using the mouse.	1,2
Follow Way-points	The Cellulan will follow predefined waypoints defined by the user.	Learner-Agent Interaction – The user can add waypoints by clicking with the mouse on the map.	1,2
Random	The Cellulan will move randomly in one direction; when it hits a wall it follows along it.	Agent-Environment Interaction – No control is given to the user.	1
Olympics	The Cellulans will first stand in line. Each Cellulan is observing the one infront of it and consider it as its leader. When its leader moves, the Cellulan will move towards the last position where the leader previously stopped.	Agent-Agent and Learner-Agent Interactions – Control over the first Cellulan in line is given to the user who can move it with their hand (in the physical modality) or with the mouse (virtual modality).	2
Snake	The Cellulans will stand in line. The first one is leader. Everyone follows. You have control over the first Cellulan in line.	Agent-Agent and Learner-Agent Interactions – Control over the first Cellulan in line is given to the user who can move it with their hand (in the physical modality) or with the mouse (virtual modality).	2
Repel	Each Cellulan will steer away from its neighbors and from the walls so that nothing is occupying its halo. The closer the neighbors are, the greater the force is.	Agent-Agent Interaction – No control is given to the user.	2
Center And Expand	The Cellulans will all go to the center first, decide where each one goes and then expand outwards to occupy the grid.	Agent-Agent and Agent-Environment Interactions – No control is given to the user.	2
Go to closest	Each Cellulan will find the closest target and move towards it. If it is occupied, it will find the second closest target.	Agent-Environment Interaction – No control is given to the user.	2

Table 4.1: The rules present in the game, their description, interaction detail and task in which they are present.

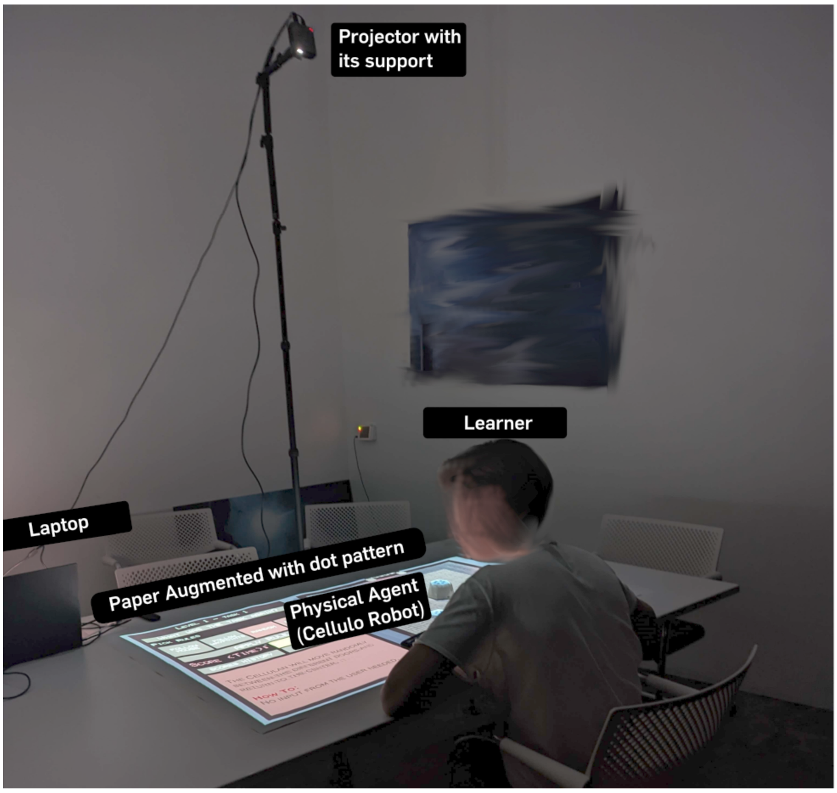


Figure 4.2: Physical modality setup

(ii) *Agent-Environment interactions:* These interactions occur when the action of an agent is directed towards the environment. In this case too, these can be centralized rules such as following predefined “spots” or “targets”, as well as decentralized ones such as leaving marks/pheromones trails. Multiple rules were implemented in the overall framework but the ones finally used in the final version of the game are explained in Table 4.1.

(iii) *Learner-Agent interactions:* In the virtual modality, these interactions are mediated by mouse or keyboard commands for moving the agents. In the physical version, the interactions can also be tangible, i.e, learner(s) move the robots by their hands (thus applying forces to the robots) and feel haptic feedback (thus feeling forces applied by the robots).

(iv) *Learner-Environment interactions:* These interactions mostly relate to the user interaction with the application itself such as UI buttons and rules-related buttons (Fig. 4.3b).

All these interactions are made possible through a cross-platform Unity application which runs the logic of the activity and uses the developed SDK futher detailed in Chapter 9.

4.5 Iterative Design

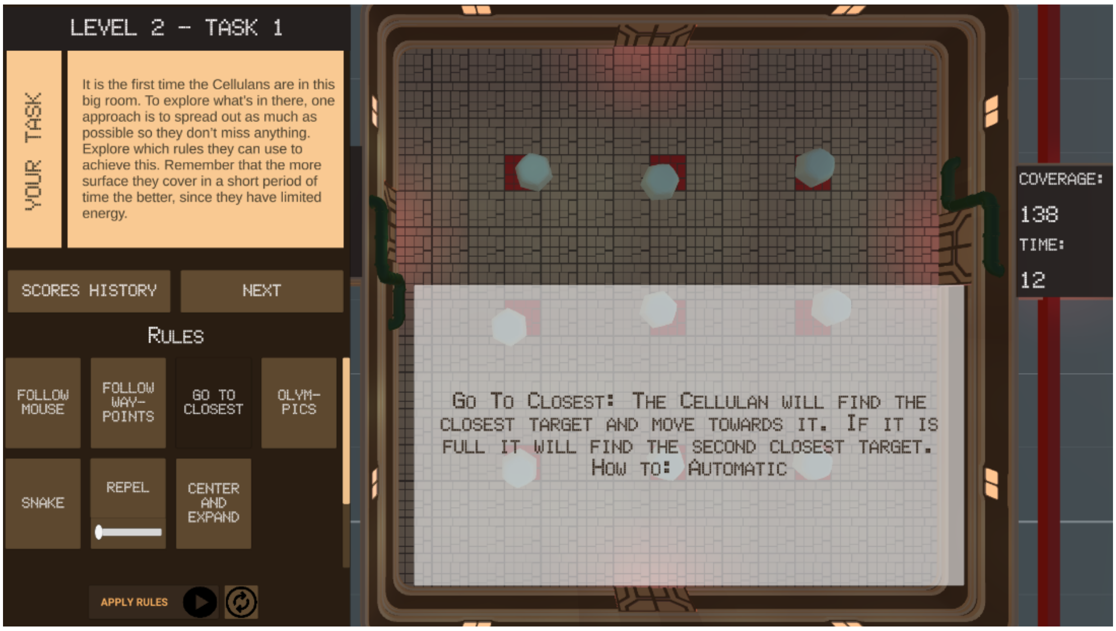
An iterative approach was employed in the development of the game and its user interface. Initially, a preliminary version was tested by colleagues to ensure that it met functional requirements and was free of any bugs that could hinder gameplay. Subsequently, a second version was tested with a sample of 13 students in a school, whose ages ranged between 15 and 18 years old. This pre-experimental test was conducted to assess the understandability of the game, as well as the usability of the UI, and the time required to complete the game. During the experiment, users' behaviors were observed, and any questions or comments they had were recorded for further analysis.

Based on the observation and feedback received from these tests, six major changes were made to the activity to enhance its functionality and improve the user experience. These modifications included changes to the UI layout, task descriptions, color scheme, coverage scoring system, and the addition of a tutorial and final task.

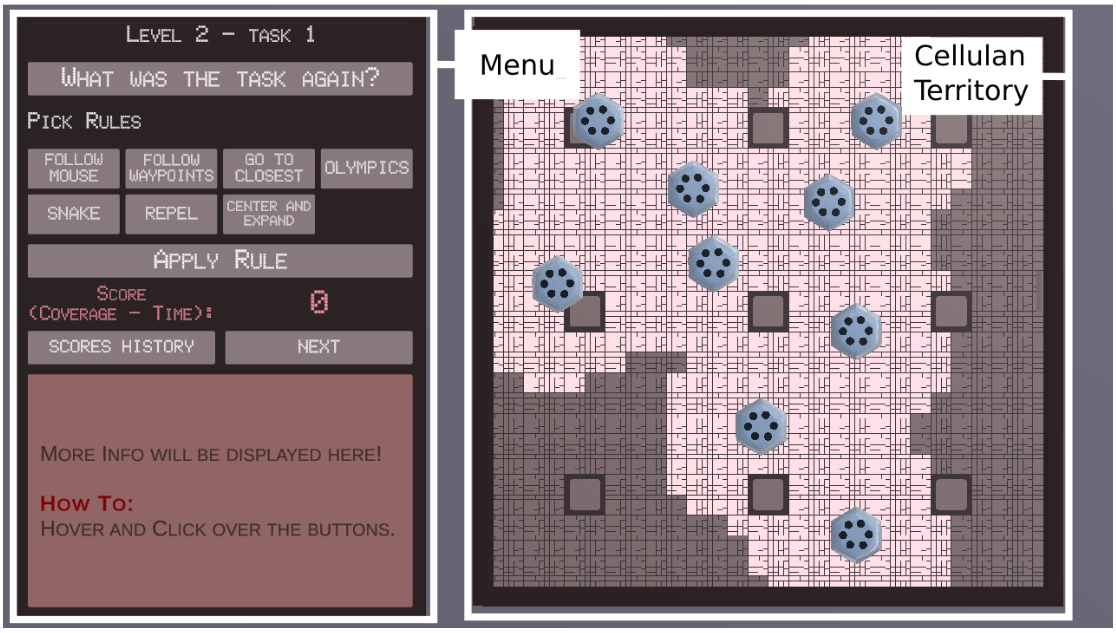
- All UI elements were moved to one side of the game screen, while the Cellulan territory was moved to the other side (Figure 4.3b). This was done to address an issue where the explanations of the game rules displayed on the Cellulan territory were hard to read in the physical modality due to the obstruction caused by the physical robots (Figure 4.3a).
- The game's task description space at the top left corner of the screen was removed and replaced with a transition character that provided a clear storyline and information on specific tasks and levels (Figure 4.4a). This made the game's transition from level to level much clearer, as users were no longer required to re-read the task description each time the level changed, which was not very obvious for them.
- The game's color scheme was changed to make it more robust when projected onto a table for use with physical robots.
- The coverage score in the game was made more concrete by adding a visual element that marked the covered tiles (as seen in Figure 4.1d), rather than simply displaying a count of the tiles covered.
- A tutorial was added to the beginning of the game to teach users how to navigate the menu (Figure 4.4b).
- A final task (Task 4, Figure 4.1h) was added to the game, serving as a closing task. From a pedagogical perspective, it also serves as an overall reflection task for the users and gives us insights into their ways of thinking.

The third version of the game underwent another round of testing with 5 users aged between 18-20 years old who were recruited on Prolific[2], an online platform for research participant

[2] www.prolific.co

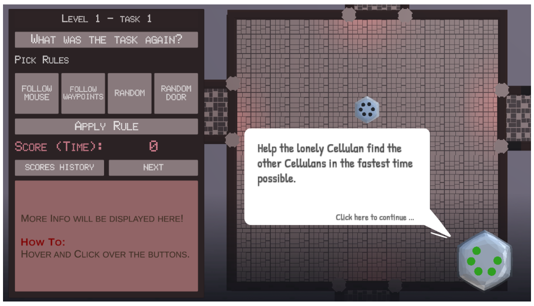


(a) First Iteration of the game UI

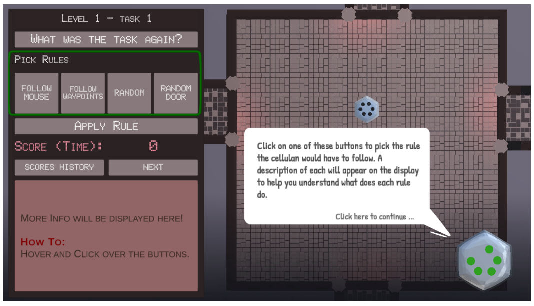


(b) Final Iteration of the game UI

Figure 4.3: Iterative Design Approach of the game UI



(a) Character providing the story line at each transition between levels.



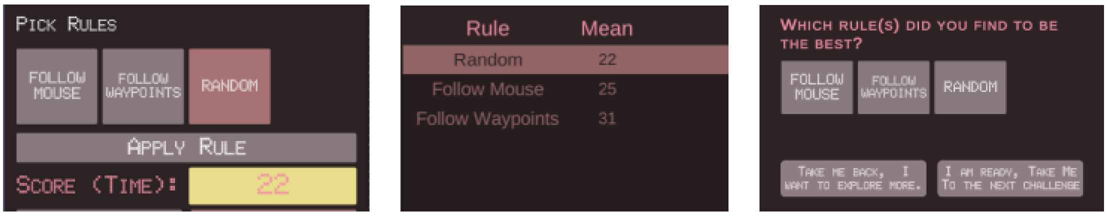
(b) Example screenshot of one instruction in the tutorial.

recruitment. Analysis of the collected behavioral data suggested a hypothesis that some users were bypassing the first levels without completing the assigned tasks. To address this issue and ensure that participants understood the game's objectives, a mandatory exploration of Task1-Level1 was enforced before they could progress to the next level.

Subsequently, the final version of the game was evaluated by another group of 5 people aged between 18-20 years old who were bachelor students at EPFL. They were asked to rate the clarity of the tasks, rules description, navigation in the menu, and robot behavior on a Likert scale ranging from 0 to 5. The results showed an average rating of around 90% on all aspects, indicating that the game is intuitive and easily understandable. Based on these results, we finalized the game design and proceeded to validate the activity's effectiveness towards fostering complex system understanding in the series of studies described in Chapters 5 and 6.

Final UI Design

The user interface (UI) for the platform is designed with a clear boundary of two distinct sections, as shown in Figure 4.3b.
The *agent territory*, which covers 60% of the space, provides a dynamic area where robots can act, move, and interact with each other, the environment and the user. On the other hand, the *menu* accounts for 40% of the space and contains all the necessary information. The bottom of the menu section features a display area, which functions as a canvas for the explanations, while the top comprises a series of buttons that allow the user to review the task, select and apply rules, and navigate through the levels. Hovering over each rule in the second group of buttons reveals a description of it in the display area (Figure 4.3b). Upon selecting a rule, its color changes to indicate the selection, and clicking on the "Apply Rule" button initiates its execution.
The platform provides a rewarding experience to the user by generating a happy sound once a task is completed, and highlighting the final score in the menu, as shown in Figure 4.4a. Additionally, in each level, the user's scores are recorded in a history table that pops out on the display area (Figure 4.4b), with entries sorted such that the best score appears in the first row. In case of multiple trials of the same rule, the history table calculates the mean score, as well



(a) Highlighting buttons/score (b) Score history table (c) Display when Next is clicked

Figure 4.4: UI components

as the best and worst trials for that rule. Finally, a "Next" button is available for the user to progress to the next level. Upon clicking the button, the user is asked to choose the best rule(s) they found (Figure 4.4c). The well-organized and visually appealing UI is designed to create a seamless user experience, facilitating learning in an engaging and interactive manner.

4.6 Conclusion

This chapter presented a comprehensive overview of the design for and implementation of a novel learning activity integrating both virtual and physical agents and aiming to foster complex systems understanding. The learning activity was developed as a game, and it draws the learning theory of conceptual change to deliver an engaging learning experience. Furthermore, the learning activity underwent an extensive iterative refinement phase before the final version was reached. This approach ensured that the game was refined and improved in response to user feedback and testing results, allowing for adjustments to be made at various stages of the game development, resulting in a usable and understandable activity.

In the following chapters, we will build on this groundwork by conducting experiments to assess the effectiveness of the game and analyze users' behaviors with the different modalities supported by the game (virtual and physical). Additionally, we will study more closely the learning effectiveness and user experience of the game.

5 Effectiveness of *Cellulan World*

5.1 Introduction

This chapter presents the first study conducted with the learning activity previously presented in Chapter 4, which aims to investigate the effectiveness of the activity in improving complex systems understanding while also examining the impact of modality on participants' behaviors, performance, and learning gain. Specifically, the following research questions are addressed:

RQ5.1 What is the effect of the learning activity on participants' learning of complex systems as measured by the score on the test instrument?

RQ5.2 What is the relationship between participants' within activity performance, behavioral patterns and their learning?

RQ5.3 What is the effect of the tangibility provided by physical robots on the participants' learning and engagement?

RQ5.4 What is the effect of the modality on participants' within activity behaviors?

The first research question (RQ5.1) aims to verify if the learning activity results in any learning gain in participants' understanding of complex systems. RQ5.2 seeks to explore the relationship between participants' within-activity performance, behavioral patterns, and their learning, which can help shed light on the learning process through the activity. RQ5.3 focuses on investigating the impact of physical robots on learning. Finally, RQ5.4 examines the effect of the learning medium (real vs virtual modality) on participants' within-activity behaviors. Answering these questions will provide insights into the effectiveness of our learning activity.

5.2 Experimental Design

5.2.1 Experimental Setup

The experiment is conducted in a sequence of the following stages :

1. **Welcome**: The experimenter (myself) welcomes the participants and provides a brief explanation of the experiment stages (which are detailed below). Participants are also asked to sign the consent form during this stage.
2. **Pre-questionnaire**: Participants are asked to complete a pre-questionnaire which collects the background metrics.
3. **Pre-test**: The participant answers the questionnaire developed as an assessment instrument (described in Chapter 3).
4. **Learning Activity**: The participant plays *Cellulan World.* In the virtual modality, they are directed to a web-app. In the real modality, the game is already setup on another table and they just needed to move few steps there.
5. **Post-questionnaire**: After finishing the game, participants are asked to fill a post-questionnaire (Table 5.1) which includes questions about game clarity and game engagement.
6. **Post-test**: The participant answers the pre-test questionnaire again, with a shuffled order of the questions.
7. **Goodbye**: The experimenter asks if participants have any questions, answers them if any, and offers them a chocolate as a token of appreciation for participating in the experiment.

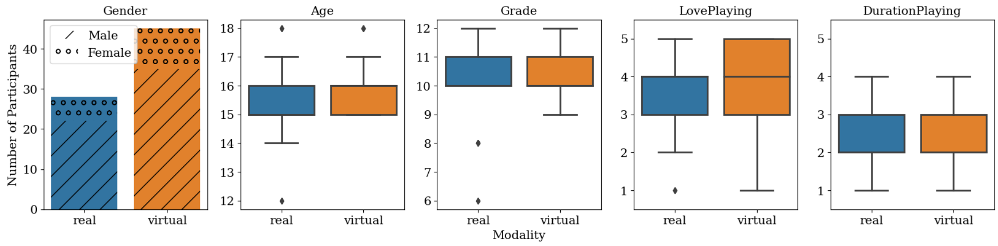
The experiment with all its stages lasts for a maximum of one hour. There is no strict time keeping for each stage.

5.2.2 Participants

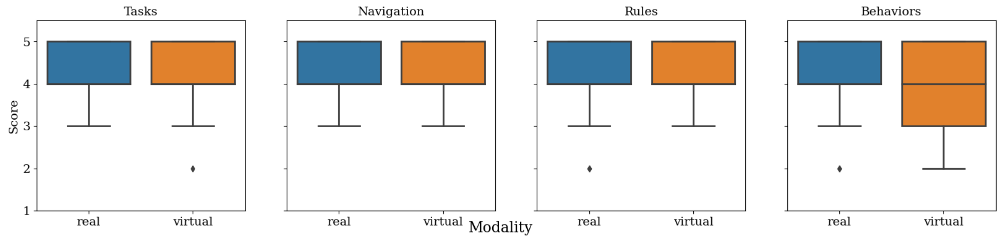
The study was conducted in three international schools in Switzerland over three weeks between the end of November and start of December 2022. A total of 87 learners participated in the study. However, 14 learners' data had to be removed due to incomplete post-test data leaving us with a dataset of 73 participants. The participants were split into two groups corresponding to each modality: real or virtual. 45 participants were in the virtual condition and 28 belonged to the real condition. Due to organizational difficulties, they were not randomized over schools. Participants in the virtual modality group all belonged to one school and were all students of a Design and Computer Science class. Participants in the real modality

Question	Category
The tasks were clear Navigation in the menu was intuitive. The description of the rules is clear. The robots behaved as expected.	Game Clarity
I enjoyed trying to find the best solution. I thought this activity was quite enjoyable.	Affective at Task Level (IMI)
I think I did pretty well at this activity. I am satisfied with my performance at this task.	Perceived Competence (IMI)
I was trying very hard to find the best solution. It was important to me to do well at this task.	Cognitive at Task Level (IMI)
I felt tense while doing this activity.	Pressure/Tension (IMI)
I felt involved in the activity.	Involvement
How many minutes do you think you spent on the part where you played the game ?	Perception of time
Additional questions for the real modality	
I liked the robots. I think the robots are distracting	Real robot behavior (Godspeed-inspired)
Imagine playing the same game with no robots. I think it would have been much less/more engaging.	Real robot engagement (Godspeed-inspired)

Table 5.1: Categorization of the questions in the post-questionnaire. A 5 points Likert scale is used for all items, except when they are asked to provide the number of minutes. More information about the items are provided in Sections 5.3 and 5.6.2



(a) Comparison of general background information between modalities.



(b) Comparison of participants' feedback on game clarity between modalities.

Figure 5.1: Comparison of control variables between modalities. No significant differences were found.

group belonged to the two other schools (5 from one and 23 from the other) and they were randomly chosen from classes by an organizer from the school.

5.3 Confounds Check

Before proceeding to answering our research questions, we check for the existence of potential confounding factors that could affect learning, performance or behaviors. These confounding variables include the background as well as game clarity metrics.

Background

As part of the pre-questionnaire, we collect the following background metrics of the participants:

- their gender
- their age
- their school grade
- their liking of playing video-games (on a scale from 1[not at all] to 5[a lot])
- the amount of time they usually spend playing video games in a week collected as a choice of 4 ranges (I don't play, 1-3 hours per week, 4-8 hours per week, more than 8 hours per week).

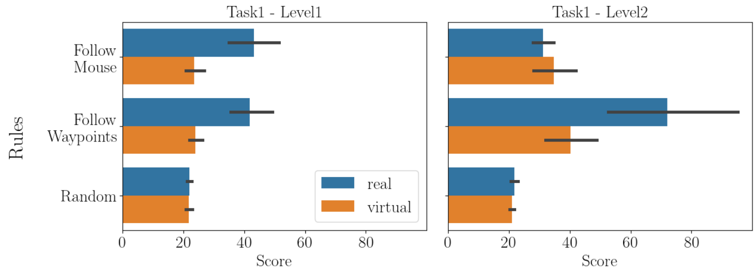
Overall, the participants included 16 Females and 57 Males, were between 12-18 years old (Mean= 15.7, SD = 1.18) and from grades 6 to 12 (Median = 10, IQR = 1). Their reported love for playing video games was 3.78/5 on average (with a SD = 1.13). The amount of playing time reported had a median of 1 to 3 hours per week. Figure 5.1a shows the distributions and the comparison between the background information of the participants in the two modalities. We compare the background metrics between the two groups and we find no significant difference, indicating that the two groups consisted of participants whose background was similar, therefore, we conclude that the general background is not a potential confound.

Game Clarity

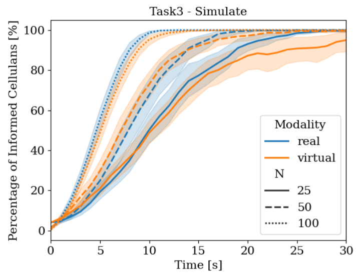
As part of the post-questionnaire, we asked for the participants' feedback on their perception of the comprehensibility of the tasks and descriptions in the game as well as comprehensibility of the robots' (real and virtual) behaviors on a Likert scale ranging from 1(not at all) to 5(very) (Table 5.1). We found no significant difference between the two modalities in terms of participants understanding of the game, showing that this is not a potential confound. Overall, the game was clear as the mean ratings on the questions were all above 4 (Figure 5.1b).

Manipulation Check

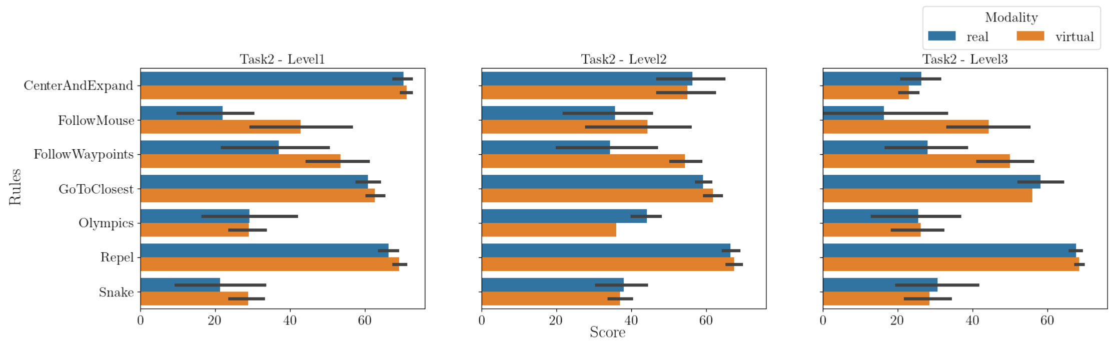
To ensure that differences in implementation between virtual and real modalities did not confound our results, we took steps to minimize these differences and adjusted the speed of the robots to be consistent across both modalities. We also conducted an additional analysis to check for any manipulation effects. This manipulation check is reasonable because it directly addresses potential confounding factors and systematically evaluates the impact of our experimental manipulation on the consistency of results between the two modalities. Specifically, we collected scores from all rules across all tasks and compared them between modalities, as shown in Figure 5.2. We found that scores were similar across all automatic rules, which did not require human control. The only differences were observed in rules that required human intervention, such as Follow Mouse or Follow Waypoints, which were related to user input rather than software speed. In Task 3, we found that simulation curves were similar between modalities. In conclusion, our analysis confirms the absence of any manipulation effect that could depend on implementation differences. We carefully controlled the speed of the robots to ensure minimal differences between the real and virtual modalities, and our comparison of scores from rules in tasks further supports this conclusion. Therefore, any observed differences in participant behavior or learning outcomes can be confidently attributed to the modality itself, rather than to differences in implementation.



(a) Scores in Task 1 (lower is better)



(b) Scores in Task 3



(c) Scores in Task 2 (higher is better)

Figure 5.2: Manipulation check: Comparison of scores for rules in the game's tasks between real and virtual modalities.

5.4 Learning Gain

In order to answer the research question about the learning efficacy of our learning activity in the context of complex systems understanding, we define the desired learning outcome as the shift towards a more complex systems mental model. As a learning metric, we use the **relative learning gain (RLG)**, a commonly used metric in learning sciences research (Hake, 1998), calculated by subtracting a participant's post-test score from their pre-test score, and dividing by the difference between the maximum score that can be achieved and the pre-test score. Concretely, this metric grasps how much the participant learned of the knowledge that he/she didn't possess before the activity. As assessment tool for the participants' mental model, we use our developed assessment instrument (described in Chapter 3) both in the pre and post tests.

Figure 5.3a shows the distribution of pre- and post-test scores of all the participants aggregated over modalities. We observe a significant increase in the scores from pre to post-test. The average relative learning gain is equal to 11.8%. We also examine potential pre-score bias, and find no correlation between relative learning gain and pretest scores (Figure 5.3b, Spearman, with $p >> 0.05$).

Digging deeper, we look at scores for each individual concept (Figure 5.3c) as well as for each individual scenario (Figure 5.3d) of the assessment instrument. The significant learning gain is observed over the *control* and *agent* concepts. Concerning the scenarios, the learning gain is significant in the *Scatter* and *Robots and Gold* scenarios. Interestingly, these scenarios are near-transfer scenarios since they closely relate to the task in the learning activity.

Conclusion: What is the effect of the learning activity on participants' learning of complex systems as measured by the score on the test instrument? → A significant positive learning gain.

5.5 Behavioral Analysis

In order to gain a deeper understanding of the learning process of participants during the learning activity, we analyze their behavioral patterns throughout the game and their performance. By analyzing the behavioral patterns, we aim to uncover the strategies that the participants employ when playing the game and thus actions done by the participants that possibly led to learning.

5.5.1 Metrics

Behavioral metrics

The behavioral dataset consists of the logs recorded during the game which includes all user clicks and keyboard input as well as the positions of all robots (virtual and real). Table 5.2

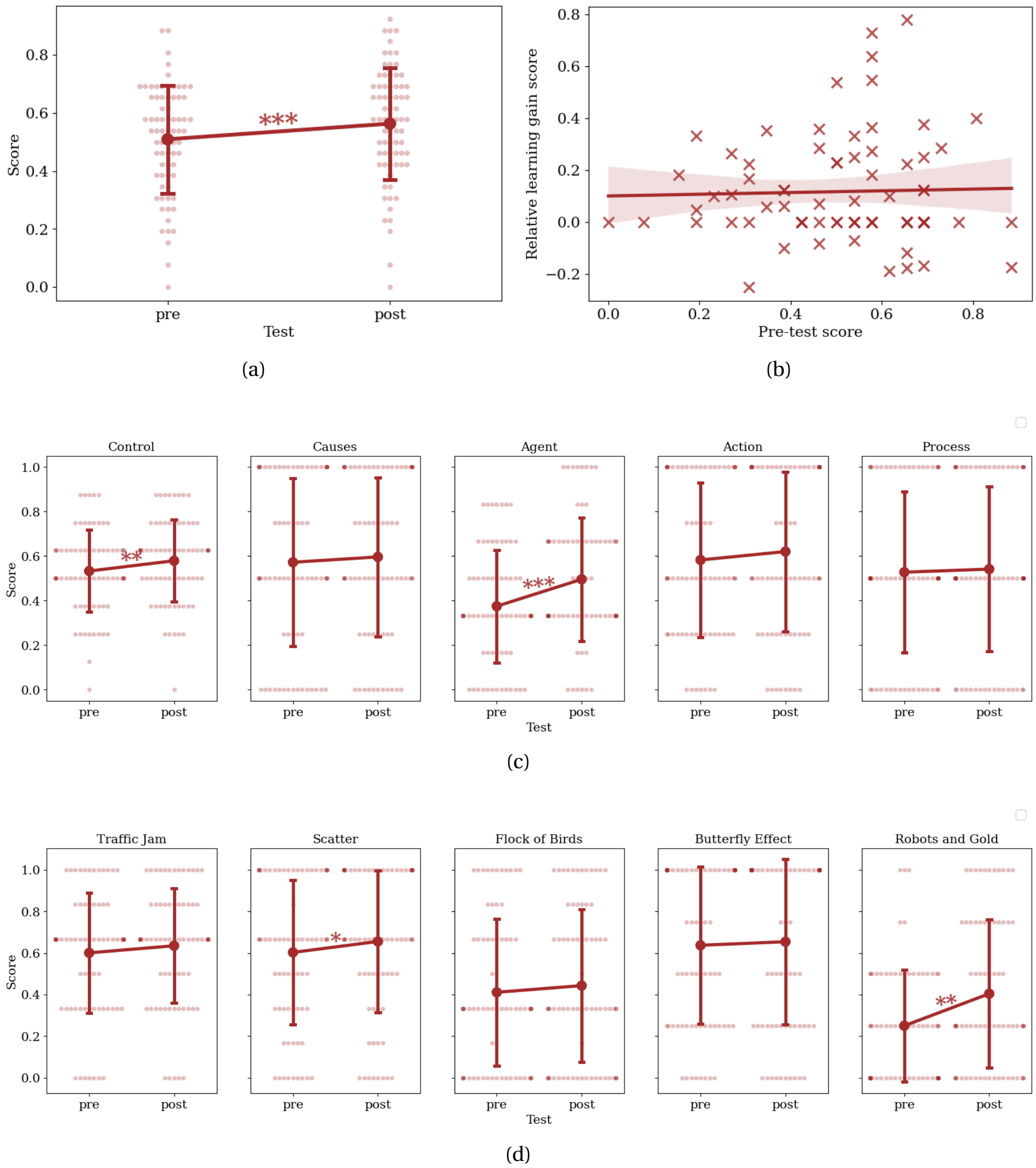


Figure 5.3: (a) Total pre- and post-test scores. (b) Correlation between pre-test scores and relative learning gain. (c) Pre- and post-test scores per concept. (d) Pre- and post-test scores per scenario. * (**) denotes statistically significant differences with p-value < 0.05 (0.005) with a paired T-test in (a) and with a Wilcoxon-test (non parametric version of a paired T-test) in (c) and (d).

Feature	Description	Tasks
time_spent	the total time spent during a level	1,2,3
percentage_reading_time	the percentage of the time spent on reading an instruction or rule description as compared to the total time spent in a level. This is estimated by calculating the time between a cursor entering and exiting a rule button in the menu.	1,2
history_check	how many times the history button is clicked.	1,2
tried_rules	how many times any rule was clicked.	1,2
complete_trials	how many times a trial was completed, i.e, the rule was applied and a score was recorded.	1,2
rules_percentage	the percentage of the unique rules tried over the number of available rules in a level.	1,2
task3_n	the count of simulations done in Task 3	3

Table 5.2: Behavioral features, their description, and the tasks they relate to

defines the computed features from the logs. The suggested features are thought to provide valuable insights into participants' learning processes by considering their involvement, effort, attention, pre-planning, reflection, and exploration. Time spent reflects the level of involvement and effort, while the percentage of time spent on reading rule descriptions might serve as an indicator of attention to rule comprehension and pre-planning. History check serves as a sign of reflection, and the remaining features represent various exploration metrics related to the application and completion of rules. Additionally, we also collect the "letter" text submitted by the participant in the last Task 4 (ref. Chapter 5, Section 4.2).

Performance

Let us consider the rules selected by a participant at the end of a level: if they include a distributed rule, this can be considered a sign of the participant's understanding of the advantages of distributed systems. We thus consider this as a relevant metric for our analysis of the user's approach to the game. Concretely, we check whether the participant's submitted best rules include a distributed one for all levels in Task 1 and 2. In Task 3, we check whether the submitted answer is "faster" or not. More formally, we first define a level-specific score as:

$$d_{tl} = \begin{cases} 1 & \text{if distributed rule chosen(Task1 and 2)/faster choice(Task3)} \\ 0 & \text{otherwise} \end{cases} \tag{5.1}$$

where $t \in \{1,2,3\}$ denotes the task and $l \in n_t$ denotes the level within the task where n_t indicates the number of levels per task.

As discussed in Chapter 4, we have carefully designed the levels to demonstrate that, although a centralized rule effectively accomplishes the task in the initial level, its efficacy diminishes

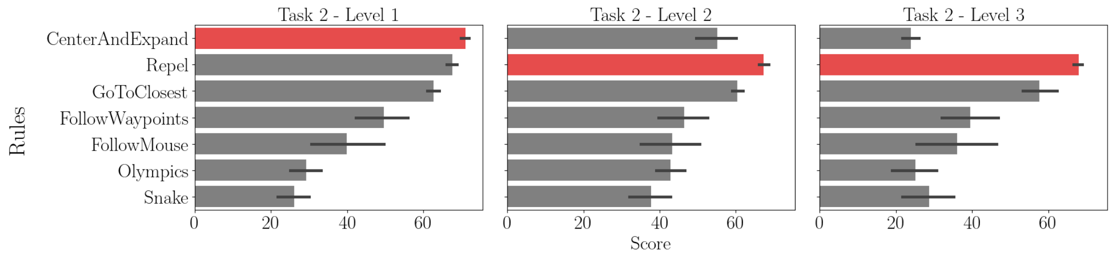


Figure 5.4: Comparing scores in the three levels in Task 2. The best rule is highlighted in red.

in subsequent levels. This highlights the superiority of a decentralized rule in those situations. For instance, in Task 2 (ref. Figure 5.4), the Center and Expand rule consistently yields the highest score during the first level, followed closely by the Repel rule, a distributed rule. In the subsequent level, where a robot breaks randomly, the previously best rule falls behind, and the Repel rule emerges as a more robust approach for accomplishing the task. By the third level, the superior performance of the Repel rule becomes even more apparent, as it is by far the best rule. As a performance metric, we aggregate the level-specific score as follows:

$$\sum_{t=1}^{t=3}\sum_{l=2}^{l=n_t} d_{t,l} \times 2^{(l-1)} \tag{5.2}$$

This formula assigns more weight to levels where a distributed rule would be more robust and perform better, which occurs in the later levels of each task. More specifically, this refers to Level 2 in Task 1, and Levels 2 and 3 in Task 2. In Task 3, a higher weight is given to the confirm level as it follows the simulation part.

The normalized performance metric is obtained by dividing the aggregate score by the maximum possible score.

5.5.2 Data Analysis

We explore the relationship between the two types of process variables, namely behaviors and performance, and their correlation with the dependent variable: learning gain. The first step of this investigation involves examining the evolution of the participants' game performance and its association with the learning gain. Subsequently, we analyze the behavioral patterns observed during the game and their link with both performance and learning gain.

Performance

Figure 5.5 shows the evolution of the level-specific score d_{tl}. Overall, we notice a trend towards the increase in d_{tl} within every task, meaning that while learners began with choosing centralized rules, they shifted towards decentralized rules in the higher levels (Figure 5.5). This links back to the design principle of conceptual change, where the first levels are designed

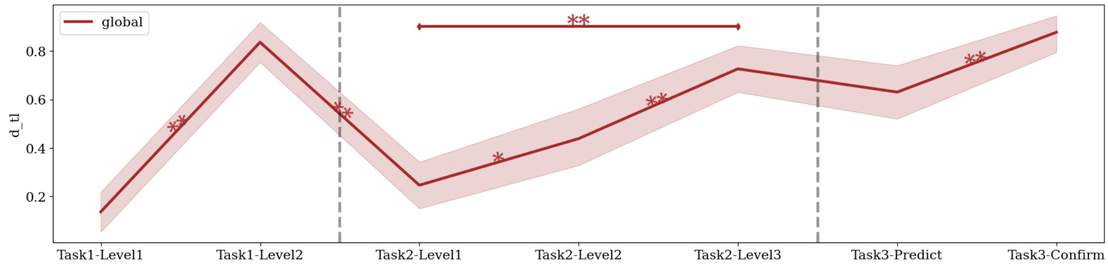


Figure 5.5: Evolution of the performance metric within levels of the activity globally. *(**) denotes statistically significant differences with p-value < 0.05 (0.005) with a Mann-Whitney test.

to be successfully completed with both centralized and distributed rules, while later ones feature changes in the environment and/or robot behavior designed as a way to nudge people towards favoring distributed rules. Similarly, in Task3, where the design principle of predict-simulate-confirm is used, we notice an increase in the choice of the correct answer ("faster") in the confirm level.

Link between Performance and Learning

To the aim of exploring the relationship between the performance and learning gain metrics, we initially conducted a Spearman correlation analysis and found no significant correlation between the two variables (r = 0.11, p = 0.33).
To further investigate this relationship, we divided the data into 4 distinct groups using a median split on the performance and relative learning gain axes (Figure 5.6). This split is in accordance with terminology and concepts commonly adopted in the field of learning and education (Kapur, 2008; Nasir, Bruno, et al., 2021). We identify the four clusters (denoted later as PLG clusters) as:

- **Productive Success (PS)**, i.e. teams that performed well in the activity and also had a positive learning gain.
- **Non-Productive Success (Non-PS)**, i.e. teams that performed well in the learning activity but had a non-positive learning gain.
- **Productive Failure (PF)**, i.e. teams that did not perform well but had a positive learning gain.
- **Non-Productive Failure (Non-PF)**, i.e. teams that neither performed well in the task nor had a positive learning gain.

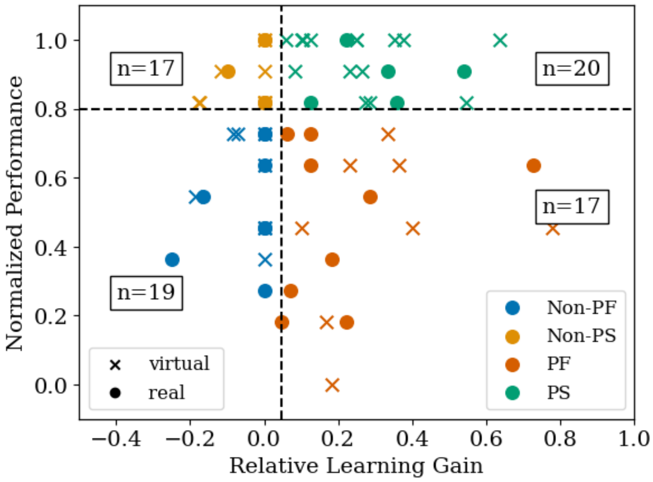


Figure 5.6: Clustering Performance and Learning gain into four categories using a median split on each of the axes. We label them as Productive Success (PS), Non-Productive Success (Non-PS), Productive Failure (PF), Non-Productive Failure (Non-PF).

In-Game Behaviors

To gain insights into the behavioral patterns of learners during the game, we first clustered participants based on the 34 behavioral features (see Table 5.2). Our pipeline consisted of the following steps:

1. Preprocessing: where we standardized the features by removing the mean and scaling to unit variance.
2. Dimensionality Reduction: where we applied PCA on the normalized data and selected the first 15 principal components which explain 90% of the variance within the features (Figure 5.7).
3. Clustering using K-means with k = 2, chosen to maximize the silhouette score.

The clustering results are shown in Figure 5.8. We then verified whether the obtained clusters showed significant differences in the original behavioral features (Figure 5.9). We found that significant differences were only present in the behavioral features of Tasks 1 and 2, indicating that the variances in behaviors were larger in these two tasks. Furthermore, we observed that when a significant difference existed, cluster (1) had a higher mean in all features, except for the percentage reading time features. We hypothesized that a high percentage reading time could either indicate participants spending more time planning and thinking about which rule to use or being slower readers. Participants in cluster (1) spent more time in Tasks 1 and 2, tried more rules, completed more trials, and had a higher percentage of rules used

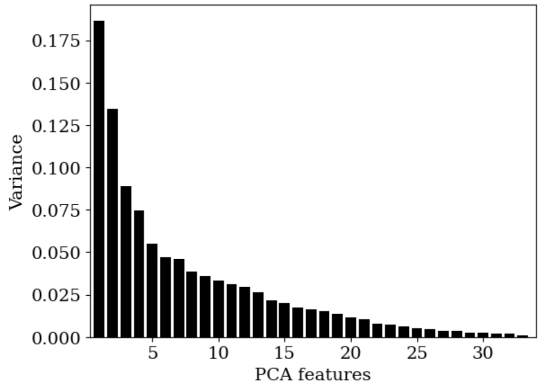


Figure 5.7: PCA - Percentage of variance explained by each principal component

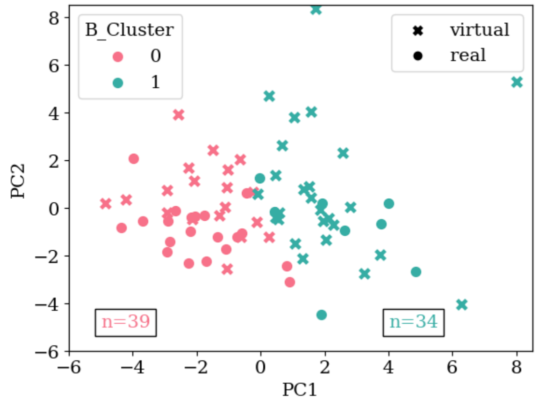


Figure 5.8: Kmeans - Clustering results based on participants' behavioral patterns

overall. Based on these findings, we suggest naming cluster (1) as the **High Explorers** group and cluster (0) as the **Low Explorers** group.

Behavioral Patterns vs Performance and Learning

Following the clustering of the behavioral patterns, we investigated the differences in performance and learning between the two behavioral clusters. The results showed a significant difference in the total performance score between the High and Low Explorers groups, with the former performing better (Figure 5.10). No significant differences were found in the total relative learning gain. Nonetheless, a significant difference in the relative learning gain was observed for Scenario "Scatter", which is closely related to Task 2 of our activity and is one of the near transfer scenarios. We also confirmed that there were no significant differences in the pre-test scores between the two clusters (Figure 5.10).

Furthermore, we examined the similarities between the PLG clusters and the behavioral clusters (Figure 5.11). The analysis revealed that the High Explorers group had a higher proportion of participants in both the Non-PS and PS clusters, while the Low Explorers group had a higher proportion of participants in both the Non-PF and PF clusters. This finding aligns with our earlier result of a significant difference in performance but not in the relative learning gain.

Reported Strategies: Task 4 Letter Analysis

In Task 4 (ref. Chapter 4, Section 4.2), participants were asked to write down their strategies or reflections on how to find the best rules. To analyze the data collected, we adopted a thematic analysis approach (Braun & Clarke, 2006). Firstly, we imported all the data to a MIRO board[1], a collaborative digital whiteboard, with each response presented as a post-it

[1] https://miro.com/

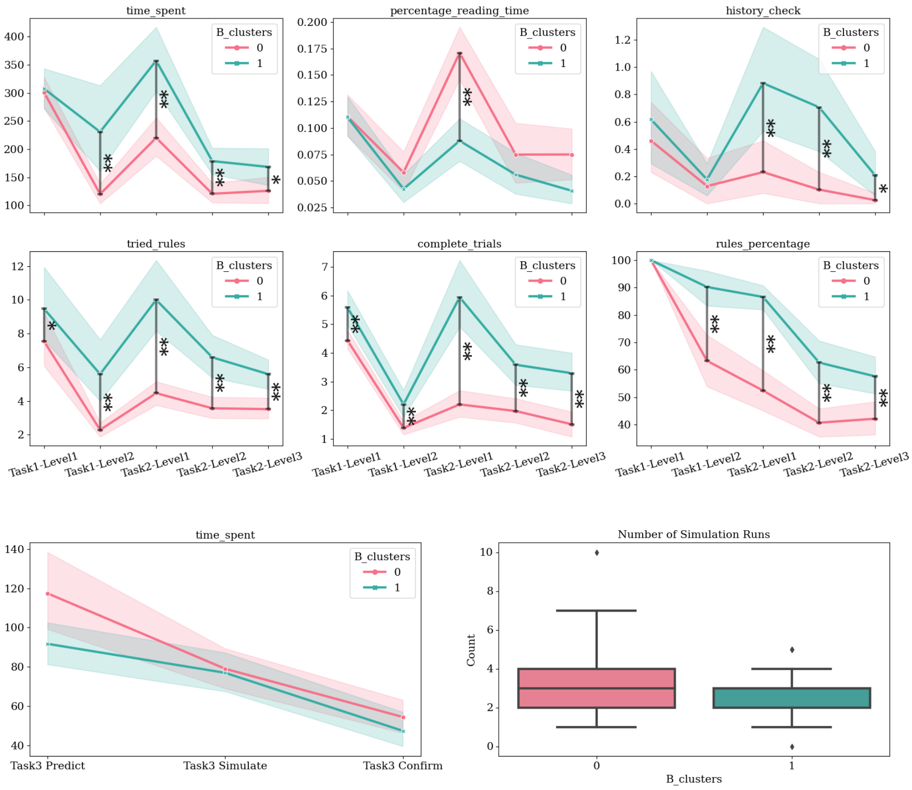


Figure 5.9: Behavioral differences between the clusters obtained by KMeans. *(**) denotes statistically significant differences with p-value < 0.05 (0.005) with a Mann-Whitney test.

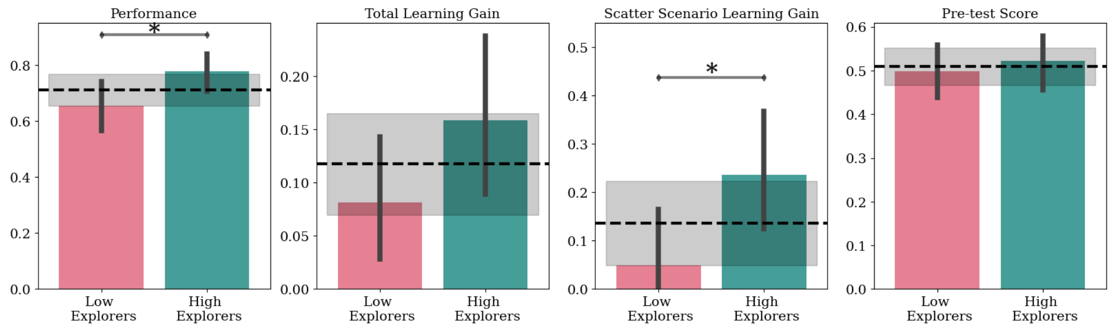


Figure 5.10: Differences between the behavioral clusters over performance and learning metrics. * denotes statistically significant differences with p-value < 0.05 with a Mann-Whitney test.

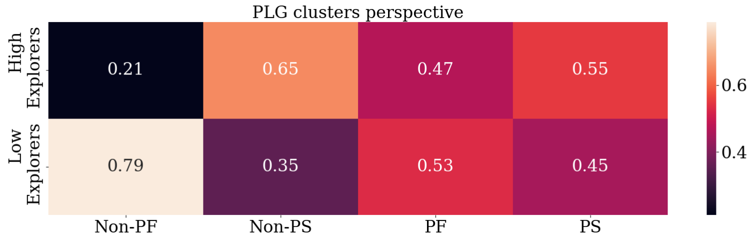


Figure 5.11: Heatmap of the relative frequency of the behavioral clusters with respect to PLG groups.

note. Secondly, we grouped together the answers that shared similarities, based on an implicit understanding of what they represent. We then created broader themes for each of these groups. In a final phase, we re-examined the data that had been placed into each theme to confirm its proper assignment and revised the assignments and categorization if necessary. Three researchers conducted this process together, iteratively until there was a complete agreement on the coding scheme. Through this analysis methodology, we were able to identify common patterns and themes in the participants' responses. The resulting themes are as follows:

- **Leader group** (6 participants): these answers mostly focused on the importance of having a leader and the centralization of decisions. E.g. *it is always important to have a leader to keep your group organized and orderly.*
- **Repel group** (9 participants): these answers focused on mentioning a specific rule (Repel) which is a distributed one. E.g. *If you repelled each other; you could cover more area because you aren't bunched up together.*
- **It depends group** (10 participants): these answers focused on the dependency of the rule on the context, i.e, choosing a rule depending on the different situations. E.g. *To choose the right method, it is necessary to analyze the circumstances and to deduce which method is not affected by the characteristics of the situation.*
- **Communication group** (8 participants): these answers focused on the importance of communication and coordination between the cellulans. E.g. *learn from pattern recognition by communicating to one another as it has proven in our journey to be the most efficient technique.*
- **Testing group** (12 participants): these answers discussed a strategy rather than a specific rule. The strategy involved doing tests with the implication of choosing the rules that score best in the repeated tests. E.g. *My technique for my answer was to try as many options as I could and then I just chose the one that worked more effectively in my opinion.*
- **Best Choice group** (15 participants): these answers were mostly about choosing the best rule, which had the higher score, implying testing different rules to do that, but

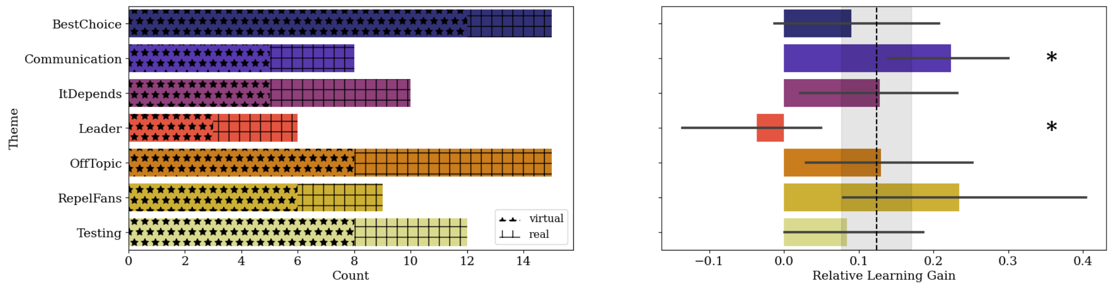


Figure 5.12: Task 4 Analysis - Theme frequency count for each theme and per modality (left). Differences in learning gain metric between themes (right).

unlike the previous group testing was not explicitly mentioned. E.g *The reasoning i had behind my actions were the ones that were simulated to be best or have the best outcome.*

- **Off-topic group** (15 participants): these answers were off-topic or very general. E.g *good luck for the next steps boys*, or *stay together don't panic look for food.*

Figure 5.12 shows the distribution of the participants between themes. Overall, the participants of each group appear to be distributed similarly over the themes. From a learning perspective, the **Leader** group had a significant lower learning gain compared to all other groups combined. Conversely, the **Communication** group had a significant higher learning gain compared to all other groups combined.

A deeper investigation into how participants learn in our game

Currently, our findings suggest that (1) participants' behaviors correlate with game performance, with the High Explorers group demonstrating superior performance, i.e, submitting more distributed rules as the best rules; (2) participants' behaviors do not fully account for the variance in total learning gain, with the exception of a significant difference in relative learning gain for a specific near transfer question in Scenario 2 (Scatter); and (3) performance and learning gain are not significantly correlated. As a result, while our investigation has provided compelling evidence of the relationship between participant behaviors and game performance, the connection between behaviors and learning remains inconclusive.

To further explore the relationship between behavioral features and learning, we follow a backward approach, moving from learning outcomes to the learning process manifested by the behavioral data. Therefore, we divide the population into two clusters of Gainers and Non-Gainers. Gainers are participants who demonstrated a positive learning gain, while Non-Gainers are those who did not show such an improvement.

Subsequently, we conducted a **decision tree classifier** analysis using the behavioral metrics as input variables and the gainers and non-gainers groups as labels. We employed a classification approach with the aim of identifying potential behavioral features that could distinguish

participants who demonstrated a positive learning gain from those who did not. We opted for a decision tree classifier due to its interpretability and ease of understanding, as it employs if-then-else decision rules. The resulting tree is shown in Figure 5.13 with a depth n = 4 and an accuracy of 97.26%. The first node (A) of the decision tree classifier is based on the history check metric at Task2-Level1, which is indicative of the level of **reflection** demonstrated by the participants. At this level, the accuracy is 64.38%, with individuals exhibiting higher history check numbers predominantly associated with gainers (more specifically, 74% of people with higher history check are gainers, whereas only 34% of people with no history check are gainers). We go through the explanation of the rest of the tree, following a depth first search approach. Moving down the tree, for participants with higher history checks, the non-gainers were mostly those who spent very little time in the subsequent Level2 of the same Task2 (Path A-C-F) or those who spent much more time in the next Level3 of the same Task2 (path A-C-G-M-V-W). For the participants with a low history check and lower time in Task2-Level1(A-B-D), the split between gainers and and non-gainers occurs on the tried rules metric. Participants who tried fewer rules were mostly non-gainers (A-B-D-H). On the other hand, participants who spent less time but meanwhile tried more rules (A-B-D-I) were mostly gainers. This suggests a link to the idea of time spent efficiently, related to the concept of "deliberate practice" (Ericsson et al., 1993). Deliberate practice refers to a specific type of practice that is focused, intentional, and involves feedback and reflection. In other words, it's not just about putting in the hours, but about how those hours are spent. Participants who spent less time but tried more rules, as suggested by the decision tree, may have been engaging in a more deliberate and efficient practice.
For the participants who had a low history check but spent higher time in Task2-Level1 (A-B-E), the split is again related to how they spent the time by exploring, this time referring to the complete trials metric. Participants who wait less for a complete run of a trial were mostly non-gainers (A-B-E-J and A-B-E-K).

To summarize, key behavioral features impacting learning include the level of reflection (history check), time spent on tasks, experimentation with different rules, efficient use of time, and thoroughness in task completion. These findings suggest that effective learning is not just about time spent, but also about intentional reflection, and efficient exploration. Therefore, a balance of these aspects seems crucial for positive learning outcomes.

5.5.3 Conclusion and Discussion

What is the relationship between participants' performance, behavioral pattern in the learning activity and their learning gain? → More exploration leads to better performance, and a higher learning gain in the near transfer scenario; however, performance is not correlated to total learning gain. Reflection, manifested by history check, seems to be one of the important factors for a successful learning. An interplay exists between reflection, efficient exploration, time spent on tasks and task completion in the participants behavior which relates to whether participants are learning gainers or non gainers.

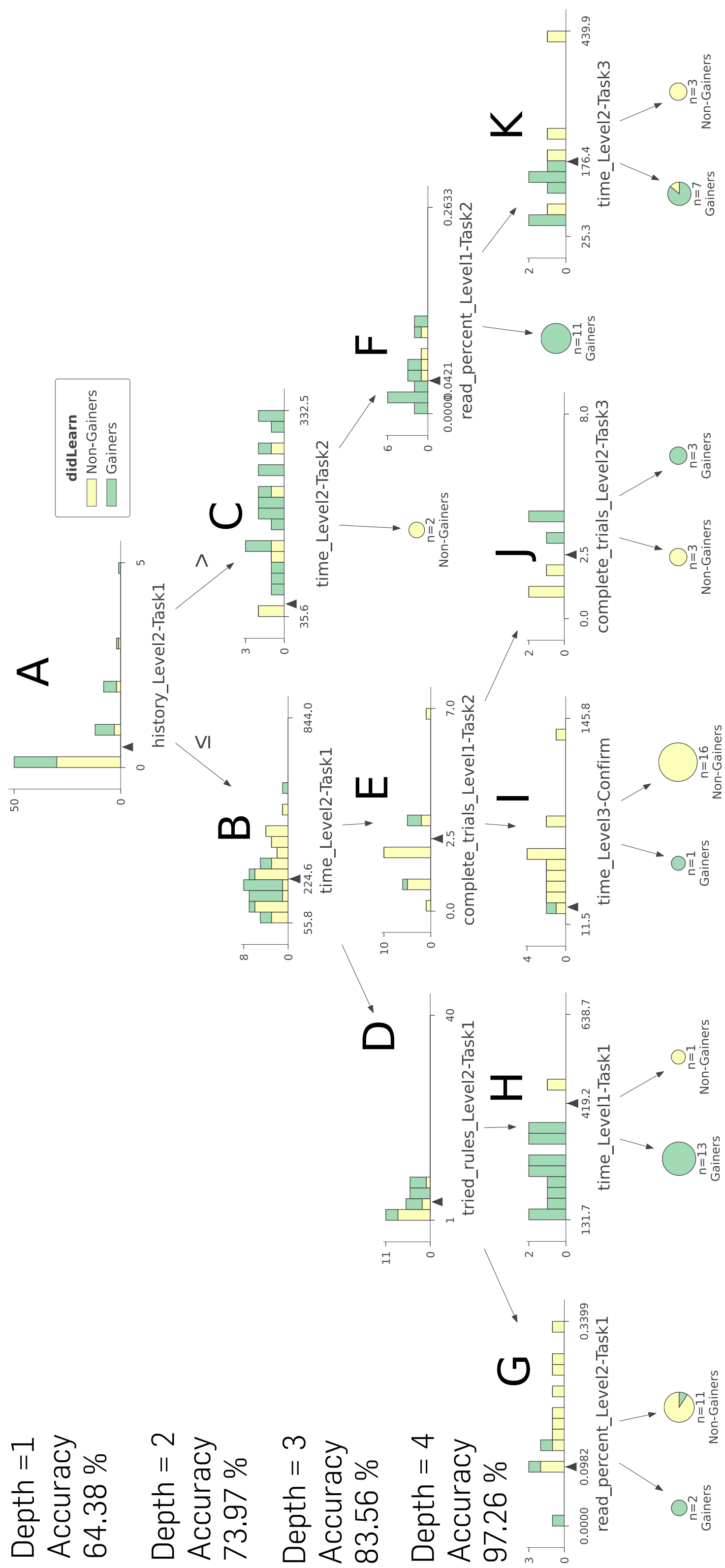


Figure 5.13: Classification decision tree with depth = 4.

These findings are consistent with prior literature in several ways. First, similar to Schneider et al. (2011)'s study, where increased exploration leads to better performance. Second, the results confirm that performance in problem-solving activities does not necessarily indicate the level of learning achieved (Loibl & Rummel, 2014; Nasir, Kothiyal, et al., 2021). Lastly, the importance of reflection (history check, in our case), reinforces prior research that highlights the significance of regulation manifested through reflection in problem-based learning environments (Do, 2012; Etkina et al., 2010; Hmelo-Silver, 2004).

5.6 Added Value of Physical Robots

We investigate two key areas to determine the added value of physical robots to the learning activity: the impact on learning outcomes and the level of participant engagement.

5.6.1 Learning

Both the virtual and real modalities produced a positive and significant learning gain (Figure 5.14a). Participants in the virtual modality scored significantly higher in the pretest compared to those in the real modality, yet there was no significant difference in the learning gain between the two modalities. Moreover, no pre-score bias was found on the relative learning gain (Figure 5.14b). Our analysis of the concepts showed a significant learning gain over the *control* and *agent* concepts, consistently over both modalities (Figure 5.14c). Participants in the virtual modality, however, had significantly higher pre-knowledge on the concepts *action* and *process* compared to their counterparts in the real modality. With respect to the scenarios (Figure 5.14d), participants learned on different scenarios depending on the modality. In the real modality, the learning gain was significant in the *Scatter* and *Flock of Birds* scenarios, whereas in the virtual modality, the learning gain was significant in the *Robots and Gold* scenario. As discussed in Chapter 3, both *Scatter* and *Flock of Birds* represent familiar scenarios, with the former derived from man-made events and the latter based on naturally observed phenomena. Yet, *Robots and Gold* serves as a thought experiment. One explanation for these findings might be that physical robots allowed students to better connect with familiar scenarios, whereas with virtual robots, participants likely associated rule selection in the game more with the thought experiment in Robots and Gold. These findings suggest that the modality used in learning activities might impact the way participants learn, with different modalities being more effective for different scenarios.

5.6.2 Engagement

As part of the post-questionnaire, 9 questions are included in order to measure participants' self-assessment of their enjoyment, engagement, competence and stress. Among them, 7 belong to the Intrinsic Motivation Inventory (IMI) (Ryan & Deci, 2001) which "is a multidimensional measurement device intended to assess participants' subjective experience related

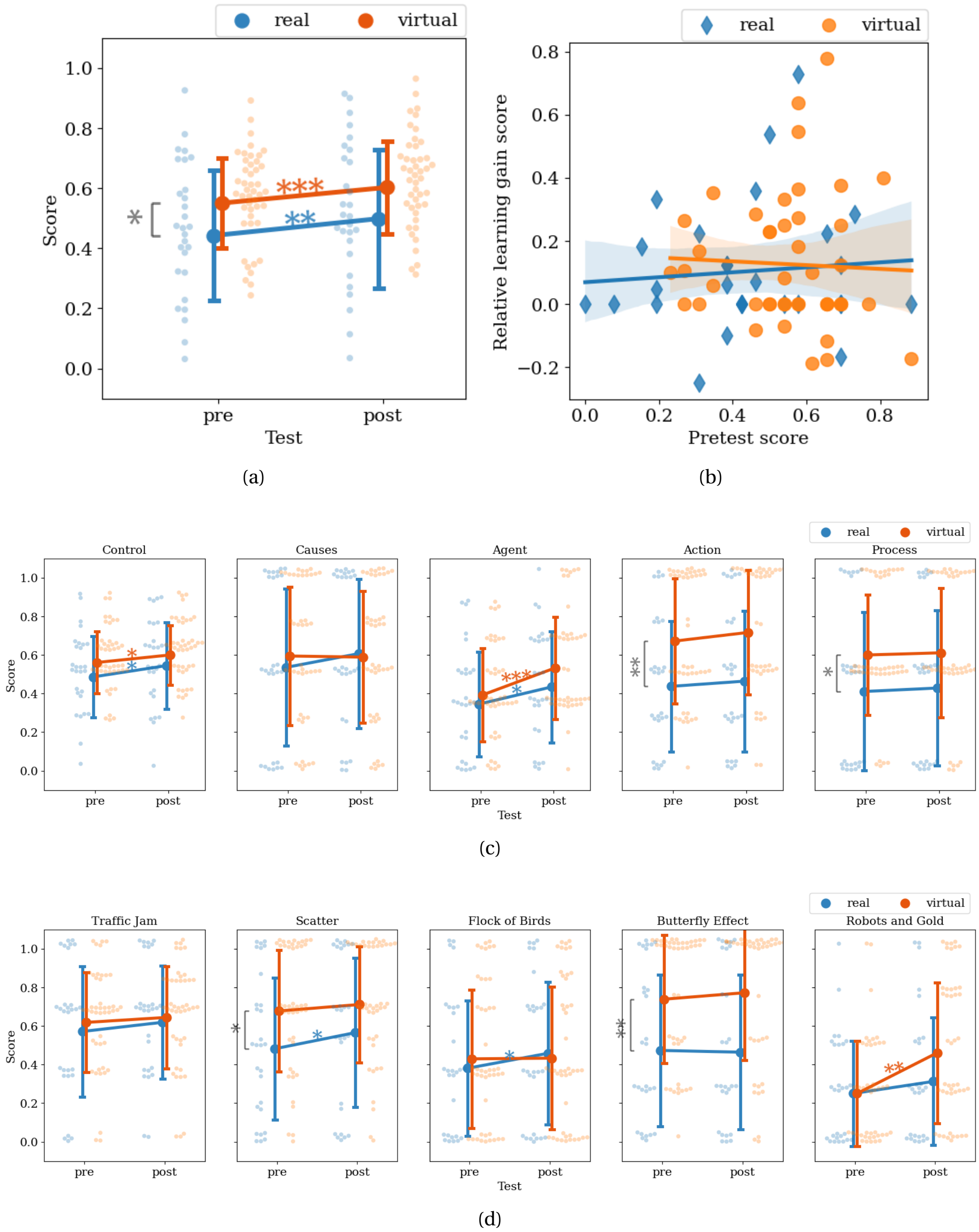


Figure 5.14: (a) Total pre- and post-test scores per modality (b) Correlation between pre-test scores and relative learning gain per modality. (c) Pre- and post-test scores per concept per modality. (d) Pre- and post-test scores per scenario per modality * (**) denotes statistically significant differences with p-value < 0.05 (0.005) with a paired T-test in (a) and with a Wilcoxon-test (non parametric version of a paired T-test) in (c) and (d).

to a target activity in laboratory experiments" and relate to engagement, own competence and stress. One is a question created for the purpose of this work for reported involvement, and one is on the perception of time elapsed. In the real condition, we also add 3 questions inspired by the Godspeed questionnaire (Bartneck et al., 2009), a widely used instrument in HRI to assess the participants' perception of the robot.

Figure 5.15 show the difference between the two modalities on the engagement metrics (ref. Section 5.2.1, and Table 5.1). Significant differences were found on the affective and cognitive scales, with the real modality rating higher. In other words, participants in the real modality enjoyed the activity more than those in the virtual modality ($[M, SD]_r$ =[4.32,0.84] vs. $[M, SD]_v$ = [3.77, 0.68]). Similarly, participants in the real modality were trying harder to perform well in the game than those in the virtual modality ($[M, SD]_r$ =[3.70,0.90] vs. $[M, SD]_v$ = [3.27, 0.70]). This suggests that the interaction with physical robots positively affects the participants' involvement with the activity.

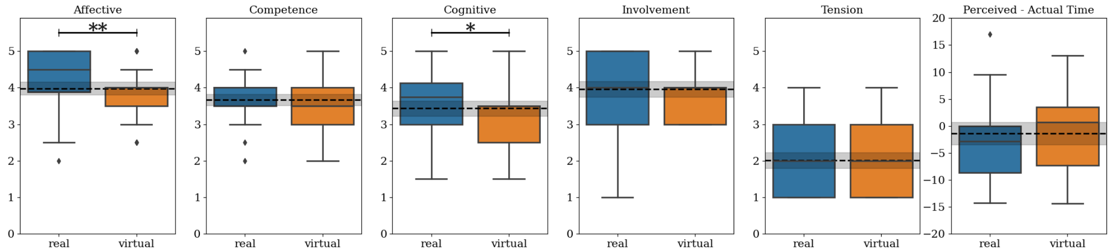


Figure 5.15: Engagement metrics per modality. * (**) denotes statistically significant differences with p-value < 0.05 (0.005) with a Mann-Whitney test.

Additionally, as shown in Figure 5.16, participants in the real condition liked the Cellulo robots (M = 4.32, SD = 0.98), did not find them distracting (M = 1.79, SD = 1.13) and thought that they would be less engaged if the robots were only virtual (M = 1.79, SD = 0.69).

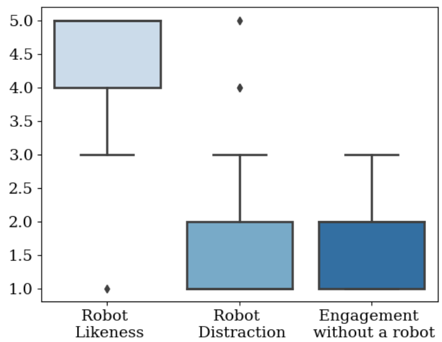


Figure 5.16: Real robot behavior metrics in the real modality corresponding to questions in Table 5.1.

5.6.3 Conclusion and Discussion

What is the effect of the tangibility provided by physical robots on the participants' learning and engagement? → Although there is no significant difference in the learning gain between the real and virtual modality, participants in the real modality reported significantly higher levels of enjoyment and cognitive engagement.

In literature, the effectiveness of Tangible User Interfaces (TUIs) with respect to learning has no clear-cut conclusion and highly depends on the context and target age. Prior experiments between graphical user interfaces (GUIs) and TUIs have shown no difference between the two in learning gains (Marshall, 2007). The findings of our study are in line with previous research as we too found no significant difference in learning gain between the real and virtual modality.

However, participants in the real modality reported significantly higher levels of enjoyment and cognitive engagement. TUIs have been shown to outperform GUIs in regard to being more inviting and promoting collaboration and child-focus, but not in regard to performance, apprehendability, or engagement (Horn et al., 2009). The impact of interactive experiences with TUIs on learners' emotions, such as engagement, initiative, playfulness, enjoyment, immersion, and confidence, has been found to be positive (Li et al., 2022). Finally, many studies have found that participants enjoy tangible learning experiences (Cook et al., 2018; Hengeveld et al., 2013; Lu et al., 2019; Rubens et al., 2020).

Although our study did not find a significant difference in learning gain between the real and virtual modalities, our results and previous research suggest that TUIs can enhance enjoyment and cognitive engagement during learning activities, thereby reinforcing intrinsic motivation, which is crucial to cognitive and social development (Ryan & Deci, 2000).

5.7 Behavioral Differences between Modalities

In order to answer RQ4, we compare the behavioral patterns and performance between the two modalities (real and virtual) to determine if there are any significant differences.

5.7.1 Performance

No significant differences between the two modalities are observed in the total performance score. We also looked more closely to the evolution of the distributed metric d_{tl} (Figure 5.17). A significant difference between the modalities (higher in virtual) is only found in Task1-Level2 (Figure 5.5). One explanation for this difference can be that in that level, only one robot is present, so for the real modality, users still preferred to control the robot themselves instead of choosing a random motion. Furthermore, an interaction effect was observed between the modalities and performance within tasks 1 and 3. Specifically, the virtual modality showed a greater increase in the performance from level 1 to level 2 in task 1. In contrast, in task3 the

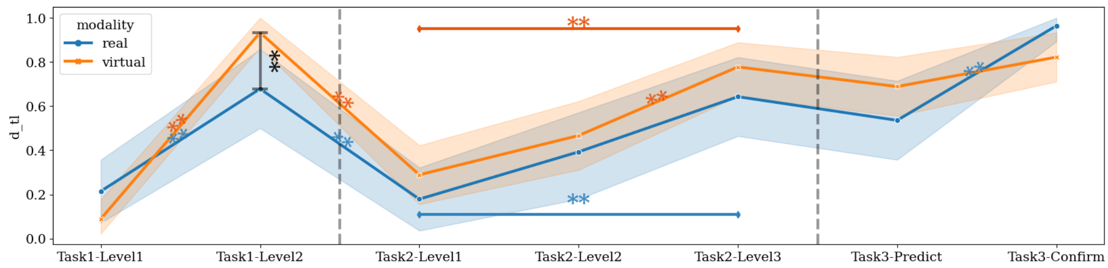


Figure 5.17: Evolution of the Performance metric within levels of the activity globally. *(**) denotes statistically significant differences with p-value < 0.05 (0.005) with a Mann-Whitney test.

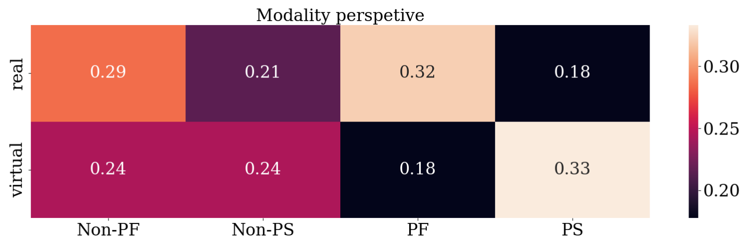


Figure 5.18: Heatmap of the relative frequency of the PLG groups with respect to modalities.

interaction effect is reverse and the increase in performance is higher for real condition than the virtual condition.

Link between Performance and Learning per modality

We examined the distribution of modalities across each of the clusters obtained based on performance and learning gain features presented in Section 5.5.2. Although not statistically significant (p = 0.33) in a chi-squared test, we observed that in the virtual condition, a higher proportion of participants belonged to the PS clusters, while in the real condition, more participants were part of the PF clusters. This finding suggests that even though there were no significant differences between the two modalities concerning performance and learning gain metrics, people learned differently in each modality. In the real modality, more participants appeared to reflect and switch to more distributed answers in the post-test, despite not performing well in the activity itself. We hypothesize that this is related to the reported higher cognitive engagement in the real modality (ref. Section 5.6.2). In contrast, in the virtual modality, more participants seemed to successfully identify the distributed rule as the best rule and transfer that knowledge to their post-test answers.

5.7.2 In-Game Behaviors

By computing the frequency table between the modality and the behavioral clusters (Figure 5.19), we can see that in the real modality, a higher proportion (71%) of the participants belongs to the Low Explorers group, whereas in the virtual modality, a slightly higher proportion (58%) belongs to the High Explorers group. From the perspective of the behavioral clusters, a higher portion (76%) of the High Explorers were in the virtual modality, while the Low Explorers were equally split between the two modalities. A chi-square test reveals a significant difference (χ^2 = 4.8, p = 0.028 < 0.05), i.e. suggesting that the modality would be a predictor for the participants' behavior cluster. In other words, this implies that individuals in the virtual modality are more likely to be high explorers, while those in the real modality are more likely to be low explorers.

Taking a closer look at individual behavioral features (Figure 5.20), we observe that participants in the virtual modality tested more rules in Level 1 of Task 1 and 2, as well as in Level 2 of Task 1. The rules they tried were distinct, as their percentage relative to the total rules was also higher in these levels. However, there are no significant differences in the number of complete trials (except for Task1-Level1), suggesting that not all the tried rules were complete trials. In the real modality, participants dedicated a significantly larger percentage of time to reading the descriptions of the rules in Task 2, and spent considerably more time in all levels of Task 3. Despite this, there were no significant differences in the number of simulation runs. Regarding Task 4 (Figure 5.12), participants seem to be distributed similarly across each theme within both modalities.

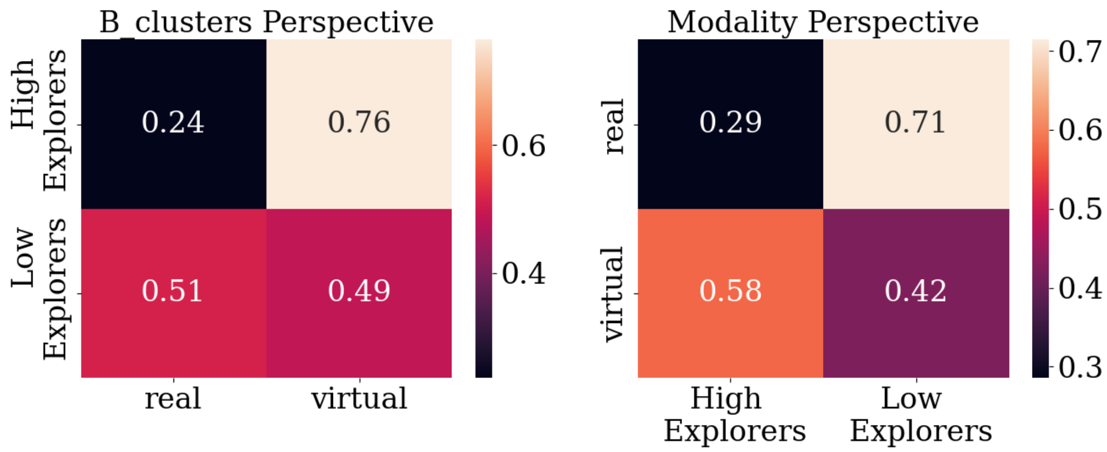


Figure 5.19: Heatmap of the relative frequency of (left) the modality groups with respect to the behavioral clusters (B_clusters) and (right) the behavioral clusters groups with respect to modalities.

5.7.3 Conclusion and Discussion

What is the effect of the modality on participants' within activity behaviors? → A difference exists in behavioral profiles with respect to the learning medium (real vs virtual modality).

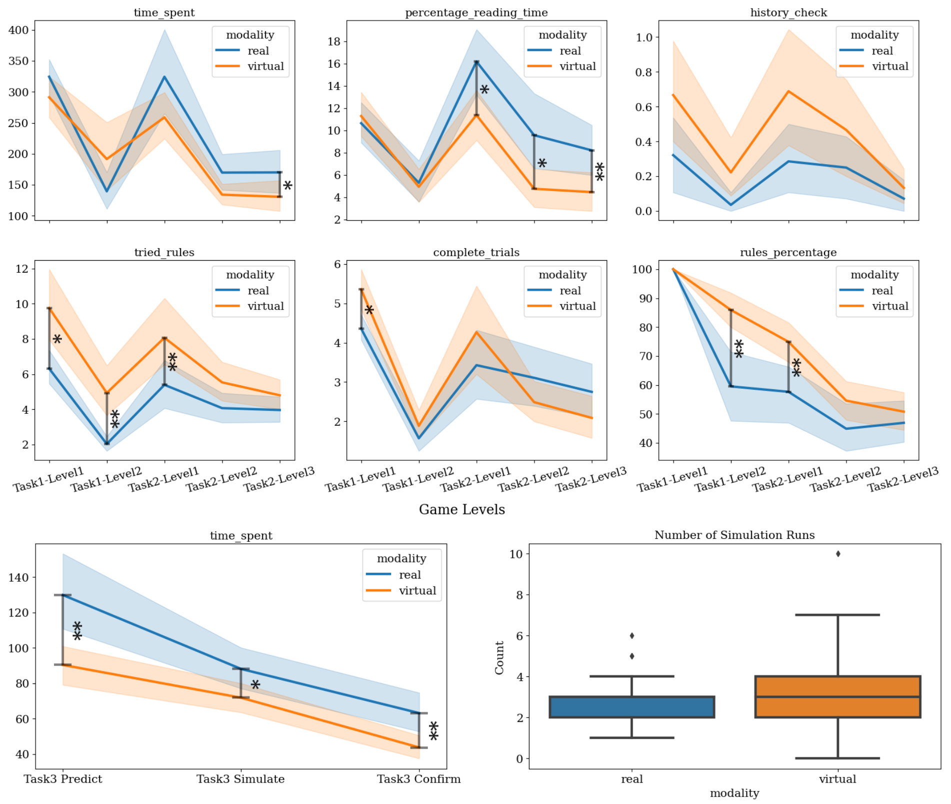


Figure 5.20: Behavioral differences between the modalities. *(**) denotes statistically significant differences with p-value < 0.05 (0.005) with a Mann-Whitney test.

Overall, higher exploration was observed in the virtual modality, and higher reading time was observed in the real modality.

The findings of this study differ from prior work in which tangibility actually fostered exploration (Hengeveld et al., 2013; Schneider et al., 2011). One main difference between our study and prior research is the type of physicality. In previous work, the tangible components were actually static objects (logistics or play-and-learn system for language development), whereas in our activity, the physicality was represented by autonomous robots that were able to move. The higher reading time in the real condition can be explained by the change of the reading medium, as in the virtual condition, it was the computer screen with which students are used to interact, while in the real modality, it was projected text on the table providing a larger and novel environment to participants.

5.8 Key Take-Aways

In this chapter we proved the suitability of the *Cellulan World* game as learning activity to foster complex systems understanding. A mean positive relative learning gain of 11.8 % was observed.

In the game, exploration led to better performance but that was not necessarily correlated to total learning gain. An interplay exists between time spent, exploration, and reflection in participants' behavior, which predicts whether they are learning gainers or non-gainers. Regulation skills depicted in reflection are critical for successful learning.

The learning medium (real vs virtual modality) influences participants' behavior, with higher exploration observed in the virtual modality and higher reading time observed in the real modality. The added-value of physical robots was proved to be an increase in enjoyment and cognitive engagement during the learning activities, reinforcing intrinsic motivation, although this may not necessarily lead to a significant difference in learning gain compared to virtual modalities. However, we observed that participants learned on different scenarios depending on the modality, suggesting a rather complementary relationship between the real and virtual modalities for learning. Further studies are needed to dig deeper into this relationship.

6 Beyond High School: Learning with *Cellulan World* Across Age

6.1 Introduction

As established in the preceding chapters, understanding complex systems is increasingly becoming an essential cognitive skill for the twenty-first century (Reed, 2020).

Research has predominantly concentrated on learning and teaching about complex systems in K-12 science education. In a meta-analysis conducted by Yoon et al. (2018), the authors compiled 75 empirical studies focusing on student learning or understanding of scientific systems that exhibit complex characteristics primarily covering seven content areas: biology, chemistry, computer science, earth science, ecology, physics, and complex systems in general. They found that the majority of the studies (79%) targeted students in a single grade group, with a fairly equal distribution across elementary, middle, and high school grades. A smaller percentage of the studies (16%) included students across grades: from elementary, middle, and/or high school grades. Comparatively, only 11 studies (15%) focused on teachers as target participants, and 7 studies (9%) involved both teachers and students in their samples. In a subsequent study, Yoon et al. (2019) proposed a learning progression of concepts spanning middle school through upper high school grades.

Another commonly considered target participant group for complex systems learning are university students, although the literature in this area is less extensive. In a study employing a hypermedia learning environment with agent-based models to investigate learning about complex systems, Jacobson et al. (2011) had undergraduate university students participants with a mean age of 22.0 years old (SD = 2.21). The participants mainly majored in engineering or science fields, with a few studying business and social sciences. In another study by Rates et al. (2022), 96 undergraduate and graduate students, primarily 19-20 years old (73%) and architecture majors (92%), engaged in an agent-based participatory simulation as part of a mid-level architecture course that emphasized the built environment's function within complex systems, such as the Chesapeake Bay Watershed.

To our knowledge, no study has been conducted across different age levels and with the general population, i.e., individuals who are not necessarily enrolled in formal education. In this chapter, we aim to evaluate the effectiveness of our developed activity, *Cellulan World*, with participants of varying ages from the general population and compare the learning gains and behaviors across different age groups.

More specifically, in this chapter, we investigate the following research questions:

RQ6.1 Is *Cellulan World* effective for participants across different age groups?

RQ6.2 Are there, and if so what are, the age-related differences in learning with *Cellulan World*?

RQ6.3 Are there, and if so what are, the age-related differences in behaviors while engaging with *Cellulan World*?

6.2 Methodology

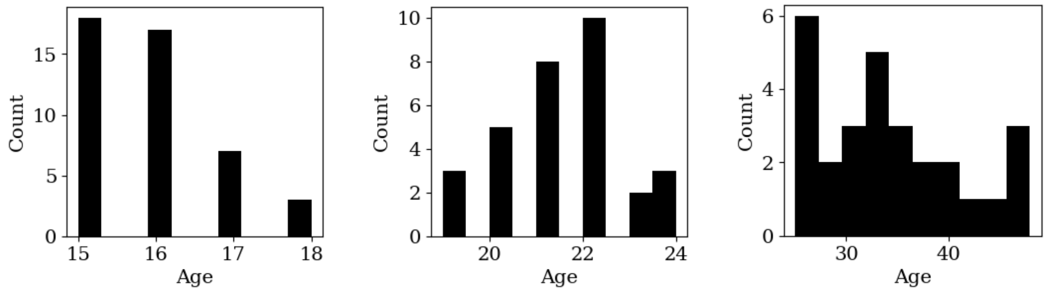
6.2.1 Participants

We included data from two main sources. The first source was our previous study on Cellulan World (Chapter 5), from which we only used data from participants in the virtual modality. This constituted a group of 45 high school students with ages ranging from 15 to 18, which we denote as Group 1. For the second source, we conducted an online study using Prolific[1], a platform for recruiting research participants. We recruited a total of 60 new participants, with ages ranging from 18 to 48 years, and compensated them with 6.5 GBP each. To participate, individuals had to be fluent in English. The data collection was done in two batches: the first batch included 30 participants aged 18 to 24, which we denote as Group 2, as this is normally the age range for individuals pursuing university degrees. The second batch included 30 participants aged 25 and older, which we denote as Group 3, as people in this age range are often already working or have finished their studies. One participant did not complete the post-test, resulting in a final sample size of 104 participants. Figure 6.1 displays the age distribution of the participants in each of the three groups. Figure 6.2 shows the distribution of participants across different levels of education and grades (if in high school). Finally, Figure 6.3 displays the distribution of majors among the participants, classified according to the UNESCO International Standard Classification of Education (ISCED)[2].

6.2.2 Procedure

We follow the same procedure described in Section 5.2.1 of Chapter 5: Welcome, Pre-questionnaire, Pre-test, Learning Activity, Post-questionnaire, Post-test, and Goodbye. The entire process is orchestrated using Google Forms, which provides the link to the activity at the appropriate

[1] www.prolific.co
[2] http://uis.unesco.org/en/topic/international-standard-classification-education-isced



(a) Group 1 (b) Group 2 (c) Group 3

Figure 6.1: Age distribution by group. This histogram shows the frequency distribution of participants in each age bin for each of the three groups: high school students (Group 1), university students (Group 2), and participants aged 25 and older (Group 3).

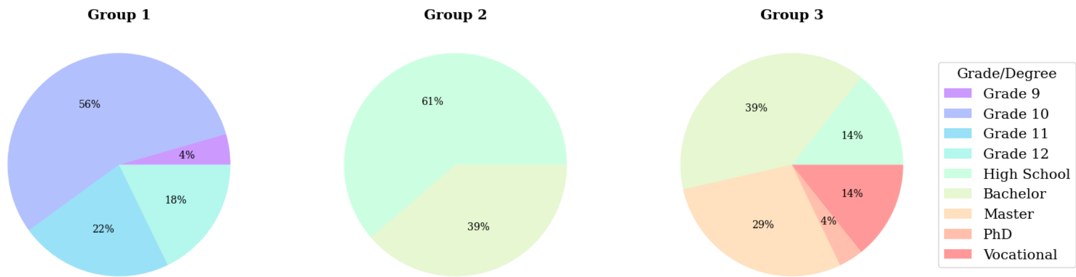


Figure 6.2: Distribution of participants by education level (highest degree obtained) and grade (for high school students).

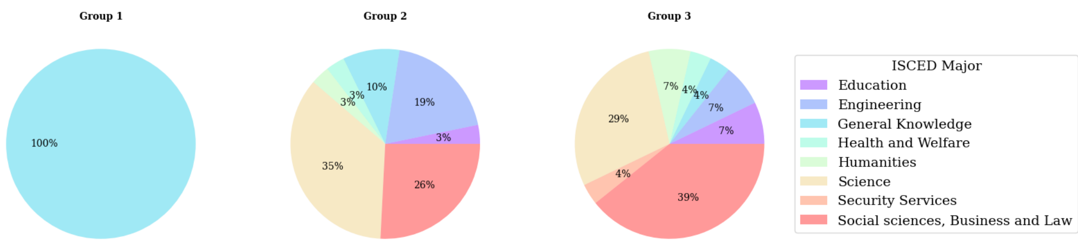
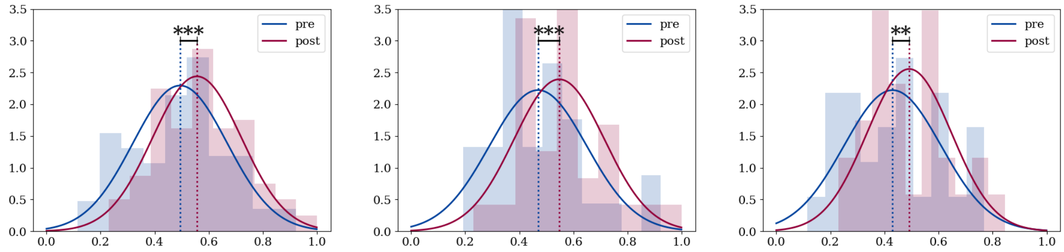


Figure 6.3: Distribution of majors by Group as classified by the UNESCO International Standard Classification of Education (ISCED).



(a) Total (b) Group 2 (c) Group 3

Figure 6.4: Pre- and post-test scores for (a) the total number of participants, (b) Group2, and (c) Group3. Group1 was already analyzed in Chapter 5. ** (***) denotes statistically significant differences with p-value < 0.01 (0.001) with a paired sample T-test.

moment, specifically upon the submission of the pre-test. Subsequently, participants are directed from the activity back to a Google Form to complete the post-questionnaires.

In this context, the Welcome stage is presented as an online text offering explanations. The consent form is also electronically signed/approved. The Goodbye stage features a thank-you message at the end, providing participants with the opportunity to leave comments if they wish. Upon completion, they are redirected back to Prolific with a unique completion code, which entitles them to receive the agreed-upon compensation.

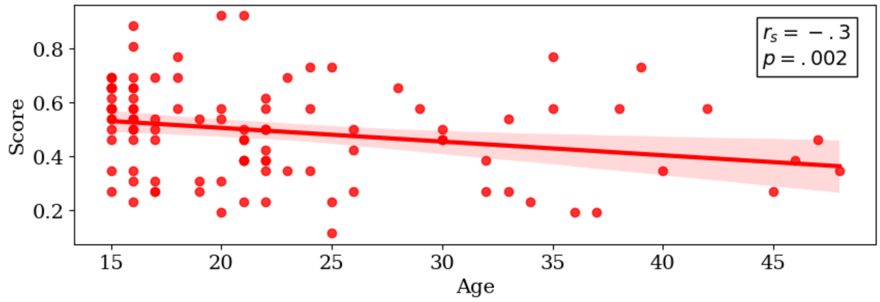
6.3 Results

We adopted a similar analysis method as described in Chapter 5, utilizing the same data extraction and collection procedures. In the graphs below, a red color graph indicates a significant result with p-value < 0.05 with Spearman correlation, a grey color indicates no significant results, and a pink color indicates a marginally significant result with p = 0.05.

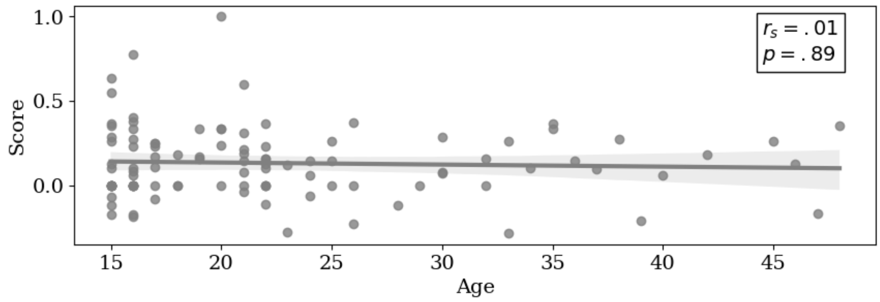
6.3.1 Effectiveness of *Cellulan World*

To address the first research question, we compared the pre- and post-test scores for (1) the group of all the 104 participants, and (2) for each of the subgroups 1, 2 and 3 (Figure 6.4). The analysis of Group 1, which is reported in Chapter 5, reveals a significant learning gain. Similarly, we observe a significant increase in scores from pre-test to post-test for the combined groups as well as within each group. The average relative learning gain is 13.1%. Among the 104 participants, 64 (62%) demonstrated a positive learning gain, indicating a shift towards a more complex systems mental model. In contrast, 25 (24%) maintained their previous score with no gain, and 15 (14%) experienced a decrease in learning, reflecting a shift towards a rather more clockwork mental model.

Conclusion: Is *Cellulan World* effective for participants across different age groups? → Yes!



(a) Pre-Test Scores



(b) Relative Learning Gain

Figure 6.5: Correlation: Scores vs Age for (a) pre-test, (b) total relative learning gain. r_s denotes Spearman correlation coefficient and p its p-value. A red color indicates a significant result with p-value < 0.05 with Spearman correlation, a grey color indicates no significant results.

6.3.2 Differences in Learning

The second research question is addressed by comparing learning across different ages. To achieve this, correlations between age and test scores, including pre-test, total relative learning gain, and relative gain for each concept and scenario, are examined using Spearman correlation, a non-parametric method that is robust to outliers and skewed data.

The analysis reveals a significant negative correlation only between age and pre-test scores, indicating a more centralized mindset as age increases (Figure 6.5a). However, no significant correlation is found for the relative learning gain (Figure 6.5b), a result consistent across the total learning gain and scores on individual concepts and scenarios.

Additionally, we perform the same correlation analysis between participants' reported grade/degree levels and their learning gains, observing similar results as with age correlations. In this analysis, higher education levels are simulated as higher grades for consistency.

Lastly, we compare the scores among different groups based on their ISCED majors. We conduct pairwise non-parametric tests (Mann-Whitney) corrected for multiple comparisons, revealing no significant differences between the groups for overall test scores, including pre-test, total relative learning gain, and learning gain across concepts and questions.

Conclusion: Are there, and if so what are, the age-related differences in learning with Cellulan World? → No significant correlation was found between age and the learning gain. Although our sample size consists of 104 participants, the data is skewed towards lower ages, warranting more extensive experimentation to derive more meaningful correlations between age and learning. Nevertheless, our findings suggest a promising outcome of positive learning gain observed across all age groups using our learning activity.

6.3.3 Differences in Behaviors

With respect to the third research question to investigate age-related differences in behaviors in the activity, we followed the same strategy and studied the correlation between age and the behavioral metrics discussed previously in Chapter 5. Figure 6.6 shows the correlation between age and time spent in each of the experimental stages, revealing a significant weak positive correlation between age and time spent in doing the pre-test and tutorial level ($r = 0.26$), and a weak negative significant correlation in time spent in Task 2 ($r = -0.22$), indicating that older people spent more time in the pre-test and tutorial and less time in Task 2.

Figure 6.7 shows the correlation with the remaining behavioral metrics within each task, averaged across levels. In Tasks 1 and 2, we see a significant negative correlation, although weak ($r = -0.27$), between age and tried rules as well as the percentage of rules tried. In Task 1, this correlation is also significant for the complete trials metric. Overall, these findings indicate that younger people tend to explore more in the tasks. In Task 2, a significant weak positive correlation ($r = 0.25$), also exists in the reading time percentage. In Task 3 (Figure 6.7c), no significant correlation between age and behavioral metrics was found.

Finally, we examined the correlation between the engagement metrics and age, and we found a marginally significant positive correlation on the affective engagement metric (Figure 6.8). Overall, the scores were high, with an average of 4 on a 5-scale.

Conclusion: Are there, and if so what are, the age-related differences in behaviors while engaging with *Cellulan World*? → The observed differences include: Younger individuals tend to explore more. Older individuals tend to read more. Older people marginally enjoyed the activity more.

6.4 Case Studies

Examining the qualitative nature of the behaviors and test responses of a cohort of participants with similar learning gains enables us to understand the underlying learning process. In our study, we selected 2 participants with highest learning gains from each group and analyzed their pre- and post-test answers. Additionally, we analyzed the responses of the participants with negative learning gains to gain a better understanding of what happened in their particular case.

6.4.1 Participants with Highest Learning Gains

When examining participants with high learning gains, we observed that 5 out of the 6 selected participants incorporated the game into their answers of the post-test. For example, participant x from Group 2 transitioned from a centralized answer in the pretest for the scenario Robots and Gold, where the given strategy focused heavily on the spaceship as the hub for communication and target assignment, to a more decentralized approach inspired by the game's concept of spreading by repelling and using local information. This shift in strategy is reflected in their post-test response, as shown in the quotes provided in Table 6.1. Similarly, another participant y from Group 3 drew inspiration from the game and incorporated it into their response for the scenario Scatter, stating, *"I am moving like Cellulans while doing the rule REPEL."* These examples illustrate how the game influenced participants' thinking and strategy development throughout the study and suggests that the participants were able to transfer their understanding from the game context to another near context, namely, Robots and Gold and Scatter scenarios. Additionally, in the response to the Butterfly Effect scenario(one of the far transfer scenarios), a 15-year-old participant w from Group 1 cited examples from the game to illustrate how a small action can have a significant impact: *"In the cellulan world game if one of the robots were broken or the number of them increased, it was only a small change but a change that required a new strategy to solve the task otherwise it would be unaffective"*.

To illustrate the actual observed behaviors in the game, an example of the participant x is presented in Figure 6.9a. The figure shows the game timeline depicting the different rules that

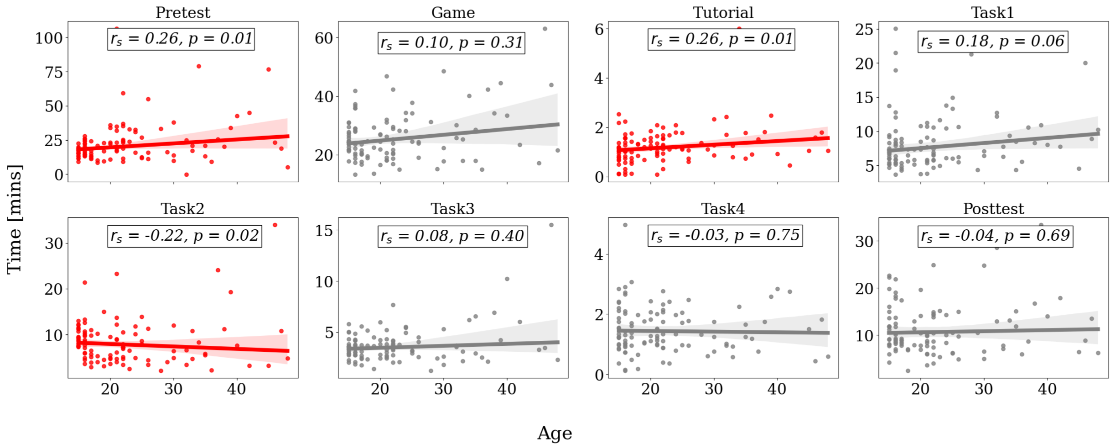
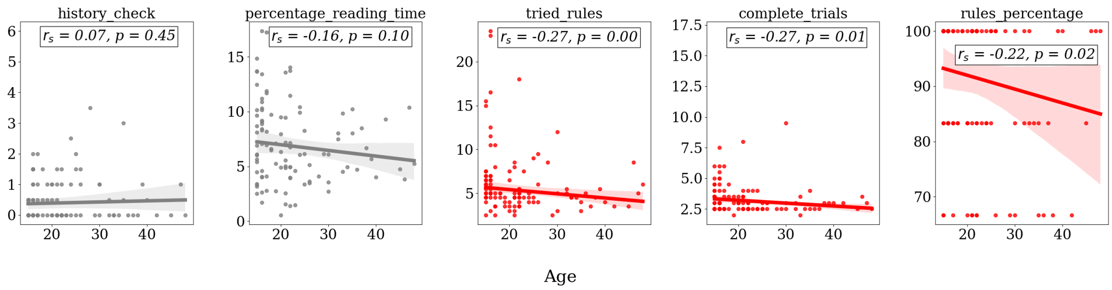
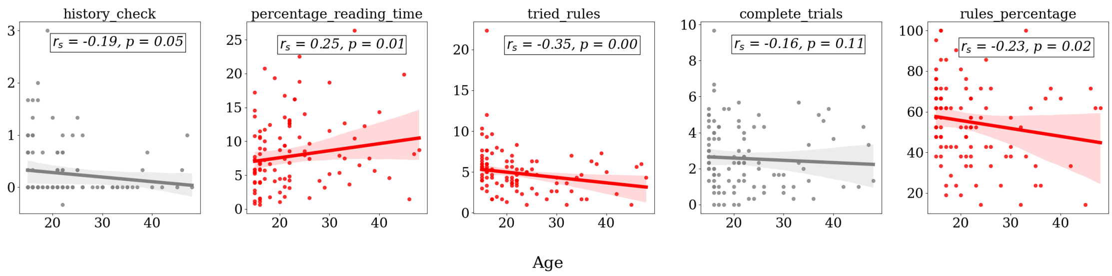


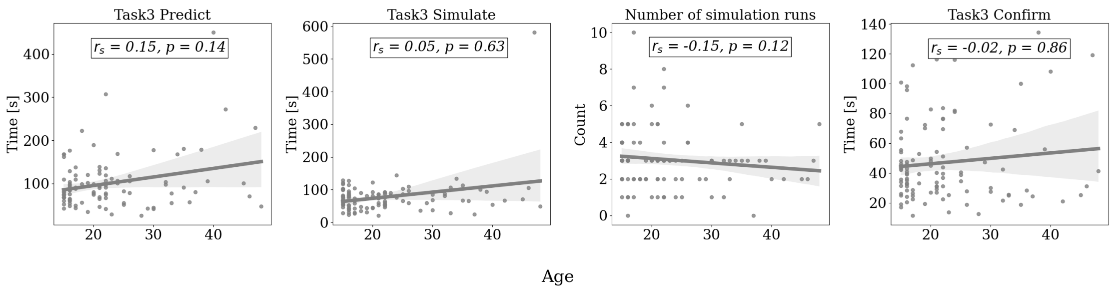
Figure 6.6: Correlation: Age vs Time of each task, as well as pre-test and post-test



(a) Task 1



(b) Task 2



(c) Task 3

Figure 6.7: Correlation: Age vs behavioral metrics in each of the Tasks 1,2 and 3. A red color indicates a significant result with p-value < 0.05 with Spearman correlation, a grey color indicates no significant results.

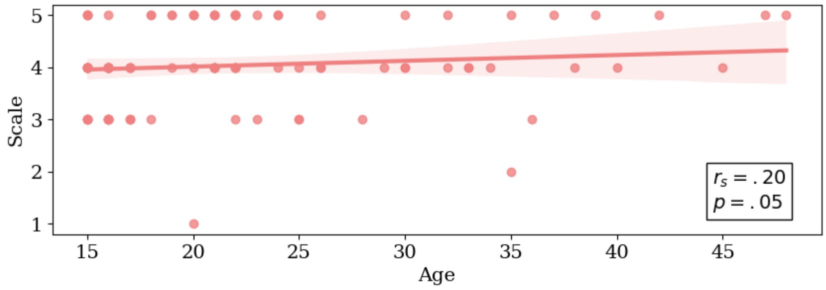


Figure 6.8: Correlation: Age vs Affective Engagement metric

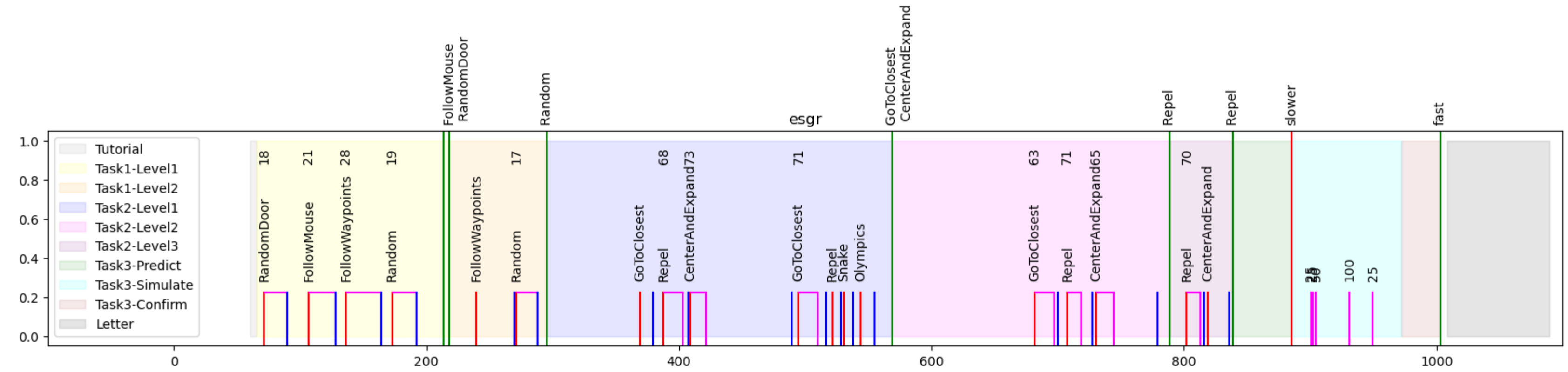
Question	**Pre**	**Post**
How should the robots move in order to find the gold?	Spaceship assigns a target area for each of the robots	*Spread according to a determinate range*
Elaborate on your choice.	*If you assign areas, you ´re covering more land and you're increasing your chances of getting gold*	*The Cellulians game, thats how they covered more land*
How would a robot inform the others if it found gold?	*Reports back to the spaceship*	*Communicate with nearby robots and inform them*
Elaborate on your choice.	*It is dangerous to let robots communicate among themselves*	*This way it is faster to communicate*

Table 6.1: Table showing the pre and post answers of participant x on the questions related to the Robots and Gold scenario (total score = 0 on pre-test and = 1 on post-test.

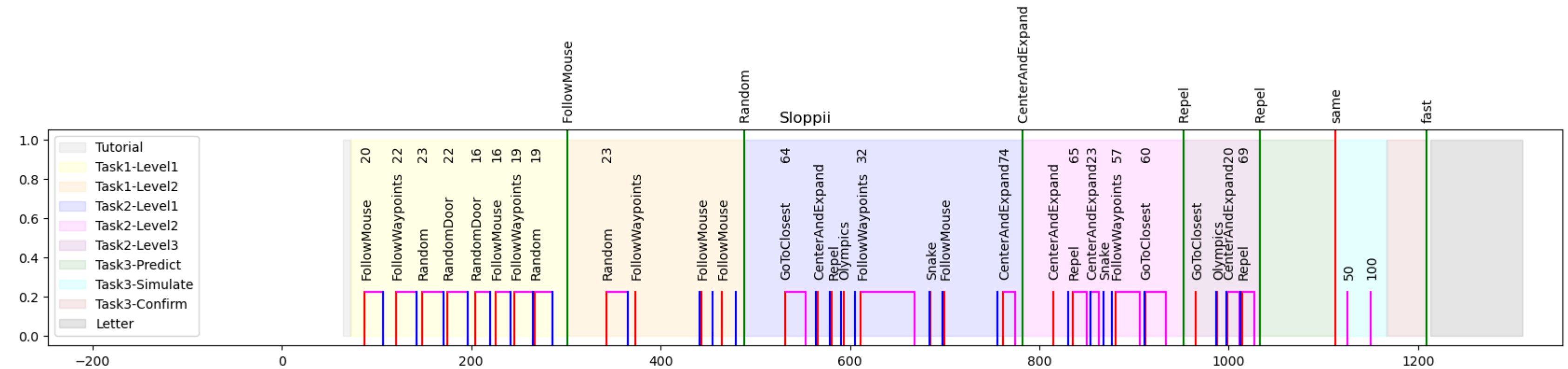
were tried, the scores that were obtained, and the rules that were ultimately chosen as the most effective. Notably, we can observe a transition from a centralized strategy to a decentralized approach as the levels progressed in Tasks 1 and 2. For instance, in Task 1, we observe a shift from a centralized rule (follow mouse) to a decentralized one (random), in Task 2, we see a transition from using rules such as Center and Expand and Go to closest to Repel, and in Task 3, a shift in answers from slower to faster in Task 3. These observed behaviors align with the theory of conceptual change and POE, which we discussed in Chapter 4, and illustrate how the game effectively promoted these learning strategies among participants.

6.4.2 Participants with Lowest Learning Gains

However, we also observed negative learning gains among some individuals. As case studies, we select the lowest 2 participants from each of the three groups. For those 6 participants, we noticed two categories of changes in their answers. The first category (4 out of 6) involved changes in responses to the scenarios Traffic Jam and Flock of Birds. For example, participant z from Group 3 switched from "Yes, a traffic jam can still form" to "No, there is no way for a traffic jam to form," while participant q from Group 2 changed their answer to the scenario Flock of Birds from "*I think there isn't one leader when flying. They try to maintain the same*



(a) participant transitioning from centralized to decentralized rules and demonstrating positive learning gain



(b) participant transitioning from centralized to decentralized rules but demonstrating negative learning gain

Figure 6.9: Timeline of Game Behaviors. The red lines indicate when a rule is applied by the user, with the corresponding rule names labeled above. Blue lines indicate when a rule is stopped, while magenta lines mark the completion of a rule, with the achieved scores noted adjacent to these lines. The green lines signify when a rule is submitted by being the best, with the rule names also provided. Colored boxes in the background represent various tasks, as detailed in the legend. The x-axis represents time in seconds.

speed to keep the group together." to "*I first said the average speed, but now I think that there might be a leader, someone who tells the route to the others. Otherwise, who would stop and start first?*"

The second category of changes involved participants switching to more centralized choices and providing examples from the game in scenario Robots and Gold. For instance, participant *t* from Group 1 stated, "*Spaceship assigns a target area for each of the robots. It is similar to the part where you would place waypoints or the robot would naturally go to the assigned spot that worked well because it has already been calculated to take up the most space.*" Similarly, participant *a* from Group 2 also chose that the spaceship assigns a target area for each of the robots, saying, "*Biggest spread, I assume that the spaceship is a bit faster in commanding robots than I was with the Cellulans :).*"

These observations highlight that this group of people referred to the game experience (similar, although in smaller percentage, to the high learners mentioned earlier), but derived the opposite lesson. Intrigued by these findings, we examined the behaviors of these participants (as shown in Figure 6.9b). Interestingly, we found that even though they chose the decentralized option as the best rule in the game, their post-test responses was in the opposite direction, suggesting they weren't able to transfer the game scenario to the test scenarios.

6.5 Discussion and Conclusion

The study in this chapter aimed to investigate the effectiveness of the *Cellulan World* learning activity for different ages and explore the differences in learning and behaviors across different age groups. The results indicate that the learning activity can be effective at all ages, with a significant increase in scores from pre-test to post-test for participants of all age groups. Furthermore, the average positive relative learning gain of 13.1% observed across all age groups is a promising outcome for the effectiveness of the designed learning activity for individuals of varying ages, including those outside of formal education settings.

Regarding the differences in learning, the correlation analysis shows no significant correlation between age and relative learning gain, indicating that the learning activity is equally effective regardless of age. Additionally, the analysis of grade/degree levels reveals similar results as the age correlations. The uniform effectiveness of the learning activity across all age demographics suggests that it may rely on factors such as conceptual change that are universally applicable, regardless of age or initial mindset.

On the other hand, there is a significant negative correlation between age and pre-test scores, which points towards a tendency for older participants to have a more centralized mindset. This trend may be due to their educational background, shaped during an era when the centralized mindset viewpoint was more prevalent. Additional research is required to provide a more comprehensive understanding of this correlation.

The study also explores the differences in behaviors among different age groups, revealing that younger individuals tend to explore more in Tasks 1 and 2, as demonstrated by the significant negative correlation between age and the percentage of rules tried and the number of rules tried in those tasks. Older individuals tend to read more as demonstrated by the significant positive correlation between age and the reading time percentage in Task 2. In contrast, no significant correlation was found in Task 3.

Our qualitative analysis provided valuable insights into the learning process of the participants in our study. Through the analysis, we observed participants with the highest learning gain demonstrated a shift from a centralized strategy to a more decentralized approach as they progressed through the tasks and transferred the game ideas into their answers on the post-test. Conversely, a few individuals demonstrated negative learning gains who switched to more centralized choices during the post-test. Interestingly, participants with negative learning gains also transferred game ideas into their answers, albeit at a smaller percentage than those with positive learning gains. This suggests that there is something deeper going on regarding how the game affects the mindset of people and how they interact with it which is worth exploring further in future studies.

A limitation of our study is the skewed age distribution of our participants, with fewer data points for older individuals and a wider range of majors, backgrounds and experiences among the older population. Future studies could aim to address this limitation by recruiting more and more diverse participants across different age groups and disciplines to be able to extract stronger correlations.

Overall, in this Chapter, we demonstrated the *Cellulan World* learning activity as being an example of an activity that fosters complex systems understanding possibly outside of traditional educational settings. Learners of all ages can benefit from the activity as it helps develop an essential skill across disciplines and across age levels.

Summary

The research questions driving Part I centered on evaluating understanding of complex systems, facilitating conceptual change, examining the advantages of physical swarm interactions, and assessing the effectiveness of the learning activity across different age groups. To address these questions, we utilized a variety of approaches, including creating an evaluation tool, iteratively designing the intervention, and performing empirical investigations.

Chapter 3 proposes an instrument for assessing a person's expertise in complex systems, which has been validated through a study with 37 participants, including experts and non-experts. The instrument relies on five scenarios and a coding scheme to grade a person's competence in five ontological concepts essential for complex systems understanding. An automatic approach for grading was also developed. The chapter concludes that the instrument, coding scheme, and automatic scoring approach advance our knowledge of assessing people's understanding of complex systems. In subsequent chapters, this assessment instrument will serve as a pre- and post-test to investigate the effectiveness of the designed activities.

Chapter 4 describes the design and implementation of a novel learning activity integrating both virtual and physical agents using the Cellulo platform and aiming to foster complex systems understanding. The learning activity is developed as a game, and it draws from design principles based on conceptual change learning theory. The chapter also discusses the extensive iterative refinement phase that the game underwent before the final version was reached. The following chapters build on this groundwork by conducting experiments to assess the effectiveness of the game and analyze users' behaviors with the different modalities supported by the game.

Chapter 5 establishes the suitability of the *Cellulan World* game as a learning activity to foster complex systems understanding. The findings show that the activity leads to a mean positive relative learning gain of 11.8%, which was observed through the game's exploration, reading, and reflection components. The findings also show how the learning medium (real vs virtual modality) influences participants' behavior, with higher exploration observed in the virtual modality. Additionally, the added value of physical robots is observed to be an increase in enjoyment and cognitive engagement during the learning activities, reinforcing intrinsic motivation, although they may not necessarily lead to a significant difference in learning gain compared to virtual modalities. Furthermore, the study shows that participants learn differently, and on different scenarios, depending on the modality, suggesting a rather complementary relationship between the real and virtual modalities for learning. Our analysis suggests that effective learning hinges on the quality and intentionality of learning behaviors. Critical among these is reflection, as indicated by the 'history check' metric underlining the significant role of reflective practice in learning. Additionally, a balance between time spent, active experimentation with different rules, and thorough task completion also contribute to effective learning.

Chapter 6 investigates the effectiveness of the *Cellulan World* learning activity for different ages and explores the differences in learning and behaviors across different age groups. The results indicate that the learning activity is effective at all ages, with a significant increase in scores from pre-test to post-test for participants of all age groups. Furthermore, the average positive relative learning gain of 13.1% observed across all age groups is a promising outcome for the effectiveness of the designed learning activity for individuals of varying ages, including those outside of formal education settings. The study also explores the differences in behaviors among different age groups, revealing that younger individuals tend to explore more, while older individuals tend to read more. The chapter concludes with a discussion of the qualitative analysis, which provides valuable insights into the learning process of the participants in the study, revealing that the game affects the mindset of people and how they interact with it which is worth exploring further in future studies.

Growing the Swarm: Part II
Towards Group Activities

Preface

In part I, we discussed the design of *Cellulan World*, which is an individual activity in a setting where one human interacts with a swarm of robots. In this part, we introduce the idea of a "double swarm" referring to a situation where multiple humans interact with several robots that are also interacting with each other.

The double swarm notion is an extension of the traditional notion of human-swarm interaction, which typically involves a single human interacting with a swarm of robots (Kolling et al., 2016). Previous research in human-robot interaction within educational applications has primarily focused on single-human single-robot systems, and to a lesser extent on multi-human single-robot systems (Sebo et al., 2020). Conversely, multi-human multi-robot systems have received little attention in the literature, similar to the use of a swarm of robots for understanding complex systems as discussed in Chapter 2. Among the few examples, Hwang and Wu (2014) designed a multi-robot system where students controlled robots to move dice, and examined their collaborative strategies and behavioral interactions. The objective was to help students learn how to carry out collaborative tasks. In another social context, Alhafnawi et al. (2022a) examined multi-human multi-robot dynamics in the context of decision-making. They deployed a swarm of robots to help a group of human participants in their decision-making process.The robots, that could move across the group, served as a visual information flow communicating the opinions of other participants.

The upcoming Part II of this thesis, which contains three chapters, examines the potential of group activities involving two or more people and their possible interactions for complex systems understanding.

Chapter 7 details a collaborative classroom activity created to explore the propagation of diseases. This activity's goal is to foster an understanding of the mechanisms behind disease outbreaks, a prime example of the complex systems we experienced during the pandemic. The activity utilizes simulation and visualization of an epidemic with a swarm of robots controlled by a group of students in a classroom setting.

Chapter 8 center on the crucial role of communication in agent-agent interactions, which has a significant impact on collective performance and emergent behaviors. It explores the impact of different communication affordances on the emergence of collaboration strategies in an online game. This chapter also provides insights on how to integrate the game into a learning activity. By experiencing the game at various levels, each with a distinct communication medium, learners can develop a deeper understanding of how communication mediums affect coordination strategies. Based on the available means of communication, learners can decide when to employ centralized or decentralized approaches.

Overall, Part II of this thesis delves into the dynamics of group activities involving two or more individuals and their potential interactions and the possibilities they offer for furthering the understanding of complex systems. The two chapters in this part aim to explore different aspects of human-swarm interactions in group settings.

7 How Disease Spread: Embodied Learning Activity with Cellulo Robots

The material of this chapter is based on the paper: Khodr, H., Brender, J., Kothiyal, A., & Dillenbourg, P. (2021). How diseases spread: embodied learning of emergence with cellulo robots. Workshop proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education, *263–272*

7.1 Introduction

This chapter marks the initial intiative into employing a double swarm setting, specifically within the context of a complex system: the spread of diseases. It details the design and assessment of an embodied learning activity using the Cellulo platform, and intended for high school science classrooms. The purpose of this activity is to enhance understanding of the emergent behavior associated to virus propagation. We tackle two research questions

RQ7.1 Do students learn the concept of virus propagation using a Cellulo-based collaborative embodied learning activity?

RQ7.2 What are the perceptions of students of conducting such an activity in a regular classroom session?

7.2 Learning Activities Design

We focus on the concept of virus propagation and present a learning activity that comprises a virtual component (using a tablet) and a physical component (using real Cellulo robots). Our learning design blends physical and imagined (virtual) manipulation in two distinct, yet interconnected learning activities that model the spread of a disease. The first activity involves tangible Cellulo robots that students can physically manipulate, providing them with a first-hand experience. These robots serve as distributed agents to illustrate the dynamics of virus propagation, or in other words, how agent-agent interaction leads to the exponential spread of the disease. In the second activity, students use a tablet to simulate the behavior of many Cellulo robots and complete a corresponding worksheet where they document their observations and interpret the results. This combination aims to facilitate a comprehensive understanding of virus propagation dynamics. While the physical activity offers a local perspective, focusing on the probability of infection through direct contact with a virus, the virtual activity provides a global view, emphasizing the non-linear aspect of widespread propagation. Past literature supports this combination, suggesting that learning is enhanced when physical and imagined manipulation are used together. Physical manipulation aids in grounding abstract concepts into tangible experiences, while imagined manipulation activates the same sensorimotor processes required to apply these grounded representations in new situations (Glenberg, 2008; Majumdar et al., 2014).

7.2.1 Activity 1: Embodied Variant with Tangible Cellulo Robots

The goal of this physical activity is for students to develop an embodied sense of how the virus propagates and to give the student an intuition of the likelihood of becoming infected through contact with a virus. In this activity, each student is responsible for one Cellulo robot which they can move around the workspace (Figure 7.1). At the start of the game, one Cellulo robot is randomly assigned as infected (the LEDs color is red). As the students move the Cellulo robots, if one infected Cellulo robot touches another robot, this will induce a vibration and a random value between 1 and 6 will be generated (with the aid of the purple LEDs) in accordance to a "probability of propagation". Thus, when an infected Cellulo robot (red) touches a healthy Cellulo robot (green) the student gets a haptic feedback and the healthy Cellulo robot is infected if the value of dice is equal to one (Figure 7.2c), or not infected (purple) for other dice values (Figure 7.2b). The students keep moving their Cellulo robots and observe how they get infected.

7.2.2 Activity 2: Simulation Variant with Virtual Cellulo Robots

The goal of this activity is to help participants understand how quickly or slowly a virus spreads in a population and what parameters impact the rate of spread. Students learn how Cellulo robots propagating a disease may infect other Cellulo robots, often more than once (and then the infected would infect others) and that therefore the number of agents that are being infected at each moment is itself increasing. This is depicted using two representations, the visualization of the simulated Cellulo robots becoming infected (Figure 7.4) and the graph which shows the growth of the infected Cellulos (Figure 7.3b). The multiple representations help to make learners aware of exponential growth and not linear spread. A brief overview of the activity can be seen in Figures 7.3 and 7.4.

The first screen (Figure 7.3a) is a menu page where the user can change attributes of the simulation by modifying the sliders in "Options". In this case, they can change the total

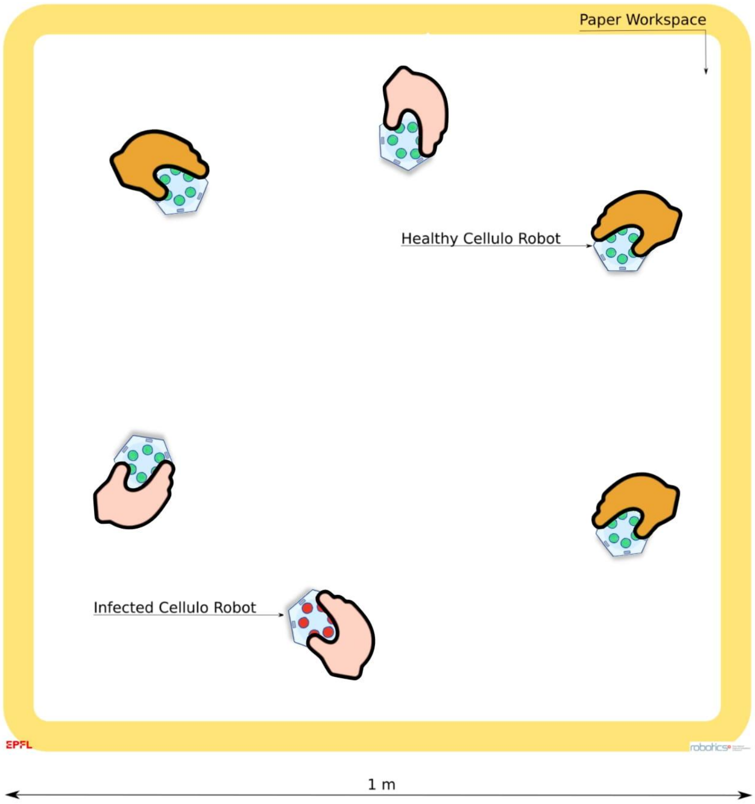
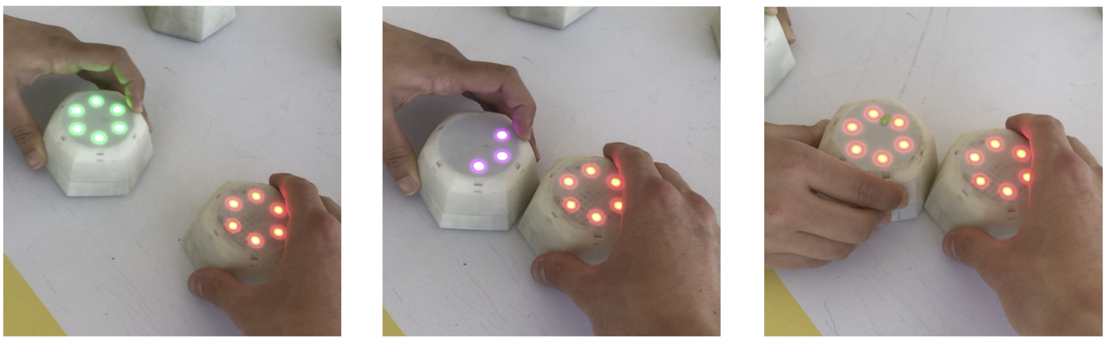


Figure 7.1: The set-up of the physical activity

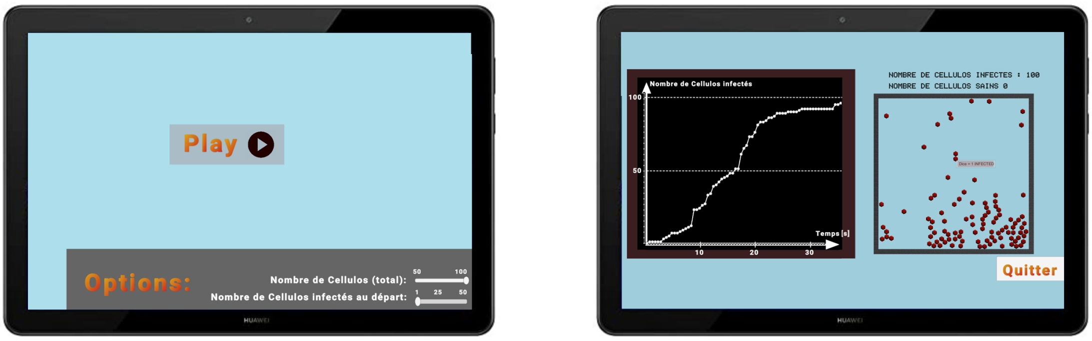


(a) Not touched.

(b) Touched but not infected (the value of dice = 3).

(c) Touched and infected (the value of dice = 1).

Figure 7.2: Examples of contact interactions between Cellulo robots



(a) Menu of the simulation with the setting put as default.

(b) 40-second time simulation.

Figure 7.3: The simulation with the virtual Cellulo robots.

number of Cellulo robots on the map or the initial number of infected Cellulo robots. When the play button is pressed, a corresponding simulation is launched (Figure 7.3b) and visualized with moving virtual Cellulo robots in a containment box along with a graph showing the evolution number of infected robots at each tick of the simulation's clock. This graph serves as a basis for students to discuss modeling the spread of the disease among the Cellulo robots and to construct models in general. They should observe an "S-curve" as the graph in Figure 7.3b depicts. The students are provided with a worksheet (ref. Appendix B) with instructions to vary the parameters, make predictions of the effect of these changes on the graphs and verify their hypotheses after interacting with the simulation. As students change the total number of Cellulo robots, they increase the number of agents per area unit, making it more difficult to avoid infection and thus more infectious interactions per time unit are seen. The graph shows this phenomenon and helps students have a better understanding of why a faster curve is observed on the graph when increasing the number of Cellulo robots.

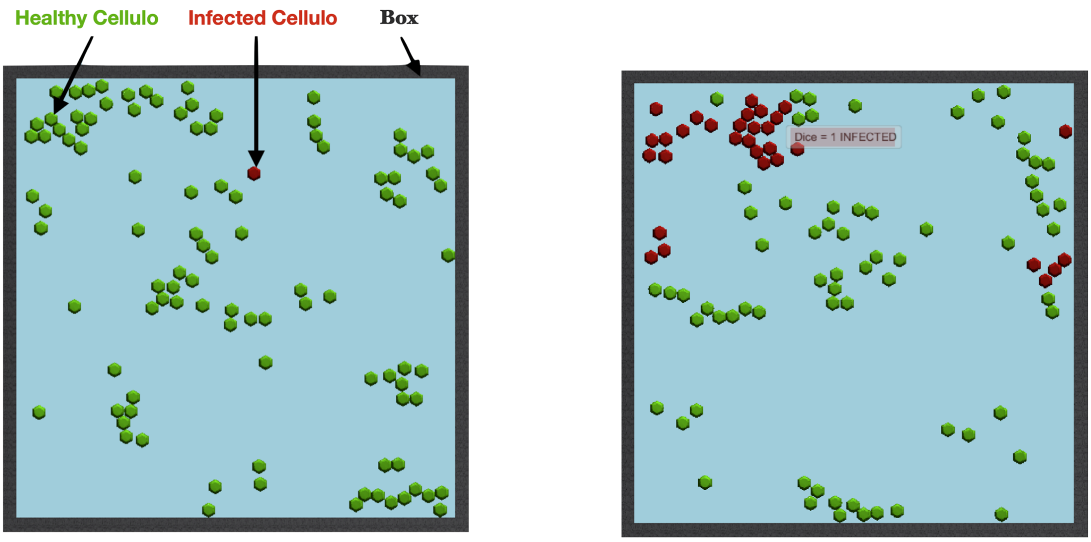
7.3 Methods and Materials

7.3.1 Participants

The study was conducted with one class of 9th grade students (13 years old) at a Swiss school. The experiment lasted 70 minutes and was run with a total of 23 students (12 male, 11 female) with no prior experience with Cellulo robots. The children were split into 6 groups of 3 or 4 students.

7.3.2 Data Sources and Analysis

To answer our RQ1, we collected participants' scores on the pre and post tests to measure their learning gain and to answer our RQ2, we collected responses to a questionnaire to measure their perception of their learning and engagement with the activity. Further in order



(a) Simulation for time = 0 second.

(b) Simulation for time = 10 seconds.

Figure 7.4: Virtual interactions between the simulation Cellulo robots as time progresses.

to understand how learners used the Cellulo robots in the physical activity, we collected logs of the robots' movements. Since at the time of this study, our assessment tool presented in Chapter 3 was not validated yet, the pre and post tests were specifically designed for the purpose of this study and only focused on the the topic of "propagation". The pretest (ref. to Appendix B) has three questions, adapted from literature (Wilensky & Abrahamson, 2006):

1. A matching question, where students were asked to associate a specific context with the corresponding graph. The context options included scenarios such as doubling the number of students every day, adding one student each day, removing one student daily, or one student having a cold and spreading it through the class. The students could select from four types of curves: linear (positive slope), linear (negative slope), exponential, and logistic.
2. A multiple-choice question, which asked students to choose the correct answer in response to a hypothetical situation. This situation revolved around the time it would take for a rumor to spread among all students under two conditions: when the number of students doubles or if the rumor is already known by half of the class. The students had to determine whether the rumor-spreading time would be faster, unchanged, or slower.
3. A graph-drawing question, where students were asked to graphically represent the context of mold propagation on a slice of bread.

The posttest (ref. to Appendix B) is similar to the pretest with a change of values, orders, and context of the questions.

At the end of the experiment, students were administered a final perception questionnaire (ref. to Appendix B) to evaluate their perception of the activities' with Cellulo from the perspectives of interest, collaboration, confidence (in solving the exercises), engagement and future interest (for including Cellulo in future mathematics lessons, as well as lessons of other disciplines). Each of these items (except the construct "interest") corresponded to a minimum of 2 questions, to acquire a more reliable estimate of the construct from the students. Internal consistency is calculated using Cronbach's alpha.

7.3.3 Protocol

We conducted a single group experimental study to assess the efficacy of the designed activity in teaching the concepts of virus propagation. Given the preliminary nature of this feasibility study and the limited pool of participants, we opted for a single pre-experimental study design. Our primary goal was to determine whether our approach had any effect whatsoever, rather than to compare its efficacy with alternative methods.

After a brief introduction and instruction session, students took the pretest which lasted for 10 minutes. Then, they were split into groups. While two groups were doing Activity 1 (physical) in a combined bigger group, four groups were each doing Activity 2 (virtual) on separate tablets. A rotation of the groups was done so everyone could experience both activities. Given the classroom scenario, we could not control the order to the physical and virtual activities. After that, a post-test was conducted. Finally, a debriefing and conclusion session was done and the students filled the perception questionnaire.

7.4 Results

7.4.1 Learning Gain

We calculate the learning gain as the percentage difference between pre- and post-test scores. On the pretest, the students obtained a medium score suggesting that the test is probably well designed for the level of these students and that no ceiling effect occurs. We found an increase from 57.42 (+- 15.14)% on the pre-test to 84.28 (+- 12.57)% on the post-test. Although one student had a negative learning gain of -7.1%, all the other students had a positive learning gain on average of 26.87% (+- 16.42)%. A t-test revealed that there was a significant difference between the pre and post test scores ($p = 1.57e\text{-}7 < 0.05$). This implies that students showed a significant improvement in their understanding of concepts of propagation (Figure 7.5).

7.4.2 Learner Perceptions

Figure 7.6 shows the results of the perception survey. Only 8 participants responded to the survey because the class had ended when the survey was given and the students were free to leave. However, the overall perception of the activity is good, especially for the future interest with 87.5 % of the students giving the maximum rating. Still, the perception of collaboration is

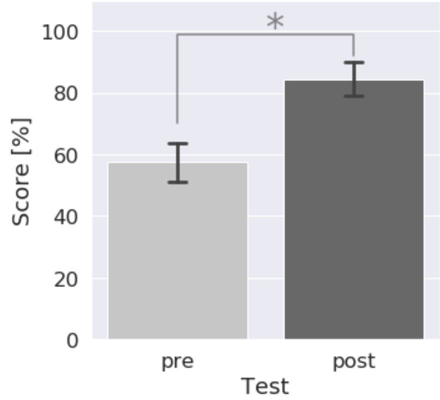


Figure 7.5: Learning Gain Results. * indicates a significant difference with p-value < 0.05 with t-test.

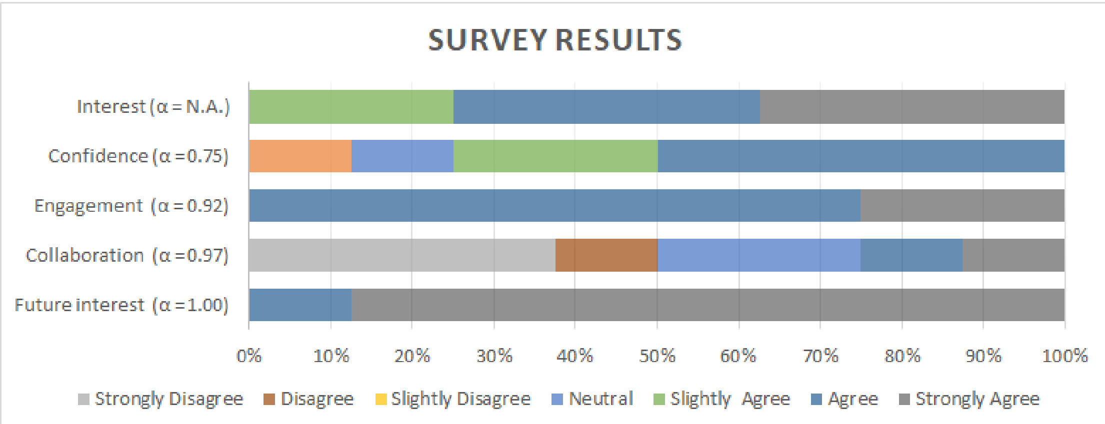


Figure 7.6: Student perceptions of the learning activity.

not very positive. At first glance, this finding could be surprising because the simulation is done with a group of 4 and the activity with Cellulo robots is done with 2 groups of 6-8 people (i.e collaboration is needed). However, the school context could be the reason for this bias where all day long, students have to be silent and the discussions between classmates are not welcome during the class. This could be the reason why the students perceive collaboration negatively.

General Feedback from teachers and students

Informal conversations with the teacher and some of the students at the end of the experiment offered further insights into the perceptions of the participating teacher and the students. The teacher appreciated the activity as a playful and motivating way to learn about virus propagation and as they reported (translated from French) "Students were interested in class, they worked well and participated in the activities, a positive welcome. Is it possible to get the software used, or where is possible to get the activity to try to do another activity?" He also added that "It would be interesting to try with other chapters (e.g. science with biology, astronomy, etc)." However, he expressed the following challenges in conducting the activity, "A little bit too much noise during the course, the students were overexcited and they perhaps 'played' too much with the Cellulo robots (instead of learning)". Thus we infer that the teacher appreciated the learning value of the activity, but admitted that there was an orchestration challenge involved. The students showed their enthusiasm for these activities, as seen by this sample quote from one of the students: "Incredible, fantastic come back [to the experimenter] when it's possible, I really enjoy the Cellulo." However, some students also expressed frustrations about the experience such as 1) "It is not cool to do the test at the beginning." 2) "Too many sheets. It is annoying to do the tests." 3) "Accuracy sometimes from the Cellulo robot is not good. And the tests are too easy." Overall, we see that the students enjoyed the experience, expressed an enthusiasm to do more such activities and their frustrations were related to the experimental protocol rather than the learning activities themselves.

7.4.3 Cellulo Movement

This section seeks to examine the problem solving strategies that emerged during the open-ended physical activity using the Cellulo robots. As the students interacted with the robots, we tracked and analyzed the trajectories of all robots on the workspace. The robot trajectories, illustrated in Figure 7.7, plot the positions and states (infected or healthy) of the robots at various times throughout the activity. Healthy robots are marked by a green hexagon, while infected ones are indicated in red. The trails represent the paths taken by the robots in the preceding 10 seconds. We notice a pattern of infected robots pursuing healthy ones, which in turn attempt to evade infection. Discontinuities in the lines, or "jumps", represent instances where students physically relocated their robot in order to catch or escape from others. These findings suggest that students developed intuitive strategies for both spreading the disease (via the infected robots) and avoiding infection (with the healthy robots).

7.5 Discussion and Conclusions

In this chapter, we presented a two-part physical-virtual embodied learning activity designed to teach the concept of virus propagation through a collaborative activity with Cellulo robots. Statistical analysis of the students' scores on the pre and post tests showed that students had a significant improvement in their scores. This shows that the two part embodied learning activity, which combines direct experience and descriptions of experience in verbal and symbolic forms, was effective for students to improve their conceptual understanding of concepts of virus propagation (Schwartz et al., 2005). Analysis of student responses to the perception questionnaire and their informal feedback suggested that students were engaged with the activity and interested in doing other similar activities. The teacher also perceived

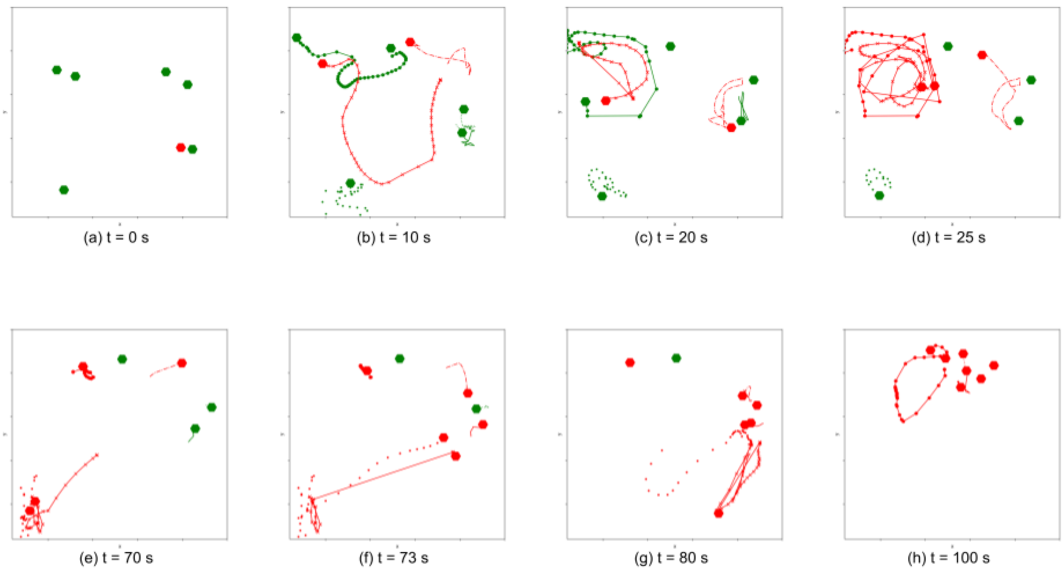


Figure 7.7: Cellulo trajectories during Activity 1.

the usefulness of the activity for learning, while also pointing out the orchestration challenge in implementing the collaborative activity in a classroom. This suggests that it is possible and useful to implement such collaborative complex-systems learning activities in a classroom for a range of topics and domains, albeit logistically challenging.

We also explored the students' trajectories of moving the robots during the physical activity in an attempt to understand how students collaborate while using the Cellulo robots. From students' trajectories we inferred that they performed the activity like a game, with the person holding the infected Cellulo being like the "catcher" and the others holding the infected Cellulo trying to "escape" the catcher. The infected Cellulo moved around trying to "catch" the healthy Cellulo robots, who moved as well in order to avoid being infected. Thus from moving the robots in this manner, getting haptic and visual feedback when they touch an infected Cellulo, and developing intuitive collaborative game strategies students understand certain concepts of emergence. These findings offer us preliminary insights into the mechanisms of collaborative learning via Cellulo robots at a classroom level.

8 Effect of Different Communication Affordances on the Emergence of Collaboration in a Multiplayer Game

The material of this chapter is based on the paper: Khodr, H., Wagner, N., Bruno, B., Kothiyal, A., & Dillenbourg, P. (2022). Effect of different communication affordances on the emergence of collaboration strategies in an online multiplayer game. Swarm Intelligence: 13th International Conference, ANTS 2022, Málaga, Spain, November 2–4, 2022, Proceedings*, 316–323*

8.1 Introduction

In collective systems, multiple agents have to coordinate and collaborate in a flexible and robust manner to achieve a shared goal. The communication between the agents is an essential factor affecting the interaction between them and therefore influencing the collective performance of the group.

As a consequence, communication and its role in collective behaviors have been studied in many different domains and under a variety of perspectives. A first lens for analysis focuses on the *medium of communication*, with verbal (Wang et al., 2013) and haptic (Khodr, Kianzad, et al., 2020) mediums being among the most commonly considered. In a collaborative virtual 2D pointing task between 2 dyads, Wang et al. (2013) compared verbal, haptic, and a combination of both modalities. Their outcomes indicate that participants using verbal only and haptic+verbal communication performed equally well while participants using haptic only communication took more time and had longer path lengths (Wang et al., 2013). Conversely, and quite interestingly, in a study on decision taking, haptically coupled dyads were found to solve a perceptual discrimination task more accurately than their best individual members and five times faster than dyads using verbal communication (Pezzulo et al., 2021). In line with this result, in a collaborative learning activity about linear functions (Khodr, Kianzad, et al., 2020), the haptic channel is found to be used for generating different collaboration strategies such as coordination in movements, or the exchange of leader-follower roles.

A second perspective is the analysis of the *emergence of communication* systems in collaborative tasks. An example of this phenomenon is the emergence of simple language (common code) when participants are involved in a coordination game with no common language made available in the beginning (Selten & Warglien, 2007). Such a behavior requires humans to know that communicative behavior is indeed communicative in nature. Scott-Phillips et al. (2009) investigates this assumption through an ad-hoc designed experimental game. The authors found that the emergence of a communication system usually involves a bootstrapping process, and that this process has an impact on the final form of the communication system. Moreover, one observed necessity for the recognition of signalhood is sufficient common ground, and the emergence of dialogue is seen as the key step in the development of a system that can be employed to attain shared goals. In (Nölle et al., 2020), through a maze game task, the authors present how particular environmental affordances (such as the structure of the mazes) drive the emergence of different communicative conventions in otherwise identical tasks, suggesting that linguistic adaptations are highly sensitive to factors of the shared task environment.

A third key aspect is the analysis of the *effects of communication on team performance*, in organizational contexts. A meta analysis on the topic reveals that communication quality has a significantly stronger relationship with team performance than communication frequency (Marlow et al., 2018). Moreover, classifying communication into commonly measured communication forms shows that information elaboration has the strongest relationship with team performance, while self-report frequency and objective frequency have the weakest relationships. Lastly, although communication can be positively correlated with team performance, the advantages of communication are dependent on the task characteristics as well as the type of communication used (O'Bryan et al., 2020).

Finally, the *scope of communication* is a crucial topic in all domains involving swarms of artificial agents, such as swarm robotics, where the locality of interactions and communication has a beneficial effect on the scalability and robustness of the system, and is thus generally preferred over the use of global communication and sensing (Bayindir & Şahin, 2007). Local communication is then further divided in 1) direct agent-to-agent communication, either using explicit messages or implicitly detecting the existence and relative location of other robots in the immediate vicinity, and 2) stigmergic communication relying on the modification of the environment (e.g pheromones) (Svennebring & Koenig, 2004).

While, as the literature review above highlights, each of the key aspects concerning communication in collective behaviors have been extensively studied, less is known about their interplay. In an effort towards bridging this gap and further expanding our understanding of the role of communication, in this chapter, we consider the above four perspectives together, specifically investigating the effect of different communication affordances (different mediums and scope) on the emergence of communication and collaboration strategies and team performance in an online multiplayer game. Our insights will further aid the design of engaging learning activities that illustrate the relationship between agent-agent communication and the emergence of global behavior. By experiencing varying communication affordances, learners can reflect on different strategies (centralized vs decentralized) based on the available communication affordances.

8.2 Methodology

For our study, we designed an online collaborative game where human participants interact via a robot avatar in a controlled experimental setup. Similar to HuGoS (Coucke et al., 2021), our online environment can capture all the interaction details among participants throughout the game. Three versions of the game, respectively allowing no verbal communication, local verbal (chat-based) communication and global (chat-based) communication among participants were designed, to allow for investigating the effects of communication medium and scope on the emerging communication system and team performance.

8.2.1 Game Design

Game Mechanics

Our online multiplayer game is based on the Unity game engine and the Photon Unity Networking (PUN2) Unity package for multiplayer games (more technical details are presented in Chapter 10).

The game involves 6 players, each represented by an avatar which is chosen to be a virtual Cellulo robot of a unique color and a limited field of view. As an example, Fig. 8.1e shows the green Cellulo robot avatar of a player. The goal of the game is to move boxes to goal positions indicated by red dots as fast as possible. There are a total of 12 boxes in the environment, most of which are not initially visible to players. Some of the boxes can be moved by a player alone (small boxes such as the one shown in Fig. 8.1e) while others require the joint action of 3 (medium boxes, see Fig. 8.1f) or even 6 players (large boxes). Players can control their avatars via the arrow keyboard keys. A player can move a box by:

1. *Pushing the box*: the player's avatar applies a repulsion force on the box when it is inside the halo surrounding the box (see Fig. 8.1f).
2. *Pulling the box*: the player's avatar applies an attraction force on the box when it is inside the halo surrounding the box and the SHIFT key is pressed.

The game is organized in three stages (see Fig. 8.2). In the first stage, each player is in a room alone and must move a small box to its goal position to unlock a door and access the second stage. In the second stage, three players are in the same part of the environment, and they must move together two medium boxes, while the three others have the same task in another part of the environment. In the third stage, all six players are in the same part of the

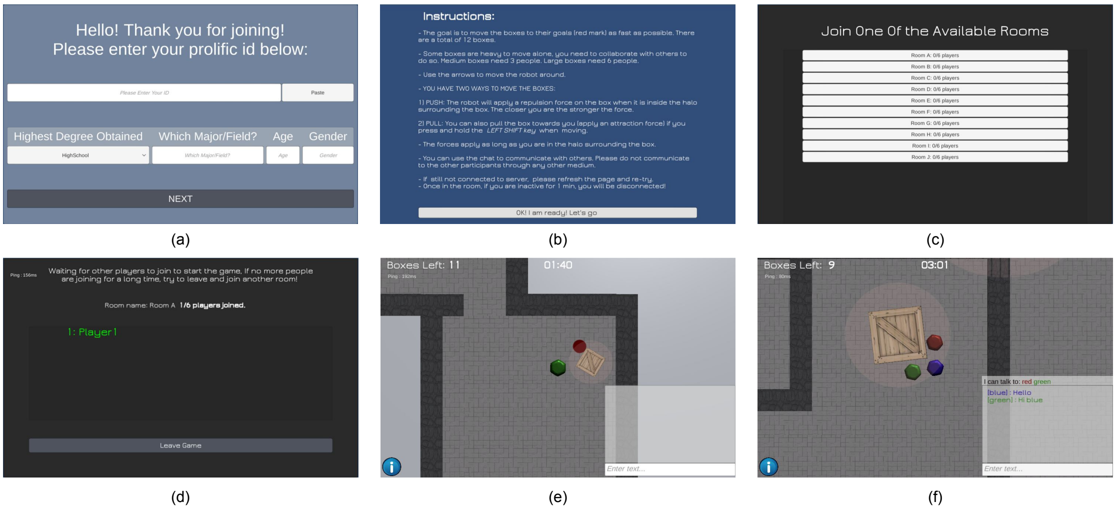


Figure 8.1: Game screenshots. (a) Landing page to enter background info. (b) Instructions page. (c) Game rooms page where the player can select a non-full room to join. (d) Waiting room where players in the same game room wait until the room is full. (e) A screenshot of the first stage with a small box, the green hexagonal Cellulo robot which is the player avatar and the red dot marking the goal position for the box. (f) A screenshot of the second stage with local chat possibilities.

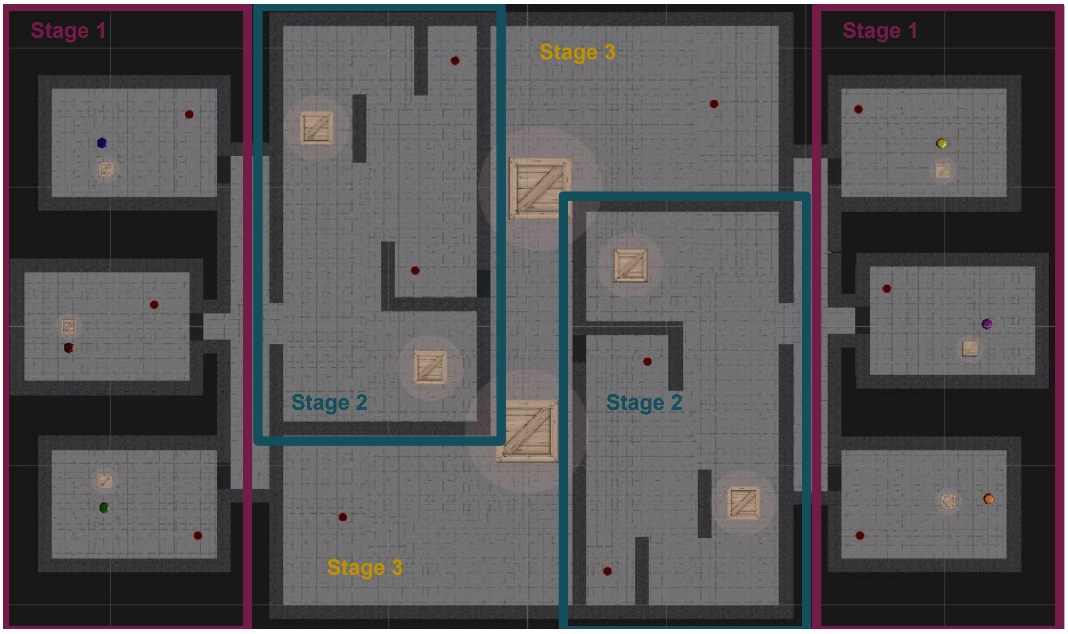


Figure 8.2: A bird view of the game map and the stages division

environment and must move two large boxes to their goals. The players can see the boxes as they move around in the environment. The game ends either when all the boxes have been placed on the goal positions, or a maximum time of 26 minutes is reached.

Game Deployment

The game is deployed on WebGL[1]. The advantage of using WebGL is that all files are hosted on the website, thus allowing players to play the game via their browsers, with no local download or installation needed.

The game data we log include: the positions of all movable objects in the game, the chat messages, the timer, and the background info of the players (age, gender, major and degrees). The data is logged every tenth of a second and uploaded to a Firebase storage[2] every 1 minute. The game is standalone, self-explanatory and can be played without the intervention of an experimenter. Upon clicking on the game link, the player lands on a page asking to provide some background information (Fig.8.1a). Once all information has been provided, the player moves to an instructions page (Fig.8.1b), where it stays at least until the connection to the server is established. Once connected the player can click next and pick one among 10 available game rooms, each displaying the number of players that have already joined it (Fig. 8.1c). A game room is full when 6 players have already joined it. After selecting an available room, the user is moved to a waiting page (Fig. 8.1d) displaying the number of players currently in that game room. As soon as 6 players are concurrently in the same game room, the game is launched automatically, with players being brought to the first stage of the game (Fig. 8.1e). During the game, an INFO button is always present at the bottom-left corner, allowing players to re-read the instructions.

Communications Affordances

We implemented three versions of the game, exclusively differing from one another in terms of the communication affordances provided to the players:

1. *Global Chat:* In this condition, a chat is included in the game. All players can communicate with everyone else by typing and receiving messages.
2. *Local Chat:* In this condition, only the players who are in the neighborhood of the sending player can receive the message. The communication range is set to be 20 % of the dimension of field of view. On the top of the chat box, each player can see with whom they are within range to communicate. As an example, Fig. 8.1f is a screenshot of the blue player, with the chat box showing that they can currently send and receive messages to/from the red and green players.
3. *No Chat:* In this condition, players cannot communicate through chat and so must

[1] and accessible at https://ants-cellulo-game.web.app/
[2] https://firebase.google.com/docs/storage/web/start

discover alternate means of communication to complete the task.

8.2.2 Experiment Design

We designed the study as a between-groups experiment, with the communication affordance as manipulated variable (thus yielding the three conditions described in Section 8.2.1) and team performance (here intended as the time taken to complete the task) as main outcome variable. The study involved 108 participants recruited via Prolific[3]. The participants self-organized in teams of 6 as described in Section 8.2.1 and the teams were split equally and randomly across the three conditions, thus yielding 6 teams per condition. The mean age of all participants was 24.81 years old (SD =5.98) with 39 females (36%), 67 males (62%) and 2 others. A Kruskal Wallis test shows no significant difference for age between the three conditions. We can therefore assume that the three groups are drawn from the same distribution and ignore the existence of a confound based on the age of the participants.

8.3 Analysis and Results

We quantitatively compare the performance of the teams in the three conditions in terms of their time of completion. To complement the statistical analyses correlating chat data with the team performance, we implemented a replay tool allowing us to perform qualitative observation of the players' behavior during the game.

8.3.1 Effects of Communication Affordances on Team Performance

The mean time to complete the task, across all conditions, was 17.2 (+/- 4.5) minutes. The fastest team took 9.6 minutes and the slowest one 25.5 minutes. Although the Local condition has the lowest average time among the three conditions, a Kruskal Wallis test shows no significant difference in performance among the three conditions (df = 2, H= 3.94, p = .14). While representing more than 100 participants, this analysis (which is done at team level), actually only accounts for 6 data points per condition: collecting more data in future studies will thus be crucial to either confirm this result (i.e., the lack of an effect of the communication affordance on performance) or reveal significant differences (e.g., between the local condition and the others).

Our preliminary result is interesting as it contradicts the common-sense hypothesis that having a (global) communication would lead to better performance. Moreover, our preliminary findings are interestingly consistent with the results of another study (Talamali et al., 2021), albeit conducted with robots and not humans. This study found that robot swarms with shorter communication ranges were better able to adapt to changing conditions in the environment compared to those with longer communication ranges. This was in contrast to the widespread

[3]The average reward was set to 6 £/hr. A bonus incentive was given to groups who finish the fastest.

belief that more communication links always improve information exchange on a network, thus enforcing the idea that constrained communication can be beneficial for robot swarms to adapt to changes in dynamic environments. The study suggests that limiting communication to a local neighborhood is a cheap and decentralized solution that allows robot swarms to adapt to previously unknown information that is locally observed by a minority of the robots (Talamali et al., 2021).

8.3.2 Emerging Communication System Analysis - Global Condition

Although we found no effect of communication affordance on team performance, a deeper analysis on the type of communication which emerged in each condition could provide useful insights on the possible causes of this result.

Particularly in the Global condition, a strong positive correlation (Spearman's $\rho = 0.94$, $p = 0.005$) is observed between the total number of messages sent during the game and the time taken to complete the task, indicating that the more messages a team sends, the longer it tends to take to finish the task. However, this correlation might be confounded by the fact that the time required to type and read messages is included in the total task completion time, thus interrupting the time available for task-oriented actions. To further investigate this, we examined the correlation between the number of characters per message sent and the total time, finding no significant correlation. This suggests that the length of individual messages doesn't appear to significantly impact the time taken to complete the task. This could explain that the time taken to type and read a single message, regardless of its length, may be relatively minor compared to the overall task time. Moreover, interestingly, the frequency of communication, defined as the number of messages sent per minute, doesn't show a significant correlation with the total task time. This suggests that the rate at which teams communicate doesn't seem to impact how quickly they complete the task, implying that teams communicating more or less frequently don't necessarily finish faster or slower, respectively.

These findings lead us to infer that (1) simply increasing the volume of chat communication doesn't seem to expedite task completion; in fact, it might even hinder it due to the interruption of the main task. Furthermore, (2) teams seem to engage in chat communication because they have the time to do so, and the frequency or length of individual messages doesn't significantly affect the task performance. In summary, in the context of this game with text-based chat, global communication doesn't appear to enhance the teams' coordination or performance on the task.

To complement the above analysis on the volume of the messages, we also analyzed their content. Fig. 8.3a shows the word cloud of all chats in the Global condition. The two most recurrent words are "need" and "push". The next recurrent words are "one", "left", and "pull". This shows that the players mainly use the chat to ask for help ("need", "one"), or give orders for the actions to be done related to the game ("push","pull") and agree on directions ("left").

(a) Global Condition

(b) Local Condition

Figure 8.3: Wordcloud analysis of the messages shared over chat.

8.3.3 Emerging Communication System Analysis - Local Condition

The total number of messages shared by teams in the Local condition is significantly lower compared to the Global condition (ttest: $t[10] = -2.3$, $p = .04$, Cohen-d = 1.33), thus suggesting that this communication modality, by design, induces a smaller volume of messages to be shared.

Contrary to the Global condition, no significant correlation is found here between the total number of messages sent during the game and the time taken to complete the task. At the same time, no significant correlation is found between chatting frequency and completion time either. Quite interestingly, however, a median split reveals that the 3 top performing teams are also the ones with the smallest overall number of messages sent, as well as the lowest chatting frequency.

These results, somewhat hinting at the importance of sharing "the right information at the right time" motivate us to look into the timing of the messages. By comparing the game events with the chat messages, we notice two trends concerning when a chat is initiated:

1. Prior to acting, e.g. before moving a box or going on a search for the goal position/missing players. Some examples are shown in Table 8.1.
2. In response to certain events, e.g. to help a player or solve a conflict. Some examples are shown is Table 8.2.

On average, teams had 2.3 chats initiated prior to acting and 2.8 initiated in response to an event/conflict, with the difference between the two not being significant.

Concerning the content of the messages, Fig. 8.3b shows the word cloud of all chats in the Local condition.

	Sender	Message	Receiver(s)
Example 1	green	where does this one have to go?	red, blue
	blue	bottom and left	green
Example 2	green	should we wait?	red
	red	yes, we need 6 for the bigger ones	green
	green	we kind have to right	blue, red
	green	yup	red
	green	lets go to the center	blue, red

Table 8.1: Examples of chats initiated before an action

	Sender	Message	Receiver(s)
Example 1	green	its stuck	red
	red	pull with shift	green
	green	oh ..	
Example 2	yellow	from the top	orange, purple
	purple	can't fit	yellow, orange
	yellow	and then left to right	orange, purple
	purple	isn't it better to push to left from the right...	yellow, orange

Table 8.2: Examples of chats initiated in response to an event/conflict

Lastly, one interesting behavior observed in few of the teams is the information transfer between players. Due to the limited communication radius, messages sent by one player could not always reach all intended recipients (e.g., when coordinating to move one of the large boxes). Few players implemented a message-relay mechanism to overcome this limitation (Table 8.3).

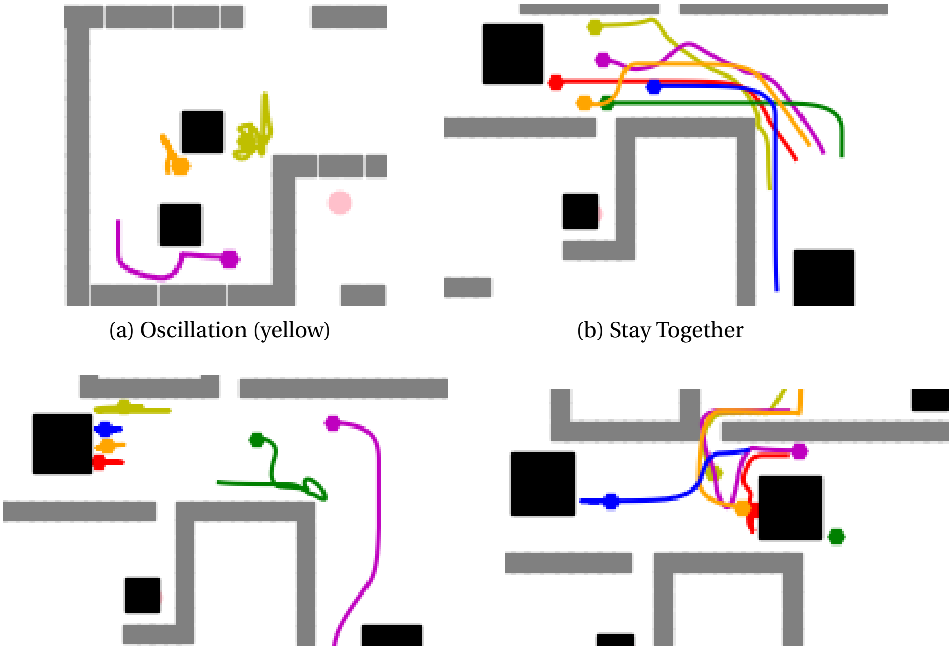
	Sender	Message	Receiver(s)
Example 1	red	tell others	purple, yellow
Example 2	red	call orange	blue,purple,yellow
	blue	come	orange

Table 8.3: Examples of the message-relay mechanism

8.3.4 Emerging Communication System Analysis - No Chat Condition

In this condition, communication is only possible implicitly, via the movement of the players' avatars. However, the chat box is kept visible for players who want to leave comments to the researcher during the game. Three players used the chat at the beginning of the game, to address other players, before realizing their mistake, and one player used it throughout the game to report who was cooperating and when.

The qualitative inspection of the games' replays allowed us to identify a number of emerging



(a) Oscillation (yellow)

(b) Stay Together

(c) Calling (green and purple)

(d) Local Voting (blue vs. others)

Figure 8.4: Emergent behaviors in the non-chat condition

behaviors and communication mechanisms.

1. "Oscillating behavior": A player moves their avatar back and forth, in a way which we hypothesize to be an indication of the direction in which they want to go (Fig. 8.4a).

2. Once together, players stay and navigate together (Fig. 8.4b).

3. "Calling behavior": A player moves close to another player, stops then moves, to signal to the other player to follow them (Fig. 8.4c).

4. "Local voting system": this behavior is particularly noticeable when the 6 players need to decide which of the two large boxes to move first. If divided between the two boxes, players in the minority group tend to go and join the majority around the other box (Fig. 8.4d). This behavior also appears when choosing in which direction to move the box. Each player chooses one side of the box and they eventually all converge to a same side.

8.4 Conclusion and Future Work

In this chapter, we study the effect of different communication affordances on a collaborative task to better understand the role played by communication medium and scope on multi-agent coordination and, concretely, the emergence of communication systems and team performance. The task was developed as an online multiplayer game where each player controls an avatar and the goal is to move boxes to target positions. Some of the boxes require the joint action of multiple players to be moved, thus requiring players' coordination. We allow for three types of communication within the game: a global chat among all players, a local chat among players within a certain distance from each other and no chat. In all conditions, no external communication among players is allowed.

The results of our experiment, involving 6 teams in each condition for a total of 108 participants, show that there was no significant difference between the three conditions in team performance (i.e., the time taken to complete the task). A deeper analysis of the messaging behavior in the Global condition reveals a significant positive correlation between the number of messages exchanged and the time taken to complete the task, suggesting that chatting is not necessarily helpful towards coordination. Following a similar trend, the 3 best performing teams in the Local condition were also the ones with the lowest number of exchanged messages, and the lowest chatting frequency. Further observation of the game replays showed the emergence of two distinct triggers for initiating communication: either to agree on an action before taking it, or in response to a conflict/difficulty. A similar qualitative analysis of the players' behavior in the No-Chat condition revealed the emergence of signaling and coordination mechanisms, based on the avatars' movements and for agreeing on an action.

We believe that our game has the potential of being a participatory learning activity for complex systems understanding, complementary to the Cellulan World activity discussed in Part I. Since, as shown in this chapter, different behaviors emerge with different communication affordances, experiencing different levels of the game with different communication affordances along with discussing the results and replays at the classroom level can help learners reflect on different types of strategies (centralized vs decentralized) depending on the means at their disposal (communication affordances).

Part III

The Evolution of the Cellulo Robotic Platform

Preface

The Cellulo platform served as the primary context for our research. In this part, we focus on augmenting and evolving the platform along three critical axes: software, hardware, and interaction capabilities.

From a software standpoint, we have developed a Software Development Kit (SDK) to facilitate the design of learning activities, thereby democratizing access for the developer community. This significant enhancement was crucial and was applied to the activities and studies outlined in Parts I and II.

From a hardware perspective, we reinvent the Cellulo robot to feature a modular design. This transformation towards modularity emerged as a valuable improvement strategy, aiming to boost the practicality and versatility of the robots, thereby broadening their potential applications.

In the domain of interaction capabilities, we extend the platform's interaction capabilities. We have enabled online interaction, integrated virtual reality (VR) features, and have also incorporated the Cellulo platform with other tangible interfaces and modular robots. This expansion aims to push the boundaries of what is currently possible, exploring new horizons in the realm of interaction capabilities of the Cellulo platform.

Chapter 9 will delve into the hardware and software dimensions from a robotic standpoint, while Chapter 10 will explore the expanded interaction dimension in greater detail.

9 Modulo Cellulo

Most of the material of this chapter is based on the paper: Khodr, H., Holdcroft, K., Wu, Y.-S., Borja, V., Sprumont, H., Bruno, B., Paik, J., & Dillenbourg, P. (2022). Modulo cellulo: modular versatile tangible educational robots. 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 3666–3671.* https://doi.org/10.1109/IROS47612.2022.9981983

9.1 Introduction

Cellulo was designed following five key design principles (Özgür, 2018): 1) Ubiquity: with the analogy of a “pen and paper", this principle refers to blending our robotic platform into the daily learning routines of classrooms. 2) Practicality: the robot is meant to be used in a classroom setting within real lessons, therefore our platform must be flexible yet reliable to allow continuous usage in such settings. 3) Versatility: rather than a tool that is only useful for teaching a single topic, a learning platform should be used in a variety of learning settings and disciplines. 4) Tangibility: inspired by Papert’s constructionist principles, the platform should inspire hands-on learning. 5) Multiplicity: the possibility to have multiple robots as well as multiple learners within the same activity, requiring the robots to enable social interaction and collaboration among learners.

We introduce Modulo Cellulo, an upgrade and a successor of the existing Cellulo, which provides modularity to the core structure. This modularity reinforces the practicality principle, by making the robot more robust to failure as it is easy and practical to change the parts. It also reinforces the ubiquity and versatility principles, by introducing task-specific functionalities which can be put together by assembling the needed modules. It further supports the tangibility principle as it enlarges the spectrum of tangible interaction possibilities. Modularity also strengthens the social collaboration between learners who might have robots with different but complementary functionalities. Introducing modularity induced mechanical, electrical, and software changes and enhancements.

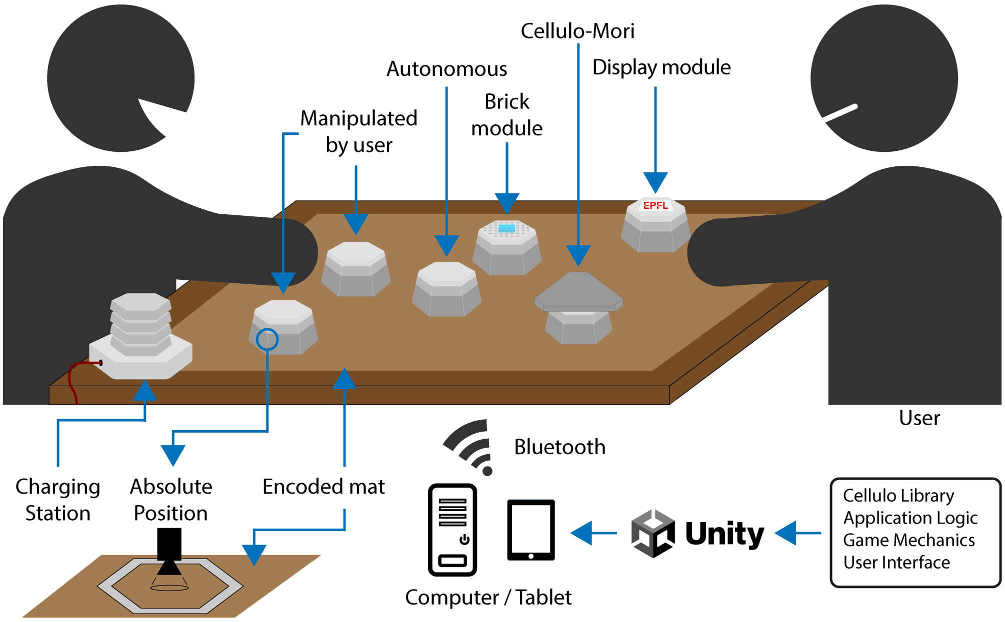


Figure 9.1: The Cellulo platform

Modulo Cellulo focuses on the design and use of robots in education where it is not to be regarded as a robot (i.e. not programmed or assembled) nor as a social tutor; but rather as an “on-demand" assistance tool for teachers, that may be utilized in a variety of disciplines and in a variety of ways for learning.

9.2 Platform Design

Figure 9.1 illustrates an overview of the Modulo Cellulo platform, which consists of three main components:

1. The *workspace*, consisting of printed sheets of paper “augmented” with a dot pattern (Hostettler et al., 2016). Depending on the activity requirements, the graphics on the paper can be designed or changed accordingly.

2. The *Modulo Cellulo robots* themselves, which are modular tangible haptic-enabled mobile robots. They can be either manipulated by a user or move autonomously, and provide versatile user interaction experiences.

3. A central *orchestrator* (usually a laptop or tablet) which connects to the robots through a star network composed of point-to-point Bluetook SPP links.

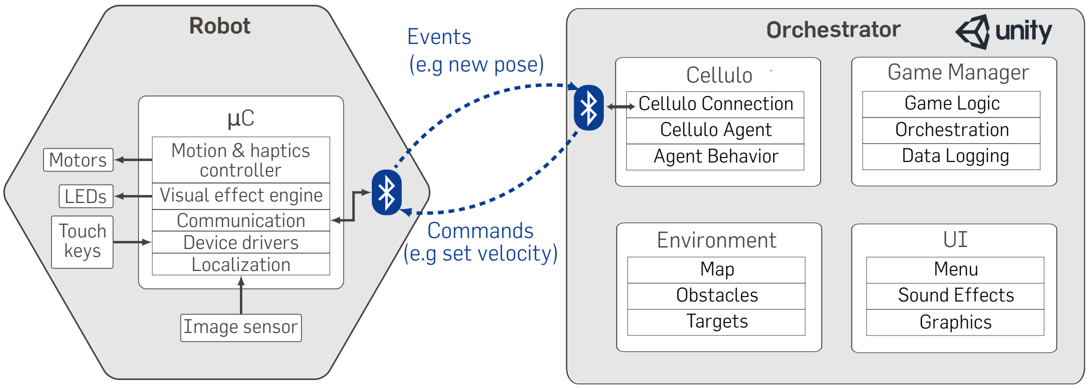


Figure 9.2: Software Architecture showing the Robot (left) with the essential on-board software components implemented in the robot firmware and the Orchestrator (right) running the Unity application corresponding to the activity.

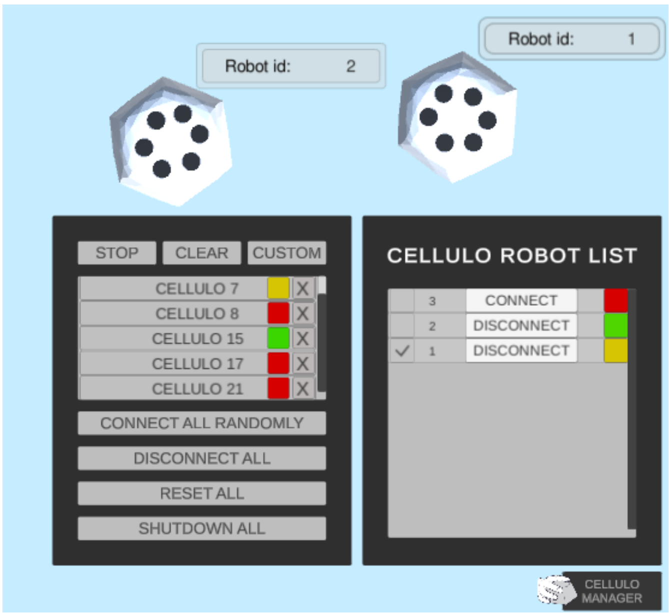


Figure 9.3: Cellulo Connection UI: green means connected, orange trying to connect, and red not connected. When scanning the list of robots available appear on the left panel. On the right panel is the list of robots in the game. It is possible to determine which physical robot each virtual robot is connected to, or randomly connect to some of the available robots.

Software Design & Development Kit

Following the principles outlined in the Introduction, we have designed our software architecture (Figure 9.2) to be driven by the activity rather than by the platform's capacities. Concretely, the robot firmware only implements on-board essential software components which are latency or bandwidth-sensitive, such as the image processing for localization, the motion controller, the haptics controllers, the battery management system, the grasp detection, and the user interaction engine. Each robot is connected through a Bluetooth Serial Port Profile (SPP) channel to a main orchestrator that runs a Unity[1] application corresponding to the activity.

We selected Unity since it provides a powerful environment for the fast development of rich activities from a software development perspective due to its game engine and built-in functionalities. Furthermore, from a deployment perspective, thanks to its portability to many platforms and devices, Unity allows to deploy applications on desktop computers running Linux, Windows, or Mac operating systems as well as consumer mobile devices or tablets, or even VR headsets (as demonstrated in Chapter 10). We developed a Software Development Kit (SDK) to streamline the process of creating learning activities[2].

As part of the SDK, we define four main modules.

Cellulo

We developed a cross-platform Unity plugin which enables a seamless connection to the Cellulo robots via Bluetooth, allowing to have the Cellulo robot as a hardware-in-the-loop in any designed activity. While our platform is centralized by design, and lacks robot-to-robot direct communication due to the star topology of our communication network, robot-to-robot interactions can always be simulated through the Unity game engine by leveraging built-in sensors. Keeping this in mind, two additional components (in addition to the connection component) are added to the main game object of Cellulo in Unity:

1. *Cellulo Agent* simulates the distance between robots or obstacles, defines neighbors, and can be affixed to a rigid body[3] to control the robots via physics simulation.
2. *Agent Behavior* component governs the actual behavior of robots and permits the addition of multiple behaviors. Two blending techniques for combining behaviors are available:
    - The *by weight* method where each behavior's output force is assigned a weight and the final output is the weighted average of all active behaviors.

[1] https://unity.com/
[2] available at https://github.com/chili-epfl/cellulo-unity-plugins.git
[3] https://docs.unity3d.com/ScriptReference/Rigidbody.html

- The *by priority* method which involves applying a cascading effect from high-priority to low-priority behaviors. This strategy is advantageous when heavier behaviors overpower the contributions of lighter ones that still need to be accounted for.

Environment

The environment is the space where the gameplay occurs, which can consist of various elements such as walls, landscape, terrain, obstacles, targets, and decorations. The design possibilities for the environment are limitless and offer creative opportunities.

In the physical part of the game, the environment is mapped onto the Cellulo workspace. This is achieved using a plugin developed as part of the SDK that allows for a one-to-one mapping between the real-world dimensions and those in the Unity environment. Developers can even mix real and virtual environments, such as adding hidden obstacles that are present in the virtual environment but not in the physical one, creating more options for game mechanics.

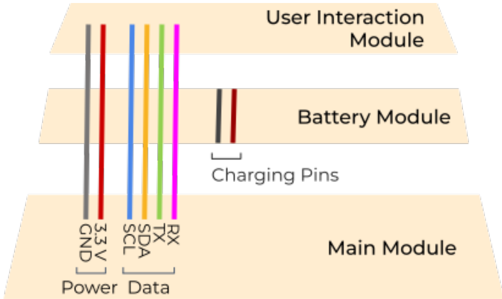
User Interface

The user interface (UI) is a critical component of games, as it significantly impacts both the gameplay experience and the visual identity of the game. Additionally, it serves as a powerful tool for interacting with and directing players throughout the game. The UI provides a menu for players to interact with the game, such as tracking progress between levels, displaying time and score, or providing hints. In the SDK, we also provide a UI for connecting to real Cellulo robots. Figure 9.3 depicts an example of the UI used to connect to a Cellulo robot. The Unity game engine allows developers to add sounds and adapt graphics to meet the needs of the activity.

Game Manager

The Game Manager module is responsible for deploying the game logic and orchestrating the game elements, such as the timeline, hint characters, and scoring mechanisms. These elements are essential in controlling the gameplay experience and allowing players to track their progress. Additionally, a logging mechanism is implemented to capture data about the player's interactions within the game. This data is particularly useful for experimental studies that evaluate the effectiveness of the game in achieving its intended outcomes or seek to identify areas for improvement during iterative game design.

(a) Main Module (bottom) and battery module (top) connected via sliding rails, magnets and pogo pins



(b) Electrical connection between modules

Figure 9.4: Mechanical and electrical connection between modules.

9.3 Modular Robot Design

9.3.1 Overview

Modulo Cellulo robots are conceptualized to be composed of three modules: A main module which contains essential functionalities, a battery module containing the power management system, and a user interaction module which can be changed and adapted based on the activity requirements.

9.3.2 Connection between Modules

From a mechanical perspective, the connection between modules need to be robust enough to not break down during use, but at the same time easy to assemble and disassemble. The connection interface between modules is designed to be intuitive for most users. The sliding rails (Figure 9.4a) provide a solid connection between modules and ensure the user correctly aligns the modules with one another. Magnets provide an extra force to align the modules and avoids unintentional sliding. As for the electronic connections, spring-loaded pogo pins touch target contacts and form a stable connection. The combination of sliding rails and pogo pins (Figure 9.4a) can thus eliminate the shortcomings of the commonly seen header pins and sockets connections (such as stacking Arduino shields). A pair of pogo pins in the middle of the connection port is the power supply for all modules, providing direct battery access. Two pairs of pogo pins on the connection port ensure data transmission with the UART and I^{2}C protocols. The two pins separate from the connection port are used for charging and only exist on the battery module (Figure 9.4a).

9.3.3 Main Module

The main module is designed to be the *base* of the whole system. The main module controls the movement of the robot, while the other modules act as *worker* devices managed by the

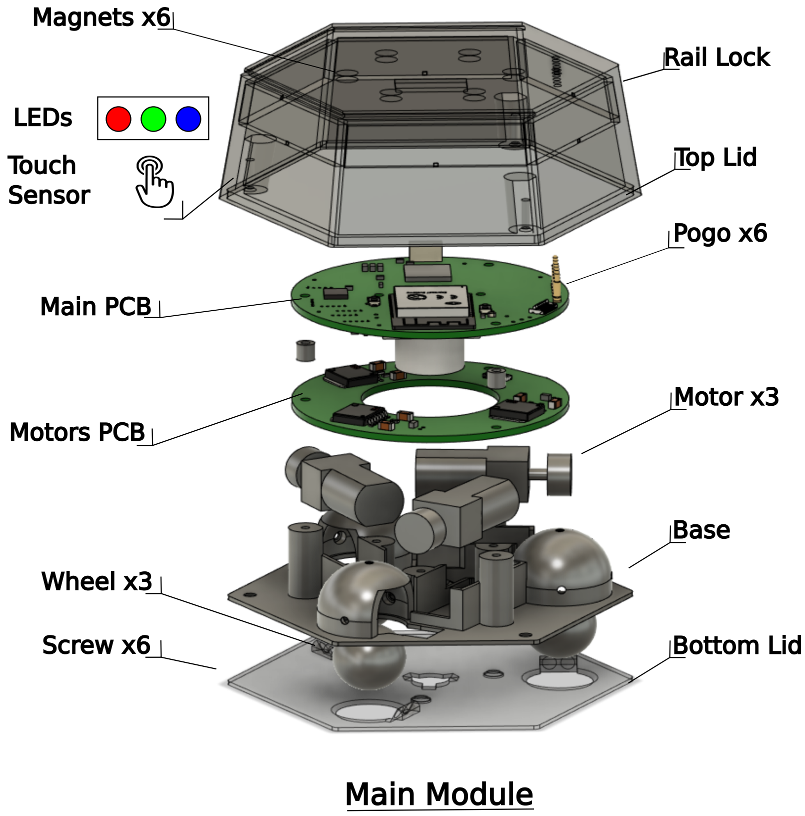


Figure 9.5: Exploded view of the Modulo Cellulo main module

main module. The main module shown in Figure 9.5 includes the main micro-controller (chosen to be a PIC32) and is responsible for all the main computational work. This includes five essential tasks: localization, locomotion, wireless communication, coordination between modules, and grasp detection.

Localization

We use the same fast and accurate localization system originally developed for Cellulo, which is based on a dense, deterministic and well-defined optical microdot pattern printable on regular office printers. More details can be found in (Hostettler et al., 2016). The main module is equipped with a global shutter image sensor facing downward which acquires an $\sim 1cm^2$ region of the printed sheet at about 93 Hz, from which a 3 DOF pose (x, y, θ) can be extracted. This approach provides a global and absolute localization with $\sim$ 0.17mm and $\sim 1.5°$ accuracy without the need for any calibration. Furthermore, thanks to the hardware design and the placement of the camera beneath the robots, the system is robust against external illumination conditions and enables instant recovery from kidnapping, thus allowing any manipulation to be performed without any adverse influence on localization performance.

Locomotion

Although locomotion doesn't have its own module mechanically, the modularity principle is followed with the electronics design, as we dedicate a separate PCB for the motor drivers for two main reasons: 1) to have a better heat distribution for the drivers and 2) to leave room for changes in the locomotion system (to enable a faster or more powerful drive, for example, or even passive, non-actuated, modules). The current locomotion drive is the permanent-magnet assisted omni-directional ball drive, detailed in (Özgür et al., 2016) and shown in Figure 9.5. In addition to providing holonomic motion and mechanical robustness against intensive user manipulation, it enables kinesthetic haptic feedback in the form of force/torque output on the learner's hand, as well as backdrivability assistance to overcome the natural impedance of the robot due to the friction of the wheels on the paper when the robot is moved by the user.

Adaptive Motion Controller

Introducing modularity to the Cellulo robot implies the possibility of dynamic changes to the physical properties of the robot, requiring an adaptive motion controller. A non-adaptive system would either be unable to overcome friction with a heavy module or overshoot any speed constraints with a lighter load. Moreover, in the case where the user interaction module is a self-reconfigurable modular robot (see Sec. 10.2), the module's weight and center of mass can dynamically change in real time. We thus implemented an adaptive controller for the Modulo Cellulo, capable of self-tuning its control parameters over time. The motion control of Cellulo was an open-loop velocity controller which incorporates wheel orientation, weight, and friction for non-holonomic motion (Özgür et al., 2016). The new controller is composed of two parts: A feedback linearization component to decouple and linearize the system, and a Model Reference Adaptive Controller (MRAC) that adjusts the parameter gains according to the online controller performance (Åström & Wittenmark, 2013). Through the feedback linearization module, the augmented system to control becomes a triplet of decoupled, perturbed, simple integrators:

$$\dot{q} = b_q \cdot (v_q - f_q(X)), \quad q \in \{x, y, \theta\} \tag{9.1}$$

where $f_q(X)$ is a term encompassing all perturbations on the system, including friction and manual user input. b_q is a value that depends on the system modeling; unknown but constant while the mass of the system is unchanged. v_q is the control input that contains the controller parameters.

The control input is composed of a feedback term, a feedforward term, and a perturbation compensation term:

$$v_q = -\hat{k_{pq}} \cdot \dot{q} + \hat{k_{rq}} \cdot r_q + \hat{\theta_q}^T \cdot \Phi_q(X) \tag{9.2}$$

where $\Phi_q(X)$ is a vector containing all the linearly separable components of $f_q(X)$ that can be measured. Most importantly, it can be extended to suit the perturbation model of specific tasks. The parameters $\hat{k_{pq}}$, $\hat{k_{rq}}$, and $\hat{\theta_q}$ are updated based on the error e_q between the system response and the reference model response:

$$\begin{aligned}
\dot{\hat{k_{pq}}} &= sign(b_q) \cdot e_q \cdot \gamma_{pq} \cdot \dot{q} \\
\dot{\hat{k_{rq}}} &= -sign(b_q) \cdot e_q \cdot \gamma_{rq} \cdot r \\
\dot{\hat{\theta_q}} &= -sign(b_q) \cdot e_q \cdot \Gamma_{\theta q} \cdot \Phi(X)
\end{aligned} \tag{9.3}$$

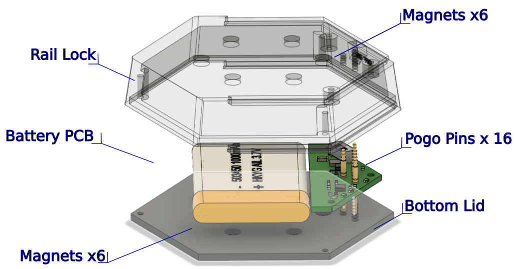
where γ_{pq}, γ_{rq}, and $\Gamma_{\theta q}$ are the learning rates. To assess the performance of the proposed adaptive controller, we compared it against the open-loop controller, on two Cellulo robots. The two robots were tasked to make back and forth motions at increasing speeds (from 50 mm/s to 150 mm/s with intervals of 25 mm/s) under 5 different added weight configurations. Each configuration was repeated 3 times, varying the initial pose. Results are shown in Table 9.1. The difference in performance is significant (p-value<<0.05) for any added weight, with the maximum improvement registered with an added weight of 200 g (100% with respect to the weight of the robot), for an average 19% performance increase using the adaptive controller.

Added weight[g]	Mean RMSE open-loop[%]	Mean RMSE adaptive[%]	p-value	Effect size
0	58.68	61.27	0.33	0.42
50	68.07	58.48	0.001	0.74
100	71.45	63.67	0.009	0.70
150	77.18	61.00	3.6e-8	0.91
200	83.30	64.38	5e-7	0.85

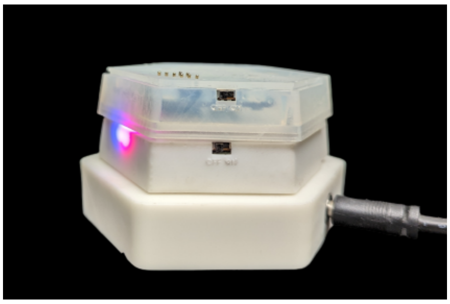
Table 9.1: RMSE percentage comparison between open-loop and adaptive controllers. p-values are given by the Mann-Whitney U test. Effect size is the Common Language Effect Sizes, i.e. the percentage of open loop runs where the error was higher than for the adaptive one.

Wireless Communication

Each robot communicates with the controller wirelessly. To do that, the main board is equipped with an ESP32 which is acting as a Bluetooth bridge. Moreover, the ESP32 can be programmed over-the-air making it easier from a deployment perspective to add new features and functionalities, including peer-to-peer communication for which work is ongoing. The wireless communication takes care of sharing information categorized along two main lines: 1) Transmitting local robot information such as robot pose, battery status, user interaction data, and debug data if requested, such as image frames, or on board profiling; 2) Receiving commands such as setting goal velocities, tracking goal poses, haptic command,



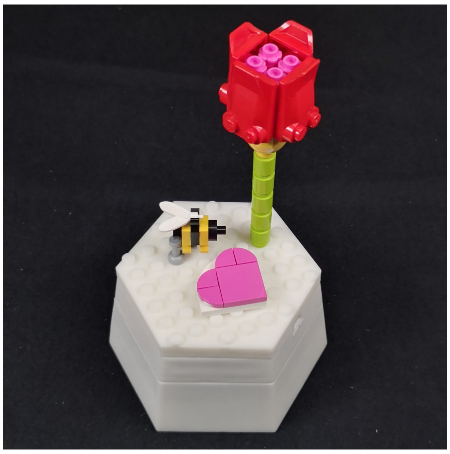
(a) Exploded view of the battery module



(b) Charging station

Figure 9.6: Battery Module and its charging station

(a) Display Module



(b) Brick Module

Figure 9.7: Examples of two types of user interaction modules

power commands (reset, sleep) and commands for the user interaction module.

Coordination between Modules

The main module is also responsible for managing and detecting which module is connected to it. The main firmware has two ways to detect the attached module: by a wireless command through the app, or by an automatic scanning of devices in the I^2C line.

Grasp Detection

To detect if the module is being moved by the user or not, conductive thin films are placed inside the sides of the main module, to act as capacitive touch sensors. This is a direct consequence of the feedback from the use of Cellulo, where the placement of the touch sensors at top was not reliable enough because users usually hold the robots from the sides in order not to cover its illuminated top.

9.3.4 Battery Module

The battery module (Figure 9.6a) is designed to be a separate unit in order to increase the practicality of the robot's usage. If the battery is drained, the user need not wait for a full recharge to continue their activity, but rather just swap the battery module with a fully charged one. This module includes a rechargeable battery and a Battery Management System (BMS). A Lithium polymer battery (LiPo) is embedded in the battery module on account of higher energy per unit volume and reliable battery safety. The BMS is composed of a BQ27441-G1 System-Side Impedance Track™ Fuel Gauge and a BQ24075 Standalone 1-Cell 1.5-A Linear Battery Charger with PowerPath. The BQ27441 provides battery information such as remaining battery capacity (mAh), state-of-charge (%), and battery voltage (mV). The BQ27441 uses I^2C serial bus for communication with the main module. The BQ24075 integrates Li-Ion linear chargers and system power path management. The LiPo battery has 1200mAh capacity: a constantly autonomously moving Cellulo would last around 1 hour. A charging station is designed to stack the battery modules above each other as illustrated in Figure 9.6b.

9.3.5 User Interaction Module

The user interaction module is the top module of Modulo Cellulo and provides the most versatile options. The basic UI module is equipped with 6 RGB-illuminated capacitive touch buttons allowing simple touch interaction and visual feedback through the LEDs. We can think hereafter of a multitude of different options for a user interface. Examples include a display module (Figure 9.7a) which provides richer information, and can be used in the activity to disclose hidden information embedded in the map as in (Johal, Tran, et al., 2019). Another option is a passive attachment with the footprint of LEGO® bricks (Figure 9.7b) allowing the addition of passive props and personalized designs on top of Modulo Cellulo. Another passive attachment, motivated by the use of Cellulo in upper limb rehabilitation scenarios (Guneysu et al., 2018), consists of a grasping aid to help people with mobility impairments grasp and hold the robot. Other UI modules currently under consideration are tactile thermal displays or a shape changing display. A thermal display could be a Peltier element thermoelectric cooler (TEC) which can generate hot and cold feeling in an application around heat transfer for example. The shape-changing display can act like Braille symbols to help visually impaired people. Apart from tactile, a sonic module including a microphone and a speaker can add a new dimension of spatial and social interaction.

9.4 Conclusion

In this chapter, we have introduced a novel upgrade to the Cellulo platform by providing a software development kit for developers to easily deploy activities and games while seamlessly connecting to the robots. We have also added modularity to the core of the versatile tangible robot now called Modulo Cellulo. We detailed its design and implementation along various facets, including mechanical, electrical, control and software design.

The resulting platform is: 1) ubiquitous by being able to seamlessly integrate into the classroom setup; 2) practical, by consisting of modular robust wireless robots operating on a paper workspace with a straightforward setup and not limited to a specific deployment method thanks to the Modulo Cellulo cross-platform activity design environment in Unity; 3) versatile, by enriching and blending into (rather than changing) the current teaching practices of various disciplines thanks to different interaction modules; 4) tangible, by enabling the users to move the robot and receive haptic feedback; 5) while asocial from a robot perspective, enabling social interactions through the possibility of having multiple robots and multiple users interacting simultaneously.

10 Expanding the Interaction Pool

The material in this chapter is based on multiple resources. The first section is based on a collaborative project between RRL and CHILI labs I was leading. Multiple people contributed to this work, including my colleague Kevin Holdcroft from RRL, and our students Jean-Etienne Charbonnet, Daniel Demko, Paul Juillard, Julien Jordan, and Nicolas Wagner.
The second section is based on the paper : Ozgur, A. G., Khodr, H., Akeddar, M., Roust, M., & Dillenbourg, P. (2022). Designing online multiplayer games with haptically and virtually linked tangible robots to enhance social interaction in therapy. 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), *358–364.* https://doi.org/10.1109/RO-MAN53752.2022.9900684.
The third section is based on the paper Khodr, H., Ramage, U., Kim, K., Guneysu Ozgur, A., Bruno, B., & Dillenbourg, P. (2020). Being part of the swarm: experiencing human-swarm interaction with vr and tangible robots. Symposium on Spatial User Interaction. https://doi.org/10.1145/3385959.3422695.

10.1 Introduction

In this chapter, we highlight several extensions designed to increase the pool of interactions between humans and the Modulo Cellulo robot. Firstly, a reconfigurable user interaction module has been added to Modulo Cellulo, resulting in a reconfigurable swarm of robots. Secondly, an online connection add-on for Cellulo has been introduced, enabling human-robot-robot-human interaction across the globe. Lastly, integration with virtual reality (VR) has been implemented to enable users to experience different perspectives. These extensions demonstrate the continued evolution of the Cellulo platform towards more versatile and interactive experiences.

(a) Cellulo-Mori

(b) Ippon[1]: Linking Mori together can lift Cellulo off the ground

Figure 10.1: Modulo Cellulo with Mori as an user interaction modules

10.2 Reconfigurable Swarm of Robots

10.2.1 Cellulo-Mori

In Chapter 9, we observed that the Cellulo robots are used to be the representation of point-like objects, independent and detached one from the other. However, adding a reconfigurable top module to Modulo Cellulo paves the way for the representation of objects (or concepts) via shapes collectively generated by multiple robots linked together. We demonstrate the benefits of Modulo Cellulo's modularity by incorporating the Mori modular robot on top. The Mori is a self-reconfigurable modular robot and consists of quasi-2D triangular modules which are able to connect to other modules with the same connector and rotate around the control joint (Belke & Paik, 2017). These modules provide a high ratio of area to volume, allowing the Mori to fold into large 3D shapes and structures with relatively little mass. By integrating Mori on top of Modulo Cellulo, we incorporate the strengths of both robots. The Mori provides Modulo Cellulo with the ability to move out of plane, as well as link to other robots to form large structures. Modulo Cellulo ensures faster planar motion and localization, along with robustness to human manipulation. The user interaction module which enables linking Modulo Cellulo and Mori is aptly named a "Cellulo-Mori" (Figure 10.1a). The module consists of a mechanical anchor and electronics to integrate a Mori onto Modulo Cellulo. The anchor spans the width of the Mori and latches onto mechanical holding points on the Mori's corners. The electronics provide power from Modulo Cellulo to the Mori and link to the Modulo Cellulo's controller, including wireless functions. Thus, the Cellulo-Mori can be operated via the same wireless protocol as Modulo Cellulo, easing simulation-to-real transitions. Example of the simulation of a game with Modulo Cellulo-Mori is shown in Figure 10.2. The Cellulo-Mori can transfer Mori modules between each other, including across multi-level planes which was

[1] https://budodragon.com/what-ippon-means-in-judo-rules/

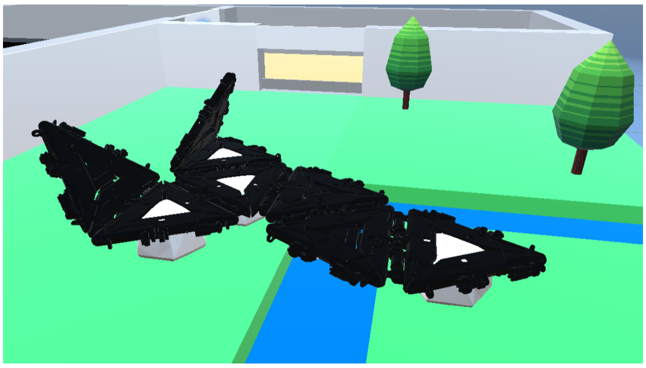


Figure 10.2: Modulo Cellulo-Mori - Example of a collaborative game to go over obstacles, inspired by the army ant behavior of building bridges

otherwise impossible. Linking Mori together can lift Cellulo off the ground (Figure 10.1b), over gaps and to different elevations. This example demonstrates the adaptivity of the Modulo-Cellulo. The user interaction module not only provides the ability to augment user interaction, but can incorporate other robots which introduce different elements to the activity design. We observe that integrating two versatile robots allows for each to benefit from the strengths of the other to create new physical operations. Modularity eases the integration of external devices onto Cellulo, providing new methods to explore a large spectrum of learning activities and disciplines.

10.2.2 Cellulo-JoJo

We also demonstrate the benefit of modularity through an alternative application, incorporating Joint-Space Joysticks (JoJos)[2] as another UI module for Modulo Cellulo. JoJos are tangible user interfaces which resemble a modular robot; in our case it emulates a Mori. Users can directly touch, move the JoJos and connect them together via mechanical pins. JoJos have magnets in each joint, which are measured by rotary hall-effect sensors of neighbouring JoJos, allowing them to connect at arbitrary angles.

There are two types of JoJo: active and passive (Figure 10.3). Active JoJos contain full functionality – wireless connectivity, sensors to measure angles between neighboring units, and power regulation circuitry. Passive JoJos have the same physical structure and magnets, but none of the electronics. If every module was to have full functionality, each joint would be measured twice due to the parallel structure of the Mori and corresponding JoJos. Connections of JoJos can thus be staggered between active and passive modules and still have a measurement for every angle. A Cellulo-JoJo unit consists of a Modulo Cellulo and a passive Mori JoJo. Active JoJos can connect to each JoJo edge, allowing users to build different shapes and to operate out-of-plane.

[2]JoJos are designed, built and developed at the Reconfigurable Robotics Lab (RRL) at EPFL and are used within a collaborative project between CHILI and RRL.

Figure 10.3: Active JoJo (right) and Passive JoJo (left) connected together via a mechanical pin.

Figure 10.4: Setup of the collaborative game Dillenbear's journey to the Farlands.

Figure 10.5: The simulation part of the game.

Learning Activity

As a proof of concept for the usability of Cellulo-JoJos, we designed a gamified learning activity. The activity, as shown in Figure 10.4, is a collaborative game involving two players. One player is responsible for physically interacting with the robots and building the connections between them. The other is responsible for controlling the virtual aspects of the game, which include selecting the rules and running the simulation. Communication between the two players is essential to successfully advance in the game.

The story line of the game involves an alien species, the “Dillenbears" who explore a world made up of islands, the “Farlands". Dillenbears are little blue bear-like characters facing multiple challenges in their exploration of the islands and asking the players for help. The activity is divided into tasks, each designed as a puzzle. The final version includes two main tasks, each organized into two levels. The goal of the two levels is to demonstrate how contextual differences, such as the number of Dillenbears involved in the activity or the information available about the environment, can lead to significant differences in the best way to solve the task. This is related to the importance of agent-environment interaction in a complex system. To solve the tasks, the player would use Cellulo-JoJos and JoJos to create a path for the Dillenbears. In the physical world, islands are linked to Cellulos, while Cellulo-JoJo and JoJos have a one-to-one mapping between the physical and the virtual views (see Table 10.1).

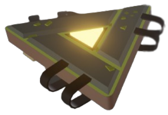
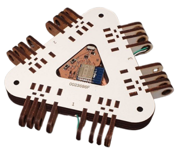
	Island	JoJo	Cellulo-Jojo
Virtual			
Physical			

Table 10.1: Game Building Blocks

Gameplay

In the first task, the objective for players is to assist the Dillenbears in reaching a portal situated on another island. The initial level comprises one Dillenbear, one Cellulo-JoJo, and a supply of 2 JoJos. The subsequent level is identical in its aim and environment but has 30 Dillenbears striving to access the portal. The players' aim is to manipulate the Cellulo-JoJo and the stock JoJos to ensure that all Dillenbears reach the portal in the shortest possible time. Following the Dillenbears' traversal of the portal, they are moved to a new dwelling atop a Cellulo-JoJo that they wish to repaint. Paint is present on neighboring islands, and the challenge is to gather as much paint as possible. In the first level, users are aware of the quantity of paint on each island. In the second level (Figure 10.5), this amount is unknown. The players' objective is to discover the optimal approach to attain the highest score in each level. The player can modify the rules of the joints by altering their colors. The available rules for players to experiment with include:

- Yellow: Buffer - predetermine the number of Dillenbears heading in each direction;
- Orange: Random - the Dillenbears move in random directions;
- Blue: Random+Attractor - the more Dillenbears bring paint from a joint, the more attractive the joint becomes (mimicking the pheromone trails used by ants);
- Green: Random+Attractor+Memory - the attractiveness of a joint is regulated by evaporation.

When the paint quantities in each location are known beforehand, a deterministic approach using the Yellow rule would work best. However, when this is unknown, a more distributed approach (using the Green rule) inspired by ant behavior would be a better option. Users are left to explore the difference between a purely random approach, a highly deterministic approach set by the user, a positive feedback with the environment, and lastly, one that is also regulated with time.

While we did not do any experiments, it would be interesting to validate the effectiveness of the game, as well as to study the social human collaboration initiated by the asocial robot.

Conclusion

The integration of reconfigurable modular robots such as Mori onto Modulo Cellulo has paved the way for representing objects or concepts via shapes collectively generated by multiple robots linked together. The Mori's ability to fold into large 3D shapes and structures, combined with Modulo Cellulo's faster planar motion and localization, create new physical operations that neither robot could achieve alone. The user interaction module, Cellulo-Mori, not only augments user interaction but also allows for the incorporation of other robots to introduce different elements to activity design. Additionally, incorporating Joint-Space Joysticks (JoJos) as another UI module for Modulo Cellulo demonstrates the flexibility and versatility of the platform. Cellulo-JoJo units can be used to build different shapes and operate out-of-plane, providing new methods to explore a large spectrum of learning activities and disciplines. The proof of concept gamified learning activity designed with Cellulo-JoJos, Dillenbear's journey to the Farlands, demonstrates the usability of the platform for educational purposes. Overall, the modularity of Modulo Cellulo offers a powerful tool for exploring new ways of interaction between humans, robots, and the physical world.

10.3 Enabling Tangible Multiplayer Interaction Over Distance

Using a tangible robot in a social setting during a pandemic can be challenging due to physical distancing requirements and the need to minimize in-person interactions. Therefore, an online adaptation of our framework would be a viable solution to alleviate some of the consequences of this situation. In this section, we introduce an add-on to our platform that enables the connection of Cellulo online.

From an implementation perspective, the software architecture is shown in Figure 10.6. We leverage the Photon Engine[3] to provide seamless and real-time data exchange, enabling online multiplayer gaming. The Photon Engine is integrated with Unity, which is the platform we utilized to build our framework, making it easy and straightforward to use. With Photon, players' data can be synchronized, and "rooms" can be created to host game sessions.

We envision two modes of the remote interaction to allow the players to readily connect and play, with or without the need for having the robotic platform on their side:

- *Virtual-to-Tangible Interaction:* Envisioned if the other party does not have the robotic platform available. An example is shown in Figure 10.7. One player must have the Cellulo hardware, including a printed map and physical robot(s), on their side. The other player can join the game remotely using their laptop. They can control the robots on the first player's side using their computer's keyboard (arrow keys) and can see the game

[3] https://www.photonengine.com/, accessed May 2023.

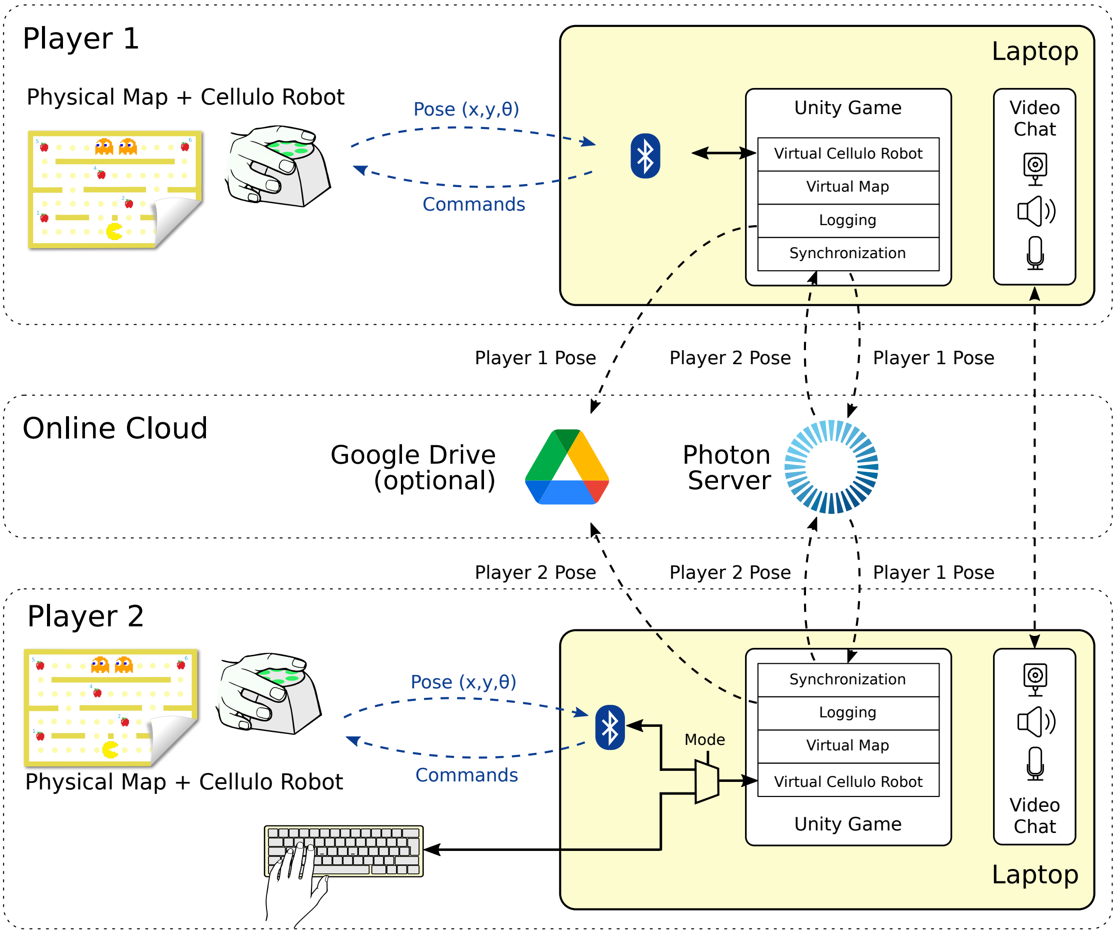


Figure 10.6: Software architecture facilitating two modes of remote interaction for the game: Virtual-to-Tangible and Tangible-to-Tangible. In the Virtual-to-Tangible mode, one player has the Cellulo hardware, including a map and physical robots, while the other player connects remotely via a computer, controlling the robots using keyboard commands. In the Tangible-to-Tangible mode, both players have the Cellulo hardware, physically controlling their robots and tracking the game on their individual maps or screens. Player poses are synchronized over a Photon server and are optionally logged at a cloud storage service such as Google Drive. A video/audio chat session can be enabled throughout the game to improve social interaction.

map and the robots' synchronized positions on the screen. This allows for real-time collaboration and competition, while minimizing the need for physical interaction.

- *Tangible-to-Tangible Interaction:* Envisioned when in need to preserve the haptic link between the two parties even though they are remote. Independent of the target user group, recent research on socially connected game experience suggests to consider augmenting the shared tabletop space with connected tangible components for a better online social game experience (Yuan et al., 2021). In this modality, each user must have the Cellulo hardware, i.e., a printed map and a physical robot, on their side. As a result, each player manually controls their robot and follows the game on the map or screen.

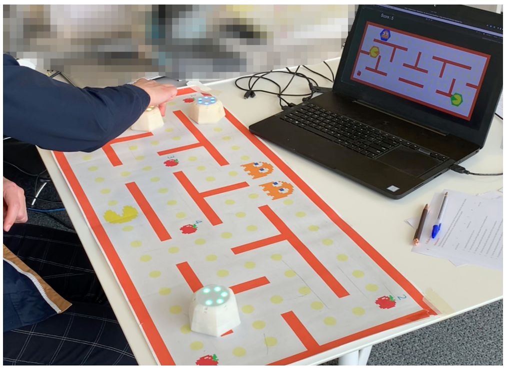
In the game application, the player has the option to create or join an online game room by entering a specific room name. Before creating the room, the master player selects the room features such as game modality and parameters. Other players can connect to the room using a standalone Unity game application (if they use tangible robots) or a web link through their internet browsers (if they use only virtual robots and do not require a connection to real robots). The browser-based option uses WebGL as a graphics backend, and no additional software needs to be installed on a modern browser. After joining the created room and confirming readiness to play, the master player can start the game for all players simultaneously. Since the game is designed to promote social interaction, the game can be played with a video or audio chat session (e.g., Zoom or Discord), enabling participants to communicate naturally.

Exemplar activity

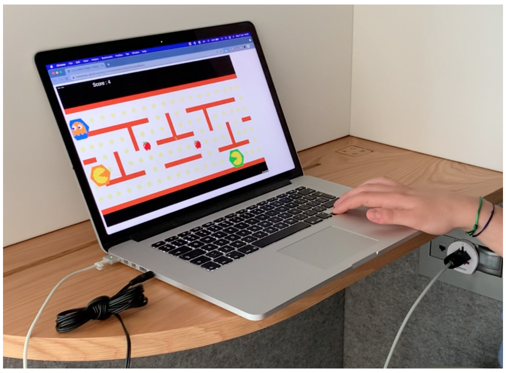
As a showcase, we have implemented a tangible Pacman game to enhance motivation and promote psychosocial well-being through social integration in home-based upper-limb therapy. Our gamified multiplayer rehabilitation platform allows relatives, children, and friends to connect and play with their loved ones remotely, supporting them with their training from anywhere in the world through the internet, which is especially relevant in view of possible social distancing measures.

In the Pacman game, there are two entities: the Pacman(s) and the Ghost. The Pacman(s) try to collect all the apples, while the Ghost chases the Pacman(s) to catch and eat all the apples. The game ends when all the apples are collected. We have two versions of the game: a collaborative one, where the two players are both Pacmans and have to collect the apples together, while an autonomous Ghost chases the closest Pacman; and a competitive one, where one player is the Pacman and the other player is the Ghost.

To evaluate the virtual-to-tangible interaction, a pilot experiment was conducted with healthy young adults with ages ranging from 20 to 40 years old to gather data and feedback on the game. 10 participants, in 5 pairs of 2 players, played 2 to 4 games in both the collaborative and competitive game versions. Every participant played on both the tangible side and virtual side of the game, at least twice per version. In total, 20 games in the competitive and 21 games



(a) "Tangible side" where the first player is playing the game on the physical map by controlling a Pac-man. One of the other robots is controlled by the remote player and the last one is an autonomous Ghost.



(b) "Virtual side" where the remote player controls either a Ghost or the second Pacman depending on the game modality. The virtual representation of the game map and all robots' synchronized positions appears on both screens.

Figure 10.7: Show case: multiplayer Pacman game with a tangible and virtual side.

Question	Average response
I enjoyed doing the activity very much	6.9
I am willing to do this activity again because I think it is somewhat useful	6.1
I felt that we cooperated with the second user	6.4
I think that doing this activity might be useful for social connection	6.6
I would recommend this activity to the elderly to play with their friends/grandchildren etc.	6.7
I think that doing this activity might be useful for arm rehabilitation at home	6.7
It is possible that this activity could be useful to improve the rehabilitation process of friends/family	6.4

Table 10.2: Qualitative survey results with mean response of the 10 participants on a 0-7 scale, 0 being "Not at all true", 7 being "Very true".

in the collaborative modality were played. The games were all played on a single large map to simplify the dimensions of the experiment. The players were placed in separate rooms and could only communicate through voice chat to simulate the conditions that the game is intended to be played in.

Participants were given a survey to measure their perceptions of the game at the end of the experiment. The survey includes seven questions on enjoyment, perceived usefulness and perceived collaboration, all measured on a Likert scale with 7 points (see Table 10.2 for the questions). Overall, highly positive feedback was received in the questionnaires and participants found the online game modalities fun and intuitive. The scores suggest that participants enjoyed the game and perceived it as useful for gamified rehabilitation, home-rehabilitation and social connection. They also gave positive responses in recommending this system to the elderly people (see Table 10.2).

Conclusion

Our framework's online adaptation presents a solution to challenges brought on by the need for physical distancing and minimizing in-person interactions, particularly when tangible robots are employed in social contexts. To illustrate the possibilities offered by our platform, we created a tangible version of the Pacman game. This adaptation can boost motivation, foster psycho-social well-being, and support remote rehabilitation for stroke patients. It can also encourage inter-generational connections and combat social isolation and loneliness among older populations. Returning to the realm of education, this online transformation of the tangible robot framework holds promising potential. It allows for the incorporation of hands-on, collaborative learning experiences, even when conducted remotely. This innovative approach reshapes our traditional understanding of online learning, which is usually confined to screen-based interactions, and introduces a new dimension of tactile engagement.

10.4 Being Part of the Swarm: Experiencing Human-Swarm Interaction with VR

We propose to combine tangible robot swarms with Virtual Reality (VR), to enable the possibility of observing and interacting with a swarm of robots from multiple perspectives. Indeed, local perception, i.e. taking the point of view of the agent, provides a highly "immersive" experience, which is one of the design principles for high embodiment in learning environments (Johnson-Glenberg et al., 2014), while global perception leads to apprehending the global emergent behavior.

The setup, as illustrated in Figure 10.8, utilizes the Cellulo framework presented in Chapter 9 with the Oculus Quest VR serving as the orchestrator in this case. Regardless of the experienced perspective, the user always acts on the other robots in the swarm by directly manipulating a Cellulo robot.

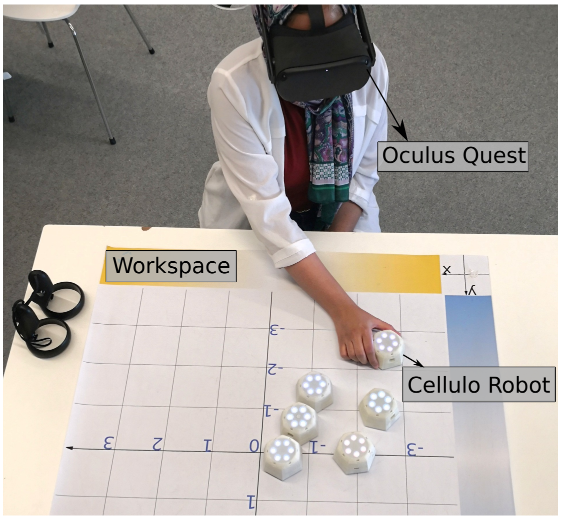


Figure 10.8: Cellulo Platfrom Setup with VR

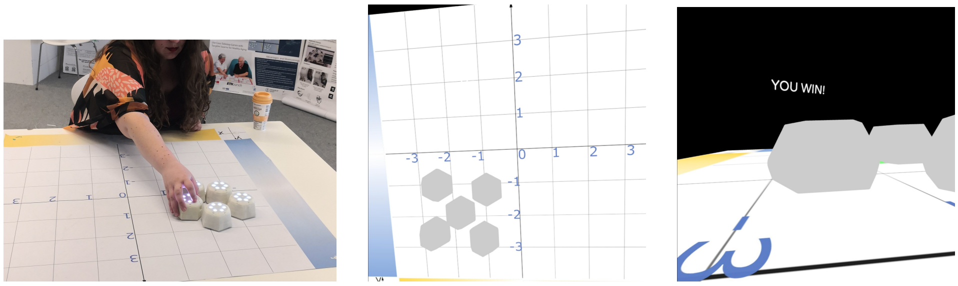
User Experiment

To evaluate the user experience of the Cellulo-VR system and to verify whether the different perspectives are perceived as significantly different by the user, a within-subject user study was performed. It involved 15 participants aged between 22 and 59 ($M = 29.73, SD = 9.05$; 7 female and 8 male), with average self-reported experience with VR technologies ($M = 3.2, SD = 1.82$ out of seven), mostly positive ($M = 4.38, SD = 1.66$, out of seven). The pilot experiment consisted of three conditions:

1. Physical condition: the learner controls one robot with the hand and directly perceives the swarm from a global perspective, without VR headset (Figure 10.9a).
2. VR top/bird view condition: the learner controls one robot with the hand and perceives a virtual representation of the swarm from a global perspective, through the VR headset (Figure 10.9b).
3. VR local view condition: the learner embodies the role of a Cellulo robot, controlling its movements with the hand and perceiving the world from its egocentric perspective through the VR headset (Figure 10.9c).

As for the activity itself, we implemented a flocking game that uses the boids algorithm introduced by Reynolds (1987). This algorithm describes an approach to simulate the natural behavior of bird-flocks, fish-schools or herds, without hard-coding each individual's path but rather by letting each agent re-arrange its course based on its current surroundings. Each agent only obeys three simple rules:

1. *Separation* (avoid collisions with nearby flock-mates),
2. *Alignment* (attempt to match the velocity of nearby flock-mates), and
3. *Cohesion* (attempt to stay close to nearby flock-mates).



(a) Physical condition (b) VR - Global view (c) VR - Local view

Figure 10.9: The three conditions tested in the experiment.

The movement of the robot manipulated by the learner is transposed to the headset in an absolute approach, i.e, the linear movement and orientation change of the camera (view in the headset) are the same of the Cellulo robot. In this case the headset is rigidly linked to the robot.

The activity relies on the Cellulo robots' haptic feedback API to render force feedback to the user's hand. A haptic feedback force is given to the user as the sum of the forces generated by the three flocking rules (separation, cohesion, and alignment). This force only depends on the interaction between the robots and is thus computed the same way in all conditions.

In all conditions, the learner is asked to reach, one after the other, a set of target points on the map. While the target positions are unknown to the user, they are known to the other robots in the swarm, as target migration point. Whenever the user reaches one point, the swarm starts moving to the next one. The user therefore needs to follow the swarm to achieve the task.

Participants completed the three conditions in a randomized order. Upon experiencing each condition, they were asked to fill the Slater-Usoh-Steed's (SUS) presence questionnaire (Usoh et al., 2000), the motion sickness questionnaire (Gianaros et al., 2001), and few custom questions related to the feeling of the haptic feedback.

Results are shown in Figure 10.10. Concretely, for the SUS presence questionnaire, the results were analyzed by taking the average answers over the 6 questions. A Kruskal-Wallis test showed a significant difference ($H = 7.85, p = 0.005$) between the local and top view in VR. The sense of presence was higher in the local condition ($M = 73.65\%, SD = 12.09\%$) than in the top condition. ($M = 56.66\%, SD = 16.52\%$). This result confirms our hypothesis that being in local view in VR would increase the sense of embodiment. For the motion sickness questionnaire, the average over all questions shows that the local view ($M = 24.12\%, SD = 16.41\%$) caused significantly higher sickness than the global view ($M = 12.63\%, SD = 3.25\%$). This result is expected since joystick-based navigation in VR has usually an increased motion sickness (Langbehn et al., 2018).

Custom questions about the feeling of the haptic feedback revealed no significant difference

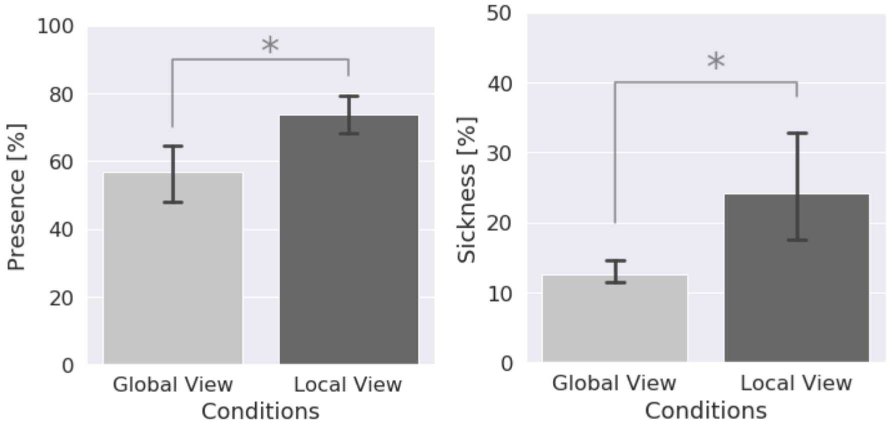


Figure 10.10: Presence and sickness percentage results in different VR conditions. * indicates significance at $p < 0.01$.

between the conditions, suggesting that the change in perspective doesn't affect the perception of force feedback.

Quantitative data from the movement of the robots were also recorded. Metrics such as task completion times ($H = 4.27, p = 0.11$) and the distance of the user-controlled robot to the swarm center were calculated ($H = 2.18, p = 0.33$). No significant difference was identified, suggesting that performance in the task is not affected by the change to VR environment, nor by the change in the perspective. However the experiment was too short (learners took around 70 seconds to complete the task in each condition) to allow any meaningful inferences regarding the performance aspect.

Conclusion

The proposal to combine tangible robot swarms with Virtual Reality (VR) enables the observation and interaction with robot swarms from multiple perspectives. In a small pilot study, we verified that the perspectives we implemented were perceived as different by learners. We observed a higher presence in the local condition, however, with the drawback of an increase in motion sickness. The proposed framework opens up opportunities to design more immersive learning activities, enabling users to switch roles and perspectives, and experience either the local interaction between agents or observe the resulting emergent global behavior. We hypothesize that experiencing swarm behaviors from multiple perspectives (top/local) is beneficial for learning the micro-macro link of emergent behavior. So it would be interesting to examine the relative benefits of the local and global views for learning in future studies.

10.5 Take-Aways

In this chapter, we present several extensions to increase the pool of interaction modalities between users and the Cellulo platform. The extensions include the addition of two reconfigurable user interaction modules to enable creating a reconfigurable swarm of robots, an online connection that enables global human-robot-robot-human interaction, and the integration with virtual reality to provide different perspectives. These extensions demonstrate the continued evolution of the Cellulo platform towards more interactive and versatile interaction experiences.

11 Synthesis and Contributions

11.1 Overview

This thesis aimed to explore the potential of utilizing swarm robots as a tool to foster complex systems understanding. To achieve this objective, the thesis proposes a comprehensive framework bridging multiple domains of learning sciences, robotics, and human-computer interaction.

Part I, being the core thread of the thesis, contributes to our understanding of assessing and fostering complex systems understanding through innovative learning activities that utilize physical and virtual agents. The proposed instrument in Chapter 3 provides a means of evaluating a person's competence in ontological concepts essential for understanding complex systems. Next we present *Cellulan World* which is designed as a novel learning framework that provides a tangible and interactive game-based learning experience addressing complex-systems related misconceptions, and thus stands on the theory of conceptual change. The learning activity, presented in Chapter 4 and evaluated in Chapters 5 and 6, demonstrates the effectiveness of a game that integrates both virtual and physical agents for fostering complex systems understanding across different age groups. Chapter 5 provides evidence for the effectiveness of using the *Cellulan World* game as a learning tool for promoting understanding of complex systems. Results showed a mean positive relative learning gain of 11.8%. It was demonstrated that incorporating physical robots in the learning activities resulted in higher levels of enjoyment and cognitive engagement, reinforcing intrinsic motivation. However, physical robots did not necessarily lead to a significant difference in learning gain compared to virtual modalities. Later in Chapter 6, we demonstrated the *Cellulan World* learning activity as being an example of an activity that fosters complex systems understanding possibly outside of traditional educational settings. Learners of all ages can benefit from the activity as it helps develop an essential skill across disciplines and across age levels. The results showed an average positive relative learning gain of 13.1% observed across all age groups, (between 15 and 50 years old in the test sample) which is a promising outcome for the effectiveness of the designed learning activity for individuals of varying ages, including those outside of formal education settings.

In Part II, we investigated group activities in a setting we named “double swarm” where multiple humans interact with multiple robots. In Chapter 7, we presented a combined physical and virtual activity on virus propagation, which is relevant to the COVID-19 pandemic, conducted in a real classroom setting. The activity resulted in a positive learning gain of 26.87% regarding the understanding of virus propagation emergent behavior. This can be considered as an initial step in deploying a collaborative activity geared towards understanding a complex system, specifically virus propagation, within a classroom environment. In an online multiplayer game presented in Chapter 8, we studied the effect of different communication affordances on a collaborative task to understand the role played by the communication medium and scope on multi-agent coordination, the emergence of communication systems, and team performance. We believe that our game has the potential to be a participatory learning activity for complex systems understanding, complementary to the Cellulan World activity discussed in Part I. As different behaviors emerge with different communication affordances, experiencing different levels of the game with different communication affordances, along with discussing the results and replays at the classroom level, can help learners reflect on different types of strategies (centralized vs decentralized) depending on the means at their disposal (communication affordance).

In Part III, we introduced a novel upgrade to the Cellulo platform, resulting in the creation of the Modulo Cellulo platform in Chapter 9. The upgrade provides a software development kit for developers to easily deploy activities and games, while the modularity of the robots enhances the versatility of the platform. This upgrade allows for easy deployment of activities and games, making it a practical and ubiquitous platform that can seamlessly integrate into various learning environments. Additionally, we present (in Chapter 10) several extensions to increase the pool of interaction between users and the Modulo Cellulo platform. The extensions include the addition of two reconfigurable user interaction modules, an online connection for global human-robot-robot-human interaction, and the integration with virtual reality to provide a different perspective. These extensions demonstrate the continued evolution of the Modulo Cellulo platform towards more interactive and versatile experiences, opening up opportunities for more immersive learning activities.

11.2 Limitations

Although this thesis presents promising findings, there are limitations that need to be acknowledged.

COVID-19 pandemic

The COVID-19 pandemic has had a profound impact on research, particularly on studies that require in-person interactions and the use of tangible robots. This, in turn, has been a significant limitation for the research presented in this thesis. The need for physical distancing measures and restrictions on in-person gatherings forced us to shift from in-person experiments to online and virtual ones. While this shift to online and virtual experiments was necessary due to the pandemic, it is essential to consider the potential limitations and challenges that arise from this change. This shift made it difficult to conduct extensive studies with the tangible part of the interaction in our research.

Novelty Effect

One potential limitation of our research is the novelty effect, which is characterized by a temporary increase in performance or engagement when participants are exposed to a new intervention or treatment (R. E. Clark, 1985). Due to the fact that our studies were conducted over only one session of approximately one hour, it was challenging to control for this phenomenon. While testing the effectiveness of educational interventions on students' long-term learning is an important aspect, single-session format are a necessary precursor to longitudinal studies.

In our *Cellulan World* study, a limited number of tasks were tested due to time constraints. However, the study design was structured in a way that allows for easy expansion into multiple sessions, including the incorporation of flocking, foraging, and synchronization behaviors. Future work in this area would provide an opportunity to more thoroughly examine the impact of incorporating these behaviors on the overall effectiveness of the intervention. Our results have shown significant effects, particularly in two of the concepts related to complex systems understanding: Control and Agent Effect. However, expanding to further tasks and behaviors could provide a more robust examination of the other concepts, resulting in a more complete and complementary set of activities. Similarly, different modalities (physical vs virtual) seem to support learning of different scenarios. Consequently, future work could also focus on exploring this finding more extensively and constructing a set of activities which integrate both modalities.

Sample Size and Participant Background

Another limitation of the studies presented in this thesis is the relatively small sample size, and the fact that participants for the physical interaction studies were drawn from a specific socio-cultural background - international school students in Switzerland. This limitation arose due to logistical constraints in organizing experiments in schools, which may have influenced the results and limit the generalizability of the findings to other populations. Additionally, the unbalanced sample size may have limited the statistical power of the studies and made it more difficult to detect significant effects. For example, in the first study of *Cellulan World* in Chapter 5, there was an imbalance of group size between the two modalities, with the virtual modality having a larger sample size. Furthermore, the fact that students were not randomized over schools means that some background factors may have been confounding variables. In the second study across age groups, there was a skewed age distribution of participants, and more data-points with a wider range of majors, backgrounds, and experiences are required among the older population.

To address these limitations, future research should aim to recruit larger and more diverse samples from a wider range of schools and backgrounds. By doing so, we can ensure that our findings are more representative and applicable to a broader range of populations. This can also help to increase the statistical power of the studies, making it easier to detect significant effects. Additionally, randomizing participants over schools could help reduce the impact of confounding variables such as educational and social background, while including participants with a broader range of ages, majors, backgrounds, and experiences could provide a more comprehensive understanding of the interaction.

11.3 Contributions Summary

Similar to how the concept of complex systems is a powerful and versatile idea that applies to multiple fields, this research has implications for a range of research areas, such as learning sciences, robotics, and human-computer interaction.

In the field of learning sciences, we have contributed with:

- An assessment instrument for understanding complex systems (presented in Chapter 3). The instrument was validated through a rigorous process grounded in literature, an expert study, and an iterative design to refine it. The final outcome is a questionnaire accompanied by a coding scheme that is available for the community to use and incorporate into their research.
- A design for a learning activity that incorporates virtual and physical agents and is based on the conceptual change theory (Chapter 4). The learning activity has been shown to be effective in both modalities and across different age groups (Chapters 5 and 6).
- An understanding of the learning processes that supports complex systems understanding though behavioral, performance and learning analysis (Chapters 5 and 6).

In the field of robotics research, we have contributed with:

- A framework for human-robot swarm interaction in an educational context, contributing to the research of swarm robots for learning.
- The upgrade and development of a modular educational platform for building robots. The resulting output is available to the community for building their own robots, as well as using the developed SDK to build their own games and activities (Chapter 9).
- The expansion of the interaction pool for tangible robots by providing a reconfigurable swarm robotic platform, an online connection for expanding interaction across the

globe, and a VR connection to robots adding tangible interaction to the virtual world, as well as adding the possibility of different perspective taking in a tangible interaction (Chapter 10).

Lastly, within the domain of human-computer interaction, we have contributed with:

- Incorporating learning sciences principles to design productive interactions for complex systems understanding by putting the learner at the center of the design process (Chapters 4 and 7).
- Bridging with swarm intelligence, in an online multiplayer game, we have shown the importance of considering the different affordances available for communication and collaboration between humans, as a global communication channel may not always be the optimal way to solve a collaborative task (Chapter 8).

11.4 Broadening the Horizon

The contributions of this thesis open various exciting avenues for future research, which can further our understanding of complex systems learning, the interaction between humans and swarm robots, and the design and implementation of effective learning activities.

The learning activities developed in this thesis have shown promising results in improving complex systems understanding. Future work can delve deeper by expanding these activities, introducing more tasks and behaviors related to complex systems. This will enable a broader and more comprehensive exploration of complex systems concepts and principles, potentially leading to more significant learning gains.

Another promising avenue to investigate further is the interplay between educational backgrounds, cultural influences, societal impacts, and diverse age groups. Understanding these intersections can provide insights into how these factors influence complex systems learning.

Furthermore, one research direction, which was initially intriguing but became restricted due to pandemic-related constraints, involves delving into the role and effect of haptic feedback within the context of human-robot-robot-human interactions and extending this into multi-human and multi-robot interactions. Preliminary results from this thesis have underscored its impact on collaboration strategies, making it particularly intriguing to build upon this area and incorporate it as a possible communication affordance in human-swarm interaction scenarios.

Finally, the enhancements to the Cellulo platform - incorporating modularity, the SDK, and the extended possibility of interaction through VR, reconfigurable robots, and online connection - open up vast possibilities for a wide range of interactions and applications across multiple domains. The platform's versatility also lends itself to various educational settings, such as online or informal learning environments, widening the scope of its potential impact.

In closing, I would like to reflect on the idea that academic research is also a perfect example of a complex system with a multitude of interconnected entities, including institutions, researchers, professors, and students from around the world. Through scientific communication, collaboration, and even competition, each entity contributes to the field in their own way, leading to the emergent behavior of scientific advancement. Similarly, this thesis builds on existing literature and we hope that our contributions have provided the community with new tools and insights to advance the fields of human-swarm robot interaction and human learning of complex systems.

At times, it seems daunting and overwhelming to understand the complex systems we are a part of, and we may be tempted to give up and lose motivation. However, if I have to give one insight after being immersed in the topic of this thesis for a few years, it would be the importance of recognizing the significance of our individual contributions. Even the smallest contributions can make a significant impact on the whole, and collectively, we can advance science and foster our understanding of the fascinating nature of complex systems.

A Assessment Instrument Material

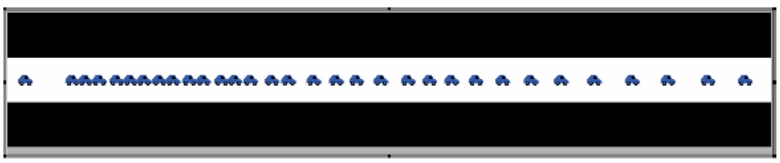
A.1 Questionnaire

Link to the questionnaire: https://forms.gle/KNwb7qDB7D1g7wY36

Why do traffic jams form?

On the road, each car followed only three rules:

1. *If there is a car close ahead of you, slow down.*
2. *If there aren't any cars close ahead of you, speed up (unless you are already moving at the speed limit).*
3. *Comply with driving regulations concerning all other elements on the road, such as: slow down if you detect a radar trap, traffic lights, accidents, entry ramps...*



1) What is causing the traffic jam? Select all applicable choices

☐ Radar trap

☐ Broken Bridge

☐ Entry Ramp with merging traffic

☐ Accident

☐ Nothing

☐ Other: ——————————-

2) Explain in your own words how your choice(s) is(are) causing the traffic jam.
——— Your answer ———

3) The radar trap is removed, and there are no accidents on the road, no broken bridge, no entry ramp or any other external events. Can a traffic jam still form?

☐ No, there is no way for a traffic jam to form

☐ Yes, a traffic jam can still form.

4) If yes, why do you think a traffic jam can still form ? If No, explain in your own words, how the cars will behave on the road.
——— Your answer ———

Scatter

Let's say you are going to the gym class. You're all standing there in a huddle waiting for instructions because you don't know what you're going to do today. Your instructor comes and says: "OK, we're going to do some warm up jumping jacks so I want you to scatter.

1) How do you know where to move?
——— Your answer ———

2) What does each person need to do in order to scatter?
——— Your answer ———

3) A new person comes to class from a town where they don't have a gym. Just before gym class, you have a chance to explain to this person what will happen when the coach tells you to scatter. How would you explain to this new person what to do when the teacher asks you to spread out?
——— Your answer ———

4) If you were up really high, and could watch your class scatter, what would you see from above when they're scattering? How would you describe this to someone that wasn't there?
——— Your answer ———

Flocks of Birds

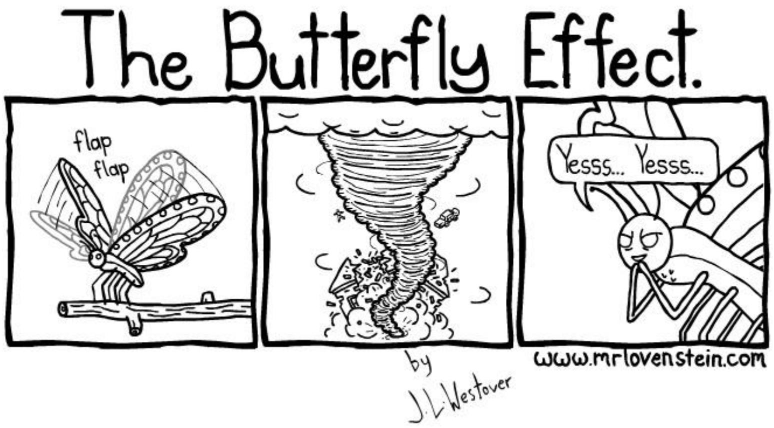
1) How is it that these birds fly in a flock? [Fill all that applies]

- ☐ There is a leader bird, the other birds follow it
- ☐ The birds are attracted to their neighbors
- ☐ The birds avoid collision with their neighbors
- ☐ The birds follow the average speed of their neighbors
- ☐ Other: ——————————-

2) Explain your answer(s) in your own words.

———— Answer ————

Butterfly Effect



It has been said that a butterfly flapping its wings in Brazil can jiggle the air and thus can help cause a snowstorm in Alaska.

1) More generally the butterfly effect is a metaphor explaining how small changes can produce large effects. Is this possible?

☐ Yes

☐ No

2) Please elaborate your choice: If yes, can you give an example of a scenario in which small changes lead to large effects. If no, explain why it is not possible.

——— Your answer ———

Robots and Gold

Large deposits of gold have been discovered on a distant planet. It is too dangerous and costly to send human astronauts to this planet, so we decide to send a spaceship with several thousand small robots. Each robot has a sensor to detect when it gets near gold, and a scoop to dig for and carry the gold. Once the spaceship lands on the planet, we want the robots to explore for gold and bring the gold back to the spaceship. Consider that a robot might break at any moment due to the unknown dangers of the environment, and similarly, communication might occasionally be disrupted. What type of rules and strategies should the robots follow?

1) How should the robots move in order to find the gold?

- ☐ Spread out randomly
- ☐ Divide robots into groups with leads and subleads
- ☐ Spaceship assigns a target area for each of the robots
- ☐ Other: ——————————-

2) What reasons led you to this choice? Elaborate on it.
————— Your answer —————

3) How would a robot inform the others if it found gold ?

- ☐ Returns back to the spaceship while leaving marks on the way
- ☐ Communicate with nearby robots and inform them
- ☐ Reports back to the spaceship
- ☐ The robots have a global communication capabity, they send a message to everyone
- ☐ Other: ——————————-

4) What reasons led you to this choice? Elaborate on it.
————— Your answer —————

A.2 Coding Scheme

Link to the coding scheme: https://forms.gle/ogPPzEGRsoBgTyHr5

Traffic Jam

Concept	Questions	Score 0	Score 0.5	Score 1
Control	3,4	Traffic jam doesn't form or still mentioning another external event	-	Traffic can still form with no other external event
Causes	1,2,4	if single cause in 1,2 and 3	if multiple causes are chosen in 1 & the explanation in 2 and 4 mentions a single cause OR if a single cause is chosen in 1, and the explanation in 2 and 3 mentions multiple causes.	if multiple causes are chosen in 1, and explanations in 2 and 3 mentions multiple causes.
Actions	2,4	mention of a car slowing down/speeding up without explicit mention of effect OR no mention of speed change effects	-	explicit mention of small changes in a car's speed and explicit mention of how they affect the speed of other cars.

Scatter

Concept	Questions	Score 0	Score 0.5	Score 1
Control	1,3	1: Follow instructions from coach, 3: follow me or follow instructor or lack of explanation	a mix between the two	1: check distances from other and 3: given with respect to others.
Agents	1,2,3	move to specific points	-	move with respect to other people. what's important is the relation between the agents, i.e, the connection of one's movement and the movement of the others.
Process	4	Description based only on global/macro level (for example expanding circle..) OR incomplete (example: random movement, chaotic)	-	Description based on micro-behaviors of agents AND relation to global/macro behavior (for example: agent moving randomly until they are well spread). Valid analogies are acceptable even not explained.

Flock of Birds

Concept	Questions	Score 0	Score 0.5	Score 1
Control	1,2	if they select leader (even if with other choices) unless see score 0.5 and energy related answers.	Leader selected + other choices and the justification the leader is just the first in the line but not organizational powers	no leader .
Causes	1	single cause selected (also if they write multiple things but refer to same cause).	-	multiple causes selected.
Agents	2	movement predictable, i.e: follow leader, V shape ..	-	movement based on the interactivity with neighbors (for example: agent checking distance to neighbors. and/or interactivity with environment (air fly in).

Butterfly Effect

Concept	Questions	Score 0	Score 0.5	Score 1
Actions	1,2	Absolute no	Maybe/describing a chain of cause-effects with no change in magnitude.	yes, also mentions non-linearity in the explanation.
Process	2	otherwise	-	mention of how the interaction between different agents generates the effect of the macro-behavior.

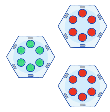
Robots and Gold

Concept	Questions	Score 0	Score 0.5	Score 1
Control	1,2,3,4	leader + global communication	leader + local communication or distributed + global communication	distributed
Agents	1,2	predictable movement of the robots (maps, specific position)	if divide in groups but no mention of how they explore	unpredictable/random movement of the robots (even with diving in groups, so 'groups moving randomly')

B Supplementary Material for Chapter 7

Classe 9 VP MAT (13-14 ans) PROPAGATION / CELLULOS Prénom: ....

LES CELLULOS

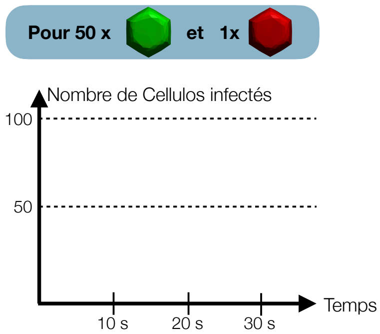


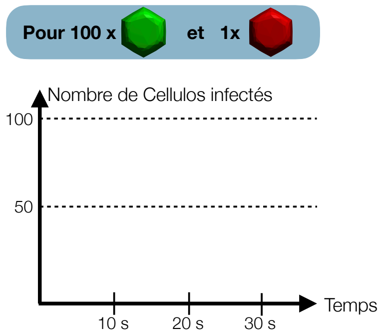
Objectif

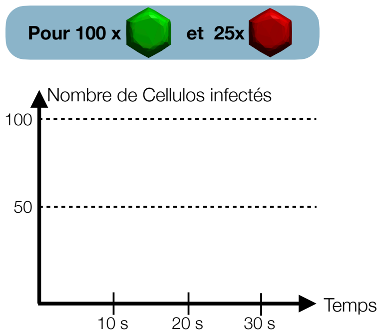
Comprendre les enjeux des différentes propagations pour un virus/rhume/rumeur.

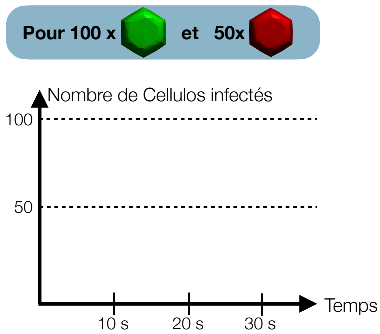
1. Simulation

A l'aide des simulations sur la Tablette, dessine l'évolution des Cellulos infectés en fonction du temps.



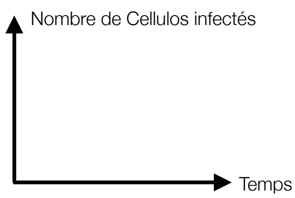






Résumé

Esquisse à la main l'allure de la courbe en général :



a) Décris en quelques mots pourquoi la courbe à tendance à avoir une forme de **S**:

b) Pourquoi le nombre d'infection augmente-t-il jusqu'à un certain point, puis ralenti vers la fin ?

A Cocher:

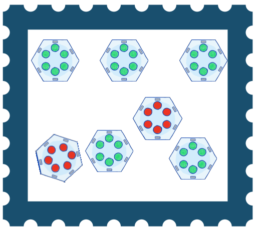
c) Si initialement on augmente le *nombre de Cellulos*, alors le temps pour que tous les Cellulos soient infectés est :

- ☐ Plus rapide
- ☐ Plus lent
- ☐ Inchangé

d) Si initialement on augmente le *nombre de Cellulos* *infectés*, alors le temps pour que tous les Cellulos soient infectés est :

- ☐ Plus rapide
- ☐ Plus lent
- ☐ Inchangé

2. Réel avec les Cellulos



a) Que se passe t-il lorsque les Cellulos se touchent ? Et pour ceux infectés ?

b) Pourquoi les Cellulos infectés ne transmettent-t-ils pas toujours l'infection aux autres ?

Classe 9 VP MAT (13-14 ans) Math/Propagation Prénom: ....

PRE-QUESTIONNAIRE

Relie les textes avec les graphiques correspondants :

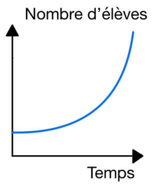
Dans une classe de 24 élèves:

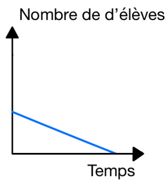
On double le nombre d'élèves chaque jour. Le nombre d'élèves est de:

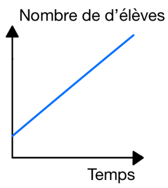
On ajoute chaque jour un élève. Le nombre d'élèves est de:

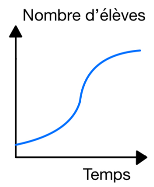
On enlève chaque jour un élève. Le nombre d'élèves est de :

1 élève a le rhume. Le rhume se transmet à travers toute la classe. Le nombre d'élèves infectés est de:









A Cocher:

Pour une classe de 24 élèves en salle 108: Le temps qu'une rumeur se propage à tous les élèves est rapide.

a) Si l'on **double le nombre d'élèves (toujours en salle 108)**, alors le temps que la rumeur se propage à tous les élèves est:

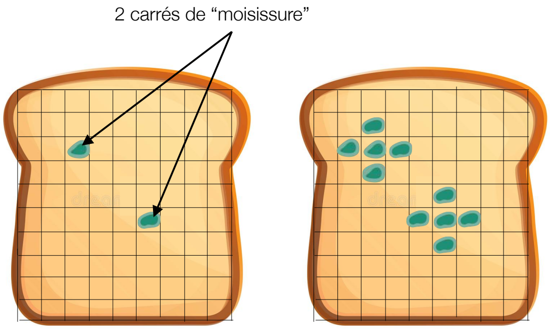
- ☐ Plus rapide
- ☐ Plus lent
- ☐ Inchangé

b) Si cette rumeur est déjà connue par la **moitié de la classe**, le temps que la rumeur se propage à tous les élèves est:

- ☐ Plus rapide
- ☐ Plus lent
- ☐ Inchangé

Classe 9 VP MAT (13-14 ans) | Math/Propagation | Prénom: ....

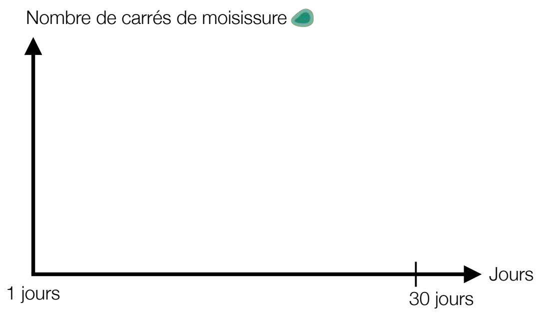
Nicolas a un toast avec 2 petits carrés de moisissure (voir le toast ci-dessous). Il a remarqué que chaque jour de nouveaux carrés de moisissure s'ajoutent et qui touchent toujours les anciens carrés moisis. Voici les deux premiers jours du toast de Nicolas:



...

Jour 1 | **Jour 2** | **Jour 3**

Dessine à la main sur le graphique ci-dessous l'évolution du nombre de carrés de moisissures :



Classe 9 VP MAT (12-13 ans) Math/Propagation Prénom: ....

POST-QUESTIONNAIRE

A Cocher:

Pour une classe de 24 élèves en salle 108: Le temps qu'un rhume se propage à tous les élèves est rapide.

a) Si la **moitié de la classe** a déjà ce rhume alors le temps pour qu'il se propage à tous les élèves est:

- ☐ Inchangé
- ☐ Plus lent
- ☐ Plus rapide

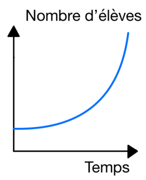
b) Si l'on **double le nombre d'élèves (toujours en salle 108)**, alors le temps que le rhume se propage à tous les élèves est:

- ☐ Inchangé
- ☐ Plus lent
- ☐ Plus rapide

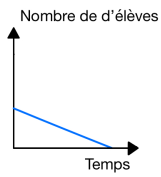
Relie les textes avec les graphiques correspondants :

Dans une classe de 24 élèves:

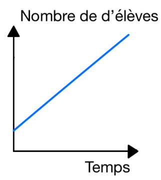
On ajoute chaque jour un élève. Le nombre d'élèves est de:



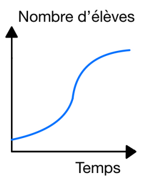
On double le nombre d'élèves chaque jour Le nombre d'élèves infecté est de:



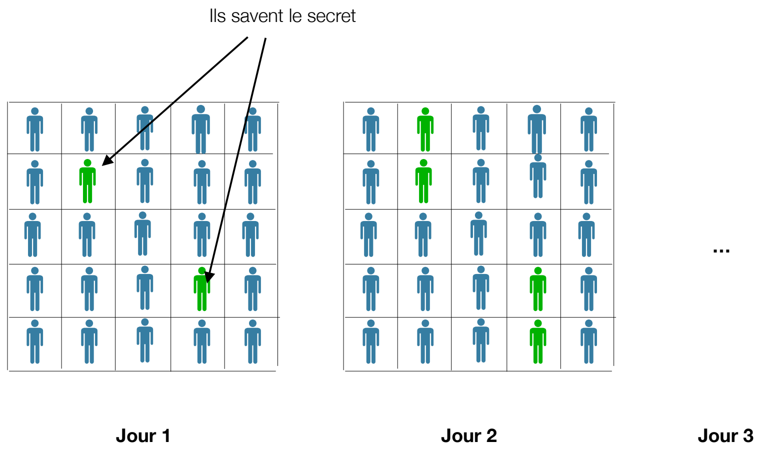
1 élève a le rhume. Le rhume se transmet à travers toute la classe. Le nombre d'élèves infectés est de:



On enlève chaque jour un élève. Le nombre d'élèves est de :



Dans une classe de 25 élèves:
Lundi, Nicolas et Marie savent un **secret** sur la classe 9 VP6. Chaque jour, Marie et Nicolas vont le dire chacun à un de leur camarade de classe, qui vont eux aussi le redire le lendemain à un camarade, et ainsi de suite. Voici les deux premiers jours:



Dessine à la main sur le graphique ci-dessous l'évolution du nombre de personnes qui savent le secret :

Figures Géométriques Planes

Mets une croix dans la colonne que tu choisis.

		Pas du tout d'accord	Pas d'accord	Plutôt pas d'accord	Ni d'accord ni en désaccord	Plutôt d'accord	D'accord	Tout à fais d'accord
0	J'aime bien les maths.							
1	J'ai trouvé cet exercice facile							
2	J'ai essayé de faire cet exercice aussi bien que possible.							
3	Je suis sûr de mes réponses pour l'exercice.							
4	j'ai discuté avec mon camarade pour résoudre							
5	J'ai trouvé cet exercice intéressant.							
6	J'ai bien réussi cet exercice.							
7	Je me suis concentré(e) dans cet exercice.							
8	J'ai discuté/ collaboré avec mon camarade pour trouver la réponse.							
9	J'aimerai refaire des activités de ce genre à l'avenir pour d'autres cours de math OS							
10	J'aimerai refaire des exercices de ce genre dans d'autres matières.							

Autres remarques: ………………………………………………………………………………

C Exemplar quotes from participants' responses

We thoroughly enjoyed going through the participants' innovative responses, and we'd like to share a few of our personal favorites.

Assessment Responses

Butterfly Effect Scenario

- If we consider for exemple, a butterfly is flapping it's wings and moving, having wonderful colors, it will attract the attention of a person crossing a road and stops for a split second in the middle of the road , a car coming his way has to slow down fast, it deviated a little bit and hit a pole on the side, no one is injured just the car a bit damaged. While waiting to the truc to arrive, a jam was created, affecting the advancement of the carnaval, millions of people needed to slow down, but the sun is high and the temperature is rising, millons of people are sweating ejecting water from there bodies and evaporating to cool off. The accumulation of humidity and vapors rise and form clouds, the rotation of earth move the clouds around, and after few cycles, clouds arriving above alaska with a powerful wind and because of very low temperatures a snowstorm starts. Nailed it !
- If everyone in a school rely on one person for homework, when that person leaves or doesn't give homework anymore, everyone will not be able to do the homework leading to a massive problem.
- A small decision, like going to the grocery store on Tuesday instead of Monday, can lead to finding a partner and getting married. That would be an example.
- If you do something new but small like go out with a new guy or other can change for ever your future
- If someone droppes a stone on a path, a runner may stub their toe on it and not run for a week. They may then not meet a potential partner on their usual route, not begin a

relationship, not have children etc.

- just being at there at the right time or bumping into someone could completely change your life as it could be a big job offer, soulmate etc.
- If the first humans on earth decided to never reproduce humanity would have never existed or be very different.
- You forgot your bag at home, you drive back to go get your bag, you are in a car accident and die. Yes, it's a bit dark but you died simply because you forgot your keys.
- For example if you kick a small rock it is a small change in the environment. But someone else might trip on that little rock and they are holding a coffe so they spill it on themselves and now they have to go to work with a spill on them which means they will have a bad day. They might be a politician that have an important voting day and their choice was a deciding one and because they have a bad they they decided to make it everyone else problem and they vote against a policy that would improve many lives because they would rather vote for one that fixes crosswalks.

Robots and Gold Scenario

- It is dangerous to let robots communicate among themselves.

Flock of Birds Scenario

- I'm not really sure, I don't care about birds that much.

Game Feedback

- that was an interesting experience, our president should try to play cellulans ;)
- It was a fun game.
- GOOD STUFF
- This was a really cool experience and game. I really enjoyed it!
- I loved the game - it's very nice and intuitive! And the graphics are surprisingly nice :)

Bibliography

Abrahamson, D., Nathan, M. J., Williams-Pierce, C., Walkington, C., Ottmar, E. R., Soto, H., & Alibali, M. W. (2020). The future of embodied design for mathematics teaching and learning. *Frontiers in Education*, *5*, 147.

(Cited on page 23)

Alhafnawi, M., Hunt, E. R., Lemaignan, S., O'Dowd, P., & Hauert, S. (2022a). Deliberative democracy with robot swarms. *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 7296–7303. https://doi.org/10.1109/IROS47612.2022.9981649.

(Cited on page 103)

Alhafnawi, M., Hunt, E. R., Lemaignan, S., O'Dowd, P., & Hauert, S. (2022b). Mosaix: a swarm of robot tiles for social human-swarm interaction. *2022 International Conference on Robotics and Automation (ICRA)*, 6882–6888. https://doi.org/10.1109/ICRA46639.2022.9811723.

(Cited on page 13)

Åström, K. J., & Wittenmark, B. (2013). *Adaptive control*. Courier Corporation.

(Cited on page 138)

Auerbach, J., Aydin, D., Maesani, A., Kornatowski, P., Cieslewski, T., Heitz, G., Fernando, P., Loshchilov, I., Daler, L., & Floreano, D. (2014). Robogen: robot generation through artificial evolution. *ALIFE 14: The Fourteenth International Conference on the Synthesis and Simulation of Living Systems*, 136–137.

(Cited on page 7)

Barnes, J., FakhrHosseini, S. M., Vasey, E., Park, C. H., & Jeon, M. (2020). Child-robot theater: engaging elementary students in informal steam education using robots. *IEEE Pervasive Computing*, *19*(1), 22–31. https://doi.org/10.1109/MPRV.2019.2940181.

(Cited on page 14)

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, *1*, 71–81.

(Cited on page 81)

Batra, S., Klingner, J., & Correll, N. (2022). Augmented reality for human–swarm interaction in a swarm-robotic chemistry simulation. *Artificial Life and Robotics*, *27*(2), 407–415.

(Cited on page 14)

Bayindir, L., & Şahin, E. (2007). A Review of Studies in Swarm Robotics. *Turkish Journal of Electrical Engineering & Computer Sciences*, *15*(2), 115–147.
(Cited on pages 7, 116)

Begel, A., & Klopfer, E. (2007). Starlogo tng: an introduction to game development. *Journal of E-Learning*, *53*(2007), 146.
(Cited on page 18)

Belke, C. H., & Paik, J. (2017). Mori: a modular origami robot. *IEEE/ASME Transactions on Mechatronics*, *22*(5), 2153–2164.
(Cited on page 144)

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa.
(Cited on page 73)

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.
(Cited on page 43)

Caprari, G., Arras, K. O., & Siegwart, R. (2000). The autonomous miniature robot Alice: from prototypes to applications. *Proceedings. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000)*, *1*, 793–798 vol.1. https://doi.org/10.1109/IROS.2000.894701.
(Cited on pages 12, 13)

Chi, M. T. H., Roscoe, R. D., Slotta, J. D., Roy, M., & Chase, C. C. (2012). Misconceived Causal Explanations for Emergent Processes [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1551-6709.2011.01207.x]. *Cognitive Science*, *36*(1), 1–61. https://doi.org/10.1111/j.1551-6709.2011.01207.x.
(Cited on page 4)

Chi, M. T. (2005). Commonsense conceptions of emergent processes: why some misconceptions are robust. *The journal of the learning sciences*, *14*(2), 161–199.
(Cited on pages 4, 34)

Cianci, C. M., Raemy, X., Pugh, J., & Martinoli, A. (2007). Communication in a swarm of miniature robots: the e-puck as an educational tool for swarm robotics. In E. Şahin, W. M. Spears, & A. F. T. Winfield (Eds.), *Swarm robotics* (pp. 103–115). Springer Berlin Heidelberg.
(Cited on pages 12, 13)

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: a systematic review and meta-analysis. *Review of educational research*, *86*(1), 79–122.
(Cited on page 23)

Clark, R. E. (1985). Confounding in educational computing research. *Journal of Educational Computing Research*, *1*(2), 137–148. https://doi.org/10.2190/HC3L-G6YD-BAK9-EQB5.
(Cited on page 161)

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). Are games effective learning tools? a review of educational games. *Journal of educational psychology*, *104*(2), 469.

(Cited on page 22)

Cook, M., Seo, J. H., Pine, M., & Sungkajun, A. (2018). Paper circuitry and projection mapping: an interactive textbook approach to veterinary education. *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, 131–135. https://doi.org/10.1145/3173225.3173290.

(Cited on page 82)

Coucke, N., Heinrich, M. K., Cleeremans, A., & Dorigo, M. (2021). HuGoS: a Virtual Environment for Studying Collective Human Behavior from a Swarm Intelligence Perspective. *Swarm Intelligence*, *15*(4), 339–376.

(Cited on page 117)

Crutcher, R. J. (1994). Telling what we know: the use of verbal report methodologies in psychological research. *Psychological Science*, *5*(5), 241–241. https://doi.org/10.1111/j.1467-9280.1994.tb00619.x.

(Cited on page 36)

DeLiema, D., Enyedy, N., & Danish, J. A. (2019). Roles, Rules, and Keys: How Different Play Configurations Shape Collaborative Science Inquiry. *Journal of the Learning Sciences*, *0*(0), 1–43. https://doi.org/10.1080/10508406.2019.1675071.

(Cited on page 21)

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

(Cited on page 44)

DiSessa, A. A., et al. (2013). A bird's-eye view of the" pieces" vs" coherence" controversy (from the" pieces" side of the fence"). *International handbook of research on conceptual change*, 31–48.

(Cited on page 26)

diSessa, A. A. (2014). A history of conceptual change research: threads and fault lines. In R. K. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (2nd ed., pp. 88–108). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.007.

(Cited on page 24)

diSessa, A. A. (2022). A history of conceptual change research: threads and fault lines. In R. K. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (3rd ed., pp. 114–133). Cambridge University Press. https://doi.org/10.1017/9781108888295.008.

(Cited on page 25)

Do, L. H. S. (2012). *Supporting reflection and classroom orchestration with tangible tabletops* (tech. rep.). EPFL.

(Cited on page 79)

Dorigo, M., Theraulaz, G., & Trianni, V. (2020). Reflections on the future of swarm robotics. *Science Robotics*, *5*(49), eabe4385.

(Cited on page 7)

Dourish, P. (2001). *Where the action is: the foundations of embodied interaction*. MIT press.

(Cited on page 23)

Driankov, D. (2001). A reminder on fuzzy logic. *Fuzzy Logic Techniques for Autonomous Vehicle Navigation*, 25–47.

(Cited on page 41)

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, *100*(3), 363.

(Cited on page 77)

Etkina, E., Karelina, A., Ruibal-Villasenor, M., Rosengrant, D., Jordan, R., & Hmelo-Silver, C. E. (2010). Design and reflection help students develop scientific abilities: learning in introductory physics laboratories. *The Journal of the Learning Sciences*, *19*(1), 54–98.

(Cited on page 79)

Farinelli, A., Iocchi, L., & Nardi, D. (2004). Multirobot systems: a classification focused on coordination. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *34*(5), 2015–2028. https://doi.org/10.1109/TSMCB.2004.832155.

(Cited on page 11)

Gianaros, P. J., Muth, E. R., Mordkoff, J. T., Levine, M. E., & Stern, R. M. (2001). A Questionnaire for the Assessment of the Multiple Dimensions of Motion Sickness. *Aviation, space, and environmental medicine*, *72*(2), 115–119.

(Cited on page 155)

Glenberg, A. M. (2008). Embodiment for education. In *Handbook of cognitive science* (pp. 355–372). Elsevier.

(Cited on page 106)

Gotwals, A. W., & Songer, N. B. (2010). Reasoning up and down a food chain: Using an assessment framework to investigate students' middle knowledge [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/sce.20368]. *Science Education*, *94*(2), 259–281. https://doi.org/10.1002/sce.20368.

(Cited on page 3)

Gourlet, P., Le Goc, M., & Follmer, S. (2017). Revisiting *Turtles* and *Termites*: an Open-ended Interactive Physical Game with Multiple Robots. *Proceedings of the 2017 Conference on Interaction Design and Children - IDC '17*, 679–682. https://doi.org/10.1145/3078072.3091979.

(Cited on page 15)

Grotzer, T. A., Powell, M. M., M Derbiszewska, K., Courter, C. J., Kamarainen, A. M., Metcalf, S. J., & Dede, C. J. (2015). Turning transfer inside out: the affordances of virtual worlds and mobile devices in real world contexts for teaching about causality across time and distance in ecosystems. *Technology, Knowledge and Learning*, *20*, 43–69.

(Cited on pages 4, 34, 43)

Grotzer, T. A., & Shane Tutwiler, M. (2014). Simplifying Causal Complexity: How Interactions Between Modes of Causal Induction and Information Availability Lead to Heuristic-Driven Reasoning [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mbe.12054]. *Mind, Brain, and Education*, *8*(3), 97–114. https://doi.org/10.1111/mbe.12054.

(Cited on page 4)

Guneysu, A., Wessel, M. J., Johal, W., Sharma, K., Özgür, A., Vuadens, P., Mondada, F., Hummel, F. C., & Dillenbourg, P. (2018). Iterative design of an upper limb rehabilitation game with tangible robots. *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, 241–250.

(Cited on page 141)

Guneysu Ozgur, A., Özgür, A., Asselborn, T., Johal, W., Yadollahi, E., Bruno, B., Skweres, M., & Dillenbourg, P. (2020). Iterative design and evaluation of a tangible robot-assisted handwriting activity for special education. *Frontiers in Robotics and AI*, *7*, 29.

(Cited on page 17)

Guo, Y., & Wilensky, U. (2016). Small bugs, big ideas: teaching complex systems principles through agent-based models of social insects. *IEEE Symposium on Artificial Life*.

(Cited on page 19)

Hake, R. R. (1998). Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *American journal of Physics*, *66*(1), 64–74.

(Cited on page 67)

Hengeveld, B., Hummels, C., van Balkom, H., Voort, R., & de Moor, J. (2013). Wrapping up linguabytes, for now. *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, 237–244. https://doi.org/10.1145/2460625.2460664.

(Cited on pages 82, 85)

Hilder, J., Horsfield, A., Millard, A. G., & Timmis, J. (2016). The psi swarm: a low-cost robotics platform and its use in an education setting. In L. Alboul, D. Damian, & J. M. Aitken (Eds.), *Towards autonomous robotic systems* (pp. 158–164). Springer International Publishing.

(Cited on pages 12, 13)

Hmelo-Silver, C. E. (2004). Problem-based learning: what and how do students learn? *Educational psychology review*, *16*, 235–266.

(Cited on page 79)

Hmelo-Silver, C. E., & Azevedo, R. (2006). Understanding complex systems: some core challenges. *The Journal of the learning sciences*, *15*(1), 53–61.

(Cited on pages 3, 33)

Hmelo-Silver, C. E., Marathe, S., & Liu, L. (2007). Fish swim, rocks sit, and lungs breathe: expert-novice understanding of complex systems. *The Journal of the Learning Sciences*, *16*(3), 307–331.

(Cited on pages 3, 33, 34)

Horn, M. S., Solovey, E. T., Crouser, R. J., & Jacob, R. J. (2009). Comparing the use of tangible and graphical programming languages for informal science education. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 975–984.

(Cited on page 82)

Hostettler, L., Özgür, A., Lemaignan, S., Dillenbourg, P., & Mondada, F. (2016). Real-time high-accuracy 2d localization with structured patterns. *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 4536–4543. https://doi.org/10.1109/ICRA.2016.7487653.

(Cited on pages 132, 137)

Hwang, W.-Y., & Wu, S.-Y. (2014). A case study of collaboration with multi-robots and its effect on children's interaction. *Interactive Learning Environments*, *22*(4), 429–443. https://doi.org/10.1080/10494820.2012.680968.

(Cited on page 103)

Ijspeert, A. J., Martinoli, A., Billard, A., & Gambardella, L. M. (2001). Collaboration through the exploitation of local interactions in autonomous collective robotics: the stick pulling experiment. *Autonomous Robots*, *11*, 149–171.

(Cited on page 15)

Jacobson, M. J. (2001). Problem solving, cognition, and complex systems: differences between experts and novices. *Complexity*, *6*(3), 41–49.

(Cited on pages 3, 35, 42)

Jacobson, M. J., Kapur, M., So, H.-J., & Lee, J. (2011). The ontologies of complexity and learning about complex systems. *Instructional Science*, *39*(5), 763–783.

(Cited on pages 35, 38, 43, 87)

Jacobson, M. J., & Wilensky, U. (2022). Complex systems and the learning sciences: educational, theoretical, and methodological implications. In R. K. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (3rd ed., pp. 504–522). Cambridge University Press. https://doi.org/10.1017/9781108888295.031.

(Cited on pages 18, 19)

Jeong, S., Elliott, J. B., Feng, Z., & Feldon, D. F. (2022). Understanding Complex Ecosystems Through an Agent-Based Participatory Watershed Simulation. *Journal of Science Education and Technology*, *31*(5), 691–705. https://doi.org/10.1007/s10956-022-09987-8.

(Cited on page 34)

Johal, W. (2020). Research trends in social robots for learning. *Current Robotics Reports*, *1*(3), 75–83.

(Cited on page 7)

Johal, W., Andersen, S., Chevalier, M., Ozgur, A., Mondada, F., & Dillenbourg, P. (2019). Learning symmetry with tangible robots. *International Conference on Robotics in Education (RiE)*, 270–283.

(Cited on page 17)

Johal, W., Peng, Y., & Mi, H. (2020). Swarm robots in education: a review of challenges and opportunities. *Proceedings of the 8th International Conference on Human-Agent Interaction*, 272–274. https://doi.org/10.1145/3406499.3418755.

(Cited on pages 7, 11, 15)

Johal, W., Tran, A., Khodr, H., Özgür, A., & Dillenbourg, P. (2019). Tip: tangible e-ink paper manipulatives for classroom orchestration. *Proceedings of the 31st Australian Conference on Human-Computer-Interaction.*

(Cited on page 141)

Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: two science studies. *Journal of Educational Psychology*.

(Cited on page 153)

Kapur, M. (2008). Productive failure. *Cognition and instruction*, *26*(3), 379–424.

(Cited on page 71)

Karim, M. E., Lemaignan, S., & Mondada, F. (2015). A review: can robots reshape k-12 stem education? *2015 IEEE International Workshop on Advanced Robotics and its Social Impacts (ARSO)*, 1–8. https://doi.org/10.1109/ARSO.2015.7428217.

(Cited on pages 7, 12)

Khodr, H., Brender, J., Kothiyal, A., & Dillenbourg, P. (2021). How diseases spread: embodied learning of emergence with cellulo robots. *Workshop proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education*, 263–272.

(Cited on page 105)

Khodr, H., Bruno, B., Kothiyal, A., & Dillenbourg, P. (2022). Cellulan world: interactive platform to learn swarm behaviors. *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, 1908–1910.

(Cited on page 49)

Khodr, H., Holdcroft, K., Wu, Y.-S., Borja, V., Sprumont, H., Bruno, B., Paik, J., & Dillenbourg, P. (2022). Modulo cellulo: modular versatile tangible educational robots. *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 3666–3671. https://doi.org/10.1109/IROS47612.2022.9981983.

(Cited on page 131)

Khodr, H., Kianzad, S., Johal, W., Kothiyal, A., Bruno, B., & Dillenbourg, P. (2020). Allohaptic: robot-mediated haptic collaboration for learning linear functions. *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 27–34.

(Cited on pages 17, 115)

Khodr, H., Kothiyal, A., Bruno, B., & Dillenbourg, P. (2022). An assessment framework for complex systems understanding. In C. Chinn, E. Tan, C. Chan, & Y. Kali (Eds.), *Proceedings of the 16th international conference on learning sciences - ICLS* (pp. 99–106).

(Cited on page 33)

Khodr, H., Ramage, U., Kim, K., Guneysu Ozgur, A., Bruno, B., & Dillenbourg, P. (2020). Being part of the swarm: experiencing human-swarm interaction with vr and tangible robots. *Symposium on Spatial User Interaction*. https://doi.org/10.1145/3385959.3422695.

(Cited on page 143)

Khodr, H., Wagner, N., Bruno, B., Kothiyal, A., & Dillenbourg, P. (2022). Effect of different communication affordances on the emergence of collaboration strategies in an online multiplayer game. *Swarm Intelligence: 13th International Conference, ANTS 2022, Málaga, Spain, November 2–4, 2022, Proceedings*, 316–323.

(Cited on page 115)

Kolling, A., Walker, P., Chakraborty, N., Sycara, K., & Lewis, M. (2016). Human interaction with robot swarms: a survey. *IEEE Transactions on Human-Machine Systems*, *46*(1), 9–26. https://doi.org/10.1109/THMS.2015.2480801.

(Cited on page 103)

Kumar, V., Tissenbaum, M. B., & Kim, T. (2021). Procedural collaboration in educational games: supporting complex system understandings in immersive whole class simulations. *Communication Studies*, *72*(6), 994–1016.

(Cited on page 21)

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

(Cited on page 45)

Langbehn, E., Lubos, P., & Steinicke, F. (2018). Evaluation of Locomotion Techniques for Room-Scale VR: Joystick, Teleportation, and Redirected Walking, 9.

(Cited on page 155)

Li, Y., Liang, M., Preissing, J., Bachl, N., Dutoit, M. M., Weber, T., Mayer, S., & Hussmann, H. (2022). A meta-analysis of tangible learning studies from the tei conference. *Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction*. https://doi.org/10.1145/3490149.3501313.

(Cited on pages 24, 82)

Loibl, K., & Rummel, N. (2014). The impact of guidance during problem-solving prior to instruction on students' inventions and learning outcomes. *Instructional Science*, *42*, 305–326.

(Cited on page 79)

Lu, Q., Ma, S.-e., Li, J., Mi, H., & Xu, Y. (2019). Irelics: designing a tangible interaction platform for the popularization of field archaeology. *Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction*, 45–54. https://doi.org/10.1145/3294109.3295647.

(Cited on page 82)

Lux, J.-D., & Budke, A. (2020). Playing with Complex Systems? The Potential to Gain Geographical System Competence through Digital Gaming [Number: 5 Publisher: Multidisciplinary Digital Publishing Institute]. *Education Sciences*, *10*(5), 130. https://doi.org/10.3390/educsci10050130.

(Cited on page 21)

Majumdar, R., Kothiyal, A., Ranka, A., Pande, P., Murthy, S., Agarwal, H., & Chandrasekharan, S. (2014). The enactive equation: exploring how multiple external representations are integrated, using a fully controllable interface and eye-tracking. *2014 IEEE sixth international conference on Technology for Education*, 233–240.

(Cited on page 106)

Marlow, S. L., Lacerenza, C. N., Paoletti, J., Burke, C. S., & Salas, E. (2018). Does Team Communication Represent a One-size-fits-all Approach?: A Meta-analysis of Team Communication and Performance. *Organizational Behavior and Human Decision Processes*, *144*, 145–170. https://doi.org/10.1016/j.obhdp.2017.08.001.

(Cited on page 116)

Marshall, P. (2007). Do Tangible Interfaces Enhance Learning? *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, 163–170. https://doi.org/10.1145/1226969.1227004.

(Cited on pages 24, 82)

Mataric, M. J., Koenig, N. P., & Feil-Seifer, D. (2007). Materials for enabling hands-on robotics and stem education. *AAAI spring symposium: Semantic scientific knowledge integration*, 99–102.

(Cited on page 7)

Miglino, O., Lund, H. H., & Cardaci, M. (1999). Robotics as an educational tool. *Journal of Interactive Learning Research*, *10*(1), 25–47.

(Cited on page 7)

Mitchell, M. (2009). *Complexity: a guided tour*. Oxford university press.

(Cited on page 2)

Mondada, F., Bonani, M., Raemy, X., Pugh, J., Cianci, C., Klaptocz, A., Magnenat, S., Zufferey, J.-C., Floreano, D., & Martinoli, A. (2009). The e-puck, a Robot Designed for Education in Engineering. Retrieved October 30, 2019, from https://infoscience.epfl.ch/record/135236.

(Cited on page 12)

Mussati, A., Giang, C., Piatti, A., & Mondada, F. (2019). A tangible programming language for the educational robot thymio. *[Proceedings of the IISA 2019 The 10th International Conference on Information, Intelligence, Systems and Applications]*. http://infoscience.epfl.ch/record/269211.

(Cited on page 24)

Nasir, J., Bruno, B., Chetouani, M., & Dillenbourg, P. (2021). What if social robots look for productive engagement? automated assessment of goal-centric engagement in learning applications. *International Journal of Social Robotics*, 1–17.

(Cited on page 71)

Nasir, J., Kothiyal, A., Bruno, B., & Dillenbourg, P. (2021). Many are the ways to learn identifying multi-modal behavioral profiles of collaborative learning in constructivist activities. *International Journal of Computer-Supported Collaborative Learning*, *16*(4), 485–523.

(Cited on page 79)

Nasir, J., Norman, U., Johal, W., Olsen, J. K., Shahmoradi, S., & Dillenbourg, P. (2019). Robot analytics: what do human-robot interaction traces tell us about learning? *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 1–7. https://doi.org/10.1109/RO-MAN46459.2019.8956465.

(Cited on page 17)

Nedjah, N., & Junior, L. S. (2019). Review of methodologies and tasks in swarm robotics towards standardization. *Swarm and Evolutionary Computation*, *50*, 100565.

(Cited on page 11)

Nölle, J., Fusaroli, R., Mills, G. J., & Tylén, K. (2020). Language as Shaped by the Environment:

Linguistic Construal in a Collaborative Spatial Task. *Palgrave Communications*, *6*(1), 1–10.

(Cited on page 116)

O'Bryan, L., Beier, M., & Salas, E. (2020). How Approaches to Animal Swarm Intelligence Can Improve the Study of Collective Intelligence in Human Teams. *Journal of Intelligence*, *8*(1). https://doi.org/10.3390/jintelligence8010009.

(Cited on page 116)

Özdemir, G., & Clark, D. B. (2007). An Overview of Conceptual Change Theories. *Conceptual Change*.

(Cited on pages 25, 27)

Ozgur, A. G., Khodr, H., Akeddar, M., Roust, M., & Dillenbourg, P. (2022). Designing online multiplayer games with haptically and virtually linked tangible robots to enhance social interaction in therapy. *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 358–364. https://doi.org/10.1109/RO-MAN53752.2022.9900684.

(Cited on page 143)

Özgür, A. (2018). *Cellulo: Tangible Haptic Swarm Robots for Learning* (Doctoral dissertation). IINFCOM. EPFL. https://doi.org/10.5075/epfl-thesis-8241.

(Cited on pages 14, 16, 17, 131)

Özgür, A., Johal, W., & Dillenbourg, P. (2016). Permanent magnet-assisted omnidirectional ball drive. *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 1061–1066.

(Cited on page 138)

Özgür, A., Johal, W., Mondada, F., & Dillenbourg, P. (2017). Windfield: learning wind meteorology with handheld haptic robots. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 156–165.

(Cited on page 17)

Özgür, A., Lemaignan, S., Johal, W., Beltran, M., Briod, M., Pereyre, L., Mondada, F., & Dillenbourg, P. (2017). Cellulo: versatile handheld robots for education. *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*, 119–127.

(Cited on page 8)

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

(Cited on pages 3, 6, 18)

Pares, N., & Eisenberg, M. (2022). Tangible and full-body interfaces in learning. In R. K. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (3rd ed., pp. 321–339). Cambridge University Press. https://doi.org/10.1017/9781108888295.020.

(Cited on page 23)

Pezzulo, G., Roche, L., & Saint-Bauzel, L. (2021). Haptic Communication Optimises Joint Decisions and Affords Implicit Confidence Sharing. *Scientific Reports*, *11*(1), 1–9.

(Cited on page 115)

Piaget, J. (1962). *Play, dreams and imitation in childhood* (Vol. 25). Routledge.

(Cited on pages 22, 25)

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational psychologist*, *50*(4), 258–283.

(Cited on page 22)

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. (1982). Toward a theory of conceptual change. *Science education*, *66*(2), 211–227.

(Cited on page 26)

Randall, A., Klingner, J., & Correll, N. (2016). Simulating Chemical Reactions Using a Swarm of Miniature Robots. In E. Tuci, A. Giagkos, M. Wilson, & J. Hallam (Eds.), *From Animals to Animats 14* (pp. 305–316). Springer International Publishing.

(Cited on page 14)

Rates, C. A., Mulvey, B. K., Chiu, J. L., & Stenger, K. (2022). Examining ontological and self-monitoring scaffolding to improve complex systems thinking with a participatory simulation. *Instructional Science*, 1–23.

(Cited on page 87)

Reed, S. K. (2020). *Cognitive skills you need for the 21st century*. Oxford University Press, USA.

(Cited on pages 3, 87)

Repenning, A., & Ioannidou, A. (2006). Agentcubes: raising the ceiling of end-user development in education through incremental 3d. *Visual Languages and Human-Centric Computing (VL/HCC'06)*, 27–34. https://doi.org/10.1109/VLHCC.2006.7.

(Cited on page 18)

Repenning, A., & Sumner, T. (1995). Agentsheets: a medium for creating domain-oriented visual languages. *Computer*, *28*(3), 17–25. https://doi.org/10.1109/2.366152.

(Cited on page 18)

Repenning, A. (2016). Transforming" hard and boring" into" accessible and exciting". *CoPDA@ NordiCHI*, 25–34.

(Cited on page 18)

Repenning, A., Ioannidou, A., Luhn, L., Daetwyler, C., & Repenning, N. (2010). Mr. vetro: assessing a collective simulation framework. *Journal of Interactive Learning Research*, *21*(4), 515–537.

(Cited on page 19)

Resnick, M. (1996). Starlogo: an environment for decentralized modeling and decentralized thinking. *Conference companion on Human factors in computing systems*, 11–12.

(Cited on page 18)

Resnick, M. (1997). *Turtles, termites, and traffic jams: explorations in massively parallel microworlds*. Mit Press.

(Cited on pages 6, 19, 34, 35, 42, 43)

Resnick, M., & Wilensky, U. (1998). Diving into Complexity: Developing Probabilistic Decentralized Thinking Through Role-playing Activities. *The Journal of the Learning Sciences*, *7*(2), 153–172.

(Cited on page 6)

Reynolds, C. W. (1987). Flocks, herds and schools: a distributed behavioral model. *Proceedings* *of the 14th Annual Conference on Computer Graphics and Interactive Techniques*, 25–34. https://doi.org/10.1145/37401.37406.

(Cited on page 154)

Rogers, C., & Portsmore, M. (2004). Bringing engineering to elementary school. *Journal of STEM Education: innovations and research*, *5*(3).

(Cited on page 7)

Rotgans, J. I., & Schmidt, H. G. (2011). Situational interest and academic achievement in the active-learning classroom. *Learning and Instruction*, *21*, 58–67.

(Cited on page 22)

Rubens, C., Braley, S., Torpegaard, J., Lind, N., Vertegaal, R., & Merritt, T. (2020). Flying lego bricks: observations of children constructing and playing with programmable matter. *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction*, 193–205. https://doi.org/10.1145/3374920.3374948.

(Cited on page 82)

Rubenstein, M., Ahler, C., Hoff, N., Cabrera, A., & Nagpal, R. (2014). Kilobot: a low cost robot with scalable operations designed for collective behaviors [Reconfigurable Modular Robotics]. *Robotics and Autonomous Systems*, *62*(7), 966–975. https://doi.org/https://doi.org/10.1016/j.robot.2013.08.006.

(Cited on pages 12, 13)

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, *55*(1), 68.

(Cited on page 82)

Ryan, R. M., & Deci, E. L. (2001). Intrinsic motivation inventory. *Recuperado de https://selfdeterminationtheory. org/intrinsic-motivation-inventory*.

(Cited on page 79)

Schelling, T. C. (2006). *Micromotives and macrobehavior*. WW Norton & Company.

(Cited on page 17)

Schneider, B., Jermann, P., Zufferey, G., & Dillenbourg, P. (2011). Benefits of a Tangible Interface for Collaborative Learning and Interaction. *IEEE Transactions on Learning Technologies*, *4*(3), 222–232. https://doi.org/10.1109/TLT.2010.36.

(Cited on pages 24, 79, 85)

Schwartz, D. L., Martin, T., & Nasir, N. (2005). Designs for knowledge evolution: towards a prescriptive theory for integrating first-and second-hand knowledge. In P. Gardenfors, P. Johansson, & N. J. Mahwah (Eds.), *Cognition, education, and communication technology* (pp. 21–54). Erlbaum Associates.

(Cited on pages 23, 112)

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. S. (2009). Signalling Signalhood and the Emergence of Communication. *Cognition*, *113*(2), 226–233. https://doi.org/10.1016/j.cognition.2009.08.009.

(Cited on page 116)

Sebo, S., Stoll, B., Scassellati, B., & Jung, M. F. (2020). Robots in groups and teams: a literature review. *Proc. ACM Hum.-Comput. Interact.*, *4*(CSCW2). https://doi.org/10.1145/3415247.

(Cited on page 103)

Seeley, T. D., Visscher, P. K., Schlegel, T., Hogan, P. M., Franks, N. R., & Marshall, J. A. (2012). Stop signals provide cross inhibition in collective decision-making by honeybee swarms. *Science*, *335*(6064), 108–111.

(Cited on page 15)

Selten, R., & Warglien, M. (2007). The Emergence of Simple Languages in an Experimental Coordination Game. *Proceedings of the National Academy of Sciences*, *104*(18), 7361–7366. https://doi.org/10.1073/pnas.0702077104.

(Cited on page 116)

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel psychology*, *64*(2), 489–528.

(Cited on page 22)

Steinkuehler, C., & Squire, K. (2022). Video games and learning. In R. K. Sawyer (Ed.), *The cambridge handbook of the learning sciences* (3rd ed., pp. 281–300). Cambridge University Press. https://doi.org/10.1017/9781108888295.018.

(Cited on page 22)

Svennebring, J., & Koenig, S. (2004). Building Terrain-covering Ant Robots: A Feasibility Study. *Autonomous Robots*, *16*(3), 313–332.

(Cited on page 116)

Talamali, M. S., Saha, A., Marshall, J. A. R., & Reina, A. (2021). When less is more: robot swarms adapt better to changes with constrained communication. *Science Robotics*, *6*(56), eabf1416. https://doi.org/10.1126/scirobotics.abf1416.

(Cited on pages 120, 121)

Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using Presence Questionnaires in Reality. *Presence: Teleoperators and Virtual Environments*, *9*(5), 497–503. https://doi.org/10.1162/105474600566989.

(Cited on page 155)

Vitanza, A., Rossetti, P., Mondada, F., & Trianni, V. (2019). Robot swarms as an educational tool: The Thymio's way. *International Journal of Advanced Robotic Systems*. https://doi.org/10.1177/1729881418825186.

(Cited on page 15)

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: a meta-analysis. *Journal of Educational Computing Research*, *34*(3), 229–243. https://doi.org/10.2190/FLHV-K4WA-WPVQ-H0YM.

(Cited on page 22)

Vosniadou, S., & Skopeliti, I. (2014). Conceptual change from the framework theory side of the fence. *Science & Education*, *23*, 1427–1445.

(Cited on page 26)

Wang, J., Chellali, A., & Cao, C. G. (2013). A Study of Communication Modalities in a Virtual Collaborative Task. *2013 IEEE International Conference on Systems, Man, and Cybernetics*, 542–546. https://doi.org/10.1109/SMC.2013.98.

(Cited on page 115)

Wilensky, U. (1999). Netlogo. http://ccl.northwestern.edu/netlogo/.

(Cited on page 18)

Wilensky, U., & Abrahamson, D. (2006). Is a disease like a lottery?: classroom networked technology that enables student reasoning about complexity. *annual meeting of the American Educational Research Association, San Francisco, CA*.

(Cited on pages 34, 35, 109)

Wilensky, U., & Rand, W. (2015). *An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with netlogo*. Mit Press.

(Cited on page 18)

Wilensky, U., & Stroup, W. (1999). Hubnet. http://ccl.northwestern.edu/netlogo/hubnet.html.

(Cited on page 19)

Wilkerson-Jerde, M. H., & Wilensky, U. J. (2015). Patterns, probabilities, and people: making sense of quantitative change in complex systems. *Journal of the Learning Sciences*, *24*(2), 204–251.

(Cited on pages 4, 34)

Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of educational psychology*, *105*(2), 249.

(Cited on page 23)

Yoon, S. A., Goh, S.-E., & Park, M. (2018). Teaching and Learning About Complex Systems in K–12 Science Education: A Review of Empirical Studies 1995–2015. *Review of Educational Research*, *88*(2), 285–325. https://doi.org/10.3102/0034654317746090.

(Cited on page 87)

Yoon, S. A., Goh, S.-E., & Yang, Z. (2019). Toward a learning progression of complex systems understanding. *Complicity: An International Journal of Complexity and Education*, *16*(1), 1–19.

(Cited on pages 6, 34, 35, 38, 42, 87)

Young, M. F., Slota, S., Cutter, A. B., Jalette, G., Mullin, G., Lai, B., Simeoni, Z., Tran, M., & Yukhymenko, M. (2012). Our princess is in another castle: a review of trends in serious gaming for education. *Review of educational research*, *82*(1), 61–89.

(Cited on page 23)

Yu, J., Han, S. D., Tang, W. N., & Rus, D. (2017). A portable, 3d-printing enabled multi-vehicle platform for robotics research and education. *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 1475–1480. https://doi.org/10.1109/ICRA.2017.7989176.

(Cited on page 12)

Yuan, Y., Cao, J., Wang, R., & Yarosh, S. (2021). Tabletop games in the age of remote collaboration: design opportunities for a socially connected game experience. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–14.

(Cited on page 151)

HALA KHODR

St-Sulpice, Vaud, Switzerland github.com/hkhodr Hala Khodr hala-khodr

EDUCATION

Swiss Federal Institute of Technology, Lausanne (EPFL)
PhD in Robotics, Control and Intelligent Systems — Dec 2018 – June 2023

Swiss Federal Institute of Technology, Lausanne (EPFL)
Master's Degree in Micro-engineering with a specialization in Robotics — Sept 2016 – Sept 2018
- GPA: 5.6/6.0
- Master Thesis: Grade: 6/6

American University of Beirut, Lebanon
Bachelor Degree in Electrical and Computer Engineering — Sept 2012 – June 2016
- GPA: 4.0/4.0
- Minor in Mathematics

CORE EXPERIENCE

Computer Human Interaction for Learning and Instruction Lab - Research Assistant

EPFL, Switzerland — Dec 2018 – Present

Human-Robot Swarm Interaction: An Exploratorive Path to Foster Complex Systems Understanding (PhD Thesis)
- Enhanced a tangible robotic platform from hardware and software perspectives.
- Designed, validated and analyzed gamified learning activities under the topic of swarm behaviors and complex systems understanding.
- Conducted experiments with more than 150 participants aged 12-45 years old.

BioRobotics Lab - Research Assistant

EPFL, Switzerland — May 2018 – Sept 2018

Roombots: Self-reconfigurable Modular Robots for Adaptive and Self-organizing Furniture (Master Thesis)
Developed an optimal planning framework for motion and shape formation and deployed it on real robotic hardware.

Synaestech - Embedded Systems Engineer

Lausanne, Switzerland — Jan 2018 – May 2018

Contributed to the design, development and testing of a wireless sensor network for a sensing platform for real-time monitoring of natural and engineered systems.

Distributed Intelligent Systems Lab - Research Assistant

EPFL, Switzerland — Feb 2017 – Sept 2017

Project: Fluid-mediated Self-assembly of Robotic Modules.
- Developed a higher-accuracy method for calibrating stochastic motion models of water-floating robots.
- Designed optimized ruleset controllers for given target structures of the robotic modules.

ADDITIONAL EXPERIENCE

Teaching Assistant

EPFL, Switzerland — Sept 2017 – Sept 2022
- Supervision of 3-4 Masters and Bachelor students per term
- Human Computer Interaction (Bachelor level), Spring 2022, Head TA.
- Introduction to Programming (Bachelor Level), Fall 2021
- Analysis I (Bachelor Level), Fall 2019, Fall 2020 (Online)
- Robotics Practicals (Masters Level), Spring 2019, Spring 2020
- Coach for Robotics Competition (Masters Level), Fall 2021
- Machine Learning Programming (Masters Level), Fall 2017

AUB, Lebanon — Sept 2015 – Jan 2016

Electronic Circuits Course (Bachelor Level)

Department of Electrical and Computer Engineer - Research Assistant

American University of Beirut (AUB), Lebanon — May 2016 – Sept 2016

Project: Energy Efficient IoT Sensor with RF Wake-up and Addressing Capability (Bachelor thesis)
Designed and implemented a complete novel design for an optimized energy-efficient IoT sensor based on the concept of radio wake-up.

Stanford Khuri-Yakub Ultrasonics Group - Research Intern

Stanford, USA — June 2015 – Aug 2015

Worked on the circuit design and simulation of a thermal noise cancelling amplifier.

Co-founder of AUB Robotics Club

American University of Beirut (AUB), Lebanon — June 2015 – Sept 2016

Established an active technical club in AUB, giving workshops and organizing engineering design challenges.

SKILLS

Programming Languages: C, C++, C#, Python, Matlab.

Robotics: Human-Robot Interaction (HRI), Control systems, Motion planning, Swarm robots, Robot Operating System (ROS), Webots

Embedded Systems: PIC32, STM32 microcontrollers, Bluetooth, Sensor integration.

Electronics Design: Low power electronics, Circuit design with Altium and Eagle.

Rapid Prototyping: Fusion 360 (CAD), 3D printing.

Game Design and Development: Unity3D, Qt, Augmented/Mixed Reality applications.

Internet of Things (IoT): Energy-efficient sensor design, Radio wake-up circuits, Data streaming, LoRa communication.

Machine Learning: Python libraries (Pandas, Scikit-learn, Pytorch, Keras), Artificial Neural Networks, Supervised and Unsupervised Learning techniques.

Project Management: Coordination of cross-functional teams, student supervision, collaborative projects.

Science Communication: Academic writing (conference and journal papers), oral presentations (conferences and public talks).

LANGUAGES

Arabic Native
English Fluent spoken and written
French Fluent spoken and written
German Beginner

AWARDS

- Teaching Assistant Award given for teaching excellence in the School of Computer and Communication Sciences at EPFL (2022).
- NCCR Robotics Tool Award presented to the CHILI lab for the development of the robot Cellulo (2022).
- NCCR Robotics Equal Opportunities Award for Career Development to attend the IEEE RAS Summer School on Multi-Robot Systems (2020).
- Distinguished Graduate Award given to a graduating senior student who demonstrates high academic achievement (2016).
- American University of Beirut - Stanford University Summer Research Internship Fellowship (2015).
- Full scholarship from the Lebanese National Council for Scientific Research for ranking 4^{th} in the National Baccalaureate (2012).

GRANTS

- Principal Investigator of NCCR Grassroot project (granted with 24,000 CHF) - a year-long collaboration involved 3 researchers, 1 engineer, and 5 undergraduate and graduate students.
- Main organizer of EPFL-ETHZ 2022 Open Science Summer School (granted with 20,000 CHF) - a week-long event featured presentations from esteemed external speakers and fostered the participation of 30 PhD students.

PUBLIC OUTREACH

- **Demos** in multiple public events such as NCCR robotics days, EPFL Open days and JOM (Journée Oser tous les métiers - day to dare all professions for kids).
- **Speaker** in a round table: "La robotique en tous genres" that aims to attract young students, and female students in particular, towards scientific and technical subjects.

PUBLICATIONS

- Khodr, Hala, Barbara Bruno, et al. (2022). "Cellulan World: Interactive platform to learn swarm behaviors". In: *AAMAS'22: Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*. CONF. IFAAMAS, pp. 1908–1910.
- Khodr, Hala, Kevin Holdcroft, et al. (2022). "Modulo Cellulo: Modular Versatile Tangible Educational Robots". In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 3666–3671.
- Khodr, Hala, Aditi Kothiyal, et al. (2022). "An Assessment Framework for Complex Systems Understanding". In: *Proceedings of the 16th international conference on Learning sciences*.
- Khodr, Hala, Nicolas Wagner, et al. (2022). "Effect of Different Communication Affordances on the Emergence of Collaboration Strategies in an Online Multiplayer Game". In: *International Conference on Swarm Intelligence*. Springer International Publishing Cham, pp. 316–323.
- Ozgur, Arzu Guneysu, Hala Khodr, Mehdi Akeddar, et al. (2022). "Designing Online Multiplayer Games with Haptically and Virtually Linked Tangible Robots to Enhance Social Interaction in Therapy". In: *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, pp. 358–364.
- Khodr, Hala, Jerome Brender, et al. (2021). "How Diseases Spread: Embodied Learning of Emergence with Cellulo Robots". In: *Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education*. CONF. ASIA PACIFIC SOC COMPUTERS IN EDUCATION.
- Ozgur, Arzu Guneysu, Hala Khodr, Barbara Bruno, et al. (2021). "Detecting Compensatory Motions and Providing Informative Feedback During a Tangible Robot Assisted Game for Post-Stroke Rehabilitation". In: *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE, pp. 243–249.
- Hauser, Simon et al. (2020). "Roombots extended: Challenges in the next generation of self-reconfigurable modular robots and their application in adaptive and assistive furniture". In: *Robotics and Autonomous Systems* 127.
- Khodr, Hala, Soheil Kianzad, et al. (2020). "Allohaptic: Robot-mediated haptic collaboration for learning linear functions". In: *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, pp. 27–34.
- Khodr, Hala, Ulysse Ramage, et al. (2020). "Being part of the swarm: Experiencing human-swarm interaction with vr and tangible robots". In: *Proceedings of the 2020 ACM Symposium on Spatial User Interaction*, pp. 1–2.
- Haghighat, Bahar, Hala Khodr, and Alcherio Martinoli (2019a). "Design and calibration of a lightweight physics-based model for fluid-mediated self-assembly of robotic modules". In: *Distributed Autonomous Robotic Systems: The 14th International Symposium*. Springer International Publishing, pp. 197–210.
- – (2019b). "Lightweight physics-based models for the control of fluid-mediated self-assembly of robotic modules". In: *Robotics and Autonomous Systems* 121, p. 103241.
- Johal, Wafa et al. (2019). "Tip: Tangible e-ink paper manipulatives for classroom orchestration". In: *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, pp. 595–598.
- Khodr, Hala, Mehmet Mutlu, et al. (2019). "An optimal planning framework to deploy self-reconfigurable modular robots". In: *IEEE Robotics and Automation Letters* 4.4, pp. 4278–4285.
- Khodr, Hala, Nour Kouzayha, et al. (2017). "Energy efficient IoT sensor with RF wake-up and addressing capability". In: *IEEE sensors letters* 1.6, pp. 1–4.